

# *Open Journal of Mathematical Optimization*

Anatoli Juditsky, Georgios Kotsalis & Arkadi Nemirovski

**Tight computationally efficient approximation of matrix norms with applications**

Volume 3 (2022), article no. 7 (38 pages)

<https://doi.org/10.5802/ojmo.19>

Article submitted on January 5, 2022, revised on November 8, 2022,  
accepted on November 8, 2022.

© The author(s), 2022.



This article is licensed under the

CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.

<http://creativecommons.org/licenses/by/4.0/>



# Tight computationally efficient approximation of matrix norms with applications

Anatoli Juditsky

LJK, Université Grenoble Alpes, 700 Avenue Centrale  
38401 Domaine Universitaire de Saint-Martin-d'Hères, France  
anatoli.juditsky@univ-grenoble-alpes.fr

Georgios Kotsalis

Georgia Institute of Technology, Atlanta, Georgia 30332, USA  
gkotsalis3@gatech.edu

Arkadi Nemirovski

Georgia Institute of Technology, Atlanta, Georgia 30332, USA  
nemirovs@isye.gatech.edu

---

## Abstract

We address the problems of computing operator norms of matrices induced by given norms on the argument and the image space. It is known that aside of a fistful of “solvable cases”, most notably, the case when both given norms are Euclidean, computing operator norm of a matrix is NP-hard. We specify rather general families of norms on the argument and the images space (“elliptic” and “co-elliptic”, respectively) allowing for reasonably tight computationally efficient upper-bounding of the associated operator norms. We extend these results to bounding “robust operator norm of uncertain matrix with box uncertainty”, that is, the maximum of operator norms of matrices representable as a linear combination, with coefficients of magnitude  $\leq 1$ , of a collection of given matrices. Finally, we consider some applications of norm bounding, in particular, (1) computationally efficient synthesis of affine non-anticipative finite-horizon control of discrete time linear dynamical systems under bounds on the peak-to-peak gains, (2) signal recovery with uncertainties in sensing matrix, and (3) identification of parameters of time invariant discrete time linear dynamical systems via noisy observations of states and inputs on a given time horizon, in the case of “uncertain-but-bounded” noise varying in a box.

Digital Object Identifier 10.5802/ojmo.19

Keywords matrix norms, uncertain matrices, robust control, system identification.

Acknowledgments The first author was supported by MIAI @ Grenoble Alpes (ANR-19-P3IA-0003). Research of the second author was supported by NIFA grant 2020-67021-31526.

## 1 Introduction

In this paper, our theoretical focus is on two problems as follows:

► **Problem A** (Approximating operator norms). *Given norms  $\|\cdot\|_{\mathcal{X}}$  and  $\|\cdot\|_{\mathcal{B}}$  with unit balls  $\mathcal{X} \subset \mathbf{R}^n$  and  $\mathcal{B} \subset \mathbf{R}^m$ , estimate the induced norm  $\|A\|_{\mathcal{B},\mathcal{X}} := \max_{x:\|x\|_{\mathcal{X}} \leq 1} \|Ax\|_{\mathcal{B}}$  of an  $m \times n$  matrix  $A$*

► **Problem B** (Approximating robust norm of uncertain matrix with box uncertainty). *With  $\|\cdot\|_{\mathcal{X}}$ ,  $\|\cdot\|_{\mathcal{B}}$  as in A, given an “uncertain  $m \times n$  matrix with box uncertainty”, set of the form*

$$\mathcal{A} = \left\{ A_{\text{nom}} + \sum_{s=1}^S \epsilon_s A_s : \|\epsilon\|_{\infty} \leq 1 \right\} \quad (A_{\text{nom}}, A_1, \dots, A_S \in \mathbf{R}^{m \times n}),$$

estimate the robust norm

$$\|\mathcal{A}\|_{\mathcal{B},\mathcal{X}} = \max_{A \in \mathcal{A}} \|A\|_{\mathcal{B},\mathcal{X}}.$$

of the uncertain matrix  $A$ .



© Anatoli Juditsky & Georgios Kotsalis & Arkadi Nemirovski;  
licensed under Creative Commons License Attribution 4.0 International

Applications motivating our interest in these problems will be discussed later; we start with outlining the research status of these problems as “academic entities” and our related results.

Aside of few special cases, e.g., the case of the spectral norm ( $\mathcal{X}$  and  $\mathcal{B}$  are unit Euclidean balls in the respective spaces),  $\mathbf{A}$  is NP-hard; this is so, e.g., when  $\|\cdot\|_{\mathcal{X}} = \|\cdot\|_p$ ,  $\|\cdot\|_{\mathcal{B}} = \|\cdot\|_r$ , and  $p \geq 2 \geq r \geq 1$  with  $p \neq r$  [41]. (12) is NP-hard already when  $\mathcal{B}$ ,  $\mathcal{X}$  are unit Euclidean balls,  $A_{\text{nom}} = 0$ , and  $A_s$  are restricted to be symmetric matrices of rank 2 [7]. Hardness of  $\mathbf{A}$ ,  $\mathbf{B}$  makes it natural to look for efficiently computable reasonably tight upper bounds on the norms in question. Below we build these bounds for the case where  $\mathcal{X}$  and *the polar*  $\mathcal{B}_*$  of  $\mathcal{B}$  are *ellitopes* (see Section 2.1 for the corresponding definitions).

Sufficient for our current purposes example of an ellitope in  $\mathbf{R}^k$  is a bounded set  $\mathcal{Z}$  cut of  $\mathbf{R}^k$  by a convex constraint on the vector  $[z^T P_1 z; \dots; z^T P_J z]$  of values of convex homogeneous quadratic forms of  $z$ :

$$\mathcal{Z} = \{z \in \mathbf{R}^k : \exists t \in \mathcal{T} : z^T P_j z \leq t_j, j \leq J\}$$

where  $P_j \succeq 0$ ,  $\sum_j P_j \succ 0$ , and  $\mathcal{T}$  is a convex compact subset of  $\mathbf{R}_+^J$  with a nonempty interior which is monotone, i.e.,  $0 \leq t' \leq t \in \mathcal{T}$  implies that  $t' \in \mathcal{T}$ . A simple example of the ellitope is the intersection of finitely many ellipsoids/elliptic cylinders centered at the origin.

We demonstrate that in the ellitopic case one can build efficiently computable upper bounds  $\Phi(A)$  on  $\|A\|_{\mathcal{B}, \mathcal{X}}$  and  $\Psi(A_1, \dots, A_N)$  on  $\|\mathcal{A}\|_{\mathcal{B}, \mathcal{X}}$  which are convex in  $A$ , resp., in  $(A_1, \dots, A_N)$ , such that

$$\|A\|_{\mathcal{B}, \mathcal{X}} \leq \Phi(A) \leq O(1) \sqrt{\ln(2K) \ln(2L)} \|A\|_{\mathcal{B}, \mathcal{X}}, \quad (1a)$$

$$\|\mathcal{A}\|_{\mathcal{B}, \mathcal{X}} \leq \Psi(A_1, \dots, A_N) \leq O(1) \sqrt{\ln(2K) \ln(2L)} \vartheta(\kappa) \|\mathcal{A}\|_{\mathcal{B}, \mathcal{X}} \quad (1b)$$

where  $K$  and  $L$  are *ellitopic sizes* (numbers of quadratic forms in the description) of  $\mathcal{X}$  and  $\mathcal{B}_*$ ,  $\kappa$  is the maximum of ranks of  $A_i$ , and  $\vartheta(\cdot)$  is a certain universal function of  $\kappa$ .

### Relation to existing literature, Problem A

$\mathbf{A}$  is the problem of maximizing a quadratic (specifically, bilinear) form on  $\mathcal{B}_* \times \mathcal{X}$ , and there exists significant literature on tractable relaxations, semidefinite and alike, of these problems. To the best of our knowledge, the most advanced existing results are those in the seminal papers [34, 35] of Yu. Nesterov. As applied to  $\mathbf{A}$ , those results, in our present language, state that when the positive semidefinite matrices participating in description of  $\mathcal{X}$  and  $\mathcal{B}_*$  are diagonal, the appropriate efficiently computable relaxation bound on  $\|A\|_{\mathcal{B}, \mathcal{X}}$  (which in fact is nothing but the bound  $\Phi$  participating in (1a)) is tight within *absolute constant factor* (for details, see Remark 4). It should be stressed that “tightness within an absolute constant” heavily exploits diagonality of the matrices describing  $\mathcal{X}$  and  $\mathcal{B}_*$ ; in the case of general ellitopes, logarithmic tightness factors in (1a) seem to be unavoidable.<sup>1</sup>

The results on tight computationally tractable upper-bounding of maxima of quadratic forms over general-type ellitopes (same as the notion of an ellitope itself) originate from [20] and are further developed in [21]. As compared to those results, dealing with bilinear rather than with general quadratic forms allows us below to refine the analysis, and, as a result, to reduce the tightness factor in (1a) to  $O(1) \sqrt{\ln(2K) \ln(2L)}$  instead of  $O(1) \ln(K + L)$  guaranteed by [21].

### Relation to existing literature, Problem B

The only known to us preceding results on bounding robust norms of uncertain matrices deal with the spectral norm ( $\mathcal{X}$  and  $\mathcal{B}$  are unit Euclidean balls), in which case the tightness factor in (1b) boils down to  $\vartheta(2\kappa)$ ; these results can be easily derived from the “Matrix Cube Theorem” in [7].

## Applications

While  $\mathbf{A}$  and  $\mathbf{B}$  look as legitimate academic problems, and the outlined results—legitimate academic results, the actual motivation for what follows stems from specific applications of problems  $\mathbf{A}$  and  $\mathbf{B}$  we are about to consider.

<sup>1</sup> For instance, it was shown in [33] that when  $\|\cdot\|_{\mathcal{B}} = \|\cdot\|_2$  and  $\mathcal{X}$  is the intersection of  $K$  “stripes” centered at the origin (i.e., the corresponding positive semidefinite matrices are of rank 1), the relaxation bounds in question can indeed be larger than the true quantity by factor  $O(\sqrt{\ln K})$ .

Our principal motivation for Problem A comes from control and is the necessity to handle *peak-to-peak* design specifications in synthesis of linear controllers. Specifically, given a linear dynamical system

$$x_{t+1} = A_t x_t + B_t u_t + D_t d_t, \quad x_0 = z, \quad y_t = C_t x_t + E_t d_t$$

with states  $x_t$ , controls  $u_t$ , observed outputs  $y_t$ , and external disturbances  $d_t$ , we want to build an affine non-anticipating controller  $u_t = g_t + \sum_{\tau=0}^t G_\tau^t y_\tau$  in such a way that the trajectory  $w^N = \{x_t, 1 \leq t \leq N; y_t, u_t, 0 \leq t < N\}$  of the closed loop system on a given time horizon satisfies a given set of design specifications. With smart nonlinear reparameterization of affine non-anticipating controllers (passing from affine output-based control to the control which is affine in *purified outputs*, see [22] and references therein), the system trajectory becomes affine function of the initial state  $z$  and the sequence  $d^N = [d_0; \dots; d_{N-1}]$  of external disturbances, with the matrices and constant terms in these functions being *affine in the vector  $\chi$  of controller's parameters* varying in certain  $\mathbf{R}^\nu$ . Bi-affinity of  $w^N$  in  $(d^N, z)$  and in  $\chi$  is the key to computationally efficient processing of design specifications of appropriate structure. In this paper, we address an important (and considered as difficult in control) specification, namely, *peak-to-peak gain* defined as follows.<sup>2</sup> Let us fix some a norm  $\|\cdot\|_{(d)}$  on the space where the disturbances  $d_t$  live, and norm  $\|\cdot\|_{(x)}$  on the space where the states  $x_t$  live. We equip the space  $D^N$  of disturbance sequences  $d^N = [d_0; \dots; d_{N-1}]$  with the norm  $\|d^N\|_{d,\infty} = \max_t \|d_t\|_{(d)}$ , and the space  $X_d^N$  of state trajectories  $x^N = [x_1; \dots; x_N]$  with the norm  $\|x^N\|_{x,\infty} = \max_t \|x_t\|_{(x)}$ . With affine in purified outputs controller  $\chi$ ,  $x^N$  is an affine function of  $d^N$  and  $z$ ; let  $X[\chi]$  be the matrix of coefficients at  $d^N$  in this affine dependence. Peak-to-peak disturbance-to-state gain stemming from  $\|\cdot\|_{(d)}$  and  $\|\cdot\|_{(x)}$  is, by definition, the norm of  $X[\chi]$  induced by the norms  $\|d^N\|_{d,\infty}$  and  $\|x^N\|_{x,\infty}$ , and the corresponding design specification is just an upper bound on this gain. Since  $X[\chi]$ , as was already mentioned, is affine in  $\chi$ , this specification is a convex constraint on  $\chi$ . However, this constraint can be difficult to handle because the operator norm in question is typically difficult to compute (this is so already when  $\|\cdot\|_{(d)}$  and  $\|\cdot\|_{(x)}$  are  $\|\cdot\|_2$ -norms). In such case, we can utilize our results on Problem A to safely approximate the design specification in question by replacing difficult-to-compute induced norm of  $X = X[\chi]$  by its efficiently computable convex in  $X$  and reasonably tight upper bound, as explained in details in Section 3.3.3.

Our main motivating application for Problem B is identification of parameters  $A$  of discrete time linear time invariant dynamical system

$$x_{t+1} = A[x_t; r_t],$$

from corrupted by noise observations of states  $x_0, \dots, x_N$  and inputs  $r_0, \dots, r_{N-1}$  on a given time horizon. We focus on the case of *uncertain-but-bounded noise*, in which deviations of entries in observations from the actual values of the corresponding entries in  $x_t$  and  $r_t$  are bounded in magnitude. We discuss an approach (to the best of our knowledge, new), heavily utilizing our results on Problem B, for computationally efficient identification of  $A$  and for generating on-line upper bounds on recovery errors.

Note that there is some literature on the first, and huge literature on the second of the just outlined applications. Instead of positioning our results with respect to this literature in the introduction, we find it more productive to postpone this positioning till appropriate parts of the main body of the paper.

*Structure* of the paper is as follows. Section 2 presents background on ellitopes. Section 3 is devoted to Problem A, and Section 4—to Problem B. Technical proofs are relegated to the appendix, where we present additional results on system identification, same as describe how our results can be extended from ellitopes to an essentially wider family of sets—*spectratopes*.

## 2 Preliminaries: ellitopes and spectratopes

Ellitopes and their extensions, spectratopes, introduced in [21], are convex compact sets well-suited for tight upper-bounding maxima of quadratic forms over the sets. To make the paper more readable, in its main body we focus on ellitopes; (always straightforward) extensions to spectratopes are relegated to Appendix.

### 2.1 Ellitopes: definition and basic examples

A *basic ellitope* is a set  $\mathcal{W}$  represented as

$$\mathcal{W} = \{w \in \mathbf{R}^p : \exists t \in \mathcal{T} : w^T T_k w \leq t_k, 1 \leq k \leq K\} \quad (2)$$

<sup>2</sup> For the sake of definiteness, we focus on “disturbance-to-state” peak-to-peak gain; peak-to-peak gains from disturbance to controls, or to outputs, or from initial state to states, etc., are defined similarly and can be processed in the same way.

where  $T_k \succeq 0$ ,  $k \leq K$ ,  $\sum_k T_k \succ 0$ , and  $\mathcal{T}$  is a convex computationally tractable compact *monotone* subset of  $\mathbf{R}_+^K$  with  $\text{int } \mathcal{T} \neq \emptyset$ , monotonicity meaning that when  $0 \leq t \leq t'$  and  $t' \in \mathcal{T}$ , we have  $t \in \mathcal{T}$  as well.

An *ellitope*  $\mathcal{X}$  is a linear image of a basic ellitope:

$$\mathcal{X} = P\mathcal{W} = \{x \in \mathbf{R}^n : \exists w \in \mathcal{W} : x = Pw\} \text{ with } \mathcal{W} \text{ given by (2)} \quad (3)$$

We call  $K$  *ellitopic size* of ellitopes (2) and (3).

Clearly, every ellitope is a convex compact set symmetric w.r.t. the origin; a basic ellitope, in addition, has a nonempty interior.

► **Example 1.**

**A.** Bounded intersection  $\mathcal{X}$  of  $K$  centered at the origin ellipsoids/elliptic cylinders  $\{x \in \mathbf{R}^n : x^T T_k x \leq 1\}$  [ $T_k \succeq 0$ ] is a basic ellitope:

$$\mathcal{X} = \{x \in \mathbf{R}^n : \exists t \in \mathcal{T} := [0, 1]^K : x^T T_k x \leq t_k, k \leq K\}$$

In particular, the unit box  $\{x \in \mathbf{R}^n : \|x\|_\infty \leq 1\}$  is a basic ellitope.

**B.** A  $\|\cdot\|_p$ -ball in  $\mathbf{R}^n$  with  $p \in [2, \infty]$  is a basic ellitope:

$$\{x \in \mathbf{R}^n : \|x\|_p \leq 1\} = \left\{ x : \exists t \in \mathcal{T} = \{t \in \mathbf{R}_+^n, \|t\|_{p/2} \leq 1\} : \underbrace{x_k^2}_{x^T T_k x} \leq t_k, k \leq K \right\}.$$

Ellitopes admit fully algorithmic “calculus”: this family is closed with respect to basic operations preserving convexity and symmetry w.r.t. the origin, e.g., taking finite intersections, linear images, inverse images under linear embedding, direct products, arithmetic summation (for details, see [21, Section 4.6]); what is missing, is taking convex hulls of finite unions.

## 2.2 Bounding maximum of quadratic form over an ellitope

The starting point of what follows is the problem

$$\text{Opt}_*(C) = \max_{x \in \mathcal{X}} x^T C x, \quad C \in \mathbf{S}^n \quad (4)$$

of maximizing a homogeneous quadratic form over a convex compact set  $\mathcal{X} \subset \mathbf{R}^n$ . It is well known that basically the only generic case when the problem is easy is the one where  $\mathcal{X}$  is an ellipsoid. It is shown in [21] that when  $\mathcal{X}$  is an ellitope, (4) admits reasonably tight efficiently computable upper bound. Specifically, when  $\mathcal{X}$  is given by (3),  $\lambda \in \mathbf{R}_+^k$  is such that  $P^T C P \preceq \sum_k \lambda_k T_k$  and  $x \in \mathcal{X}$ , one has for some  $t \in \mathcal{T}$

$$x^T C x = w^T P^T C P w \leq w^T \left[ \sum_k \lambda_k T_k \right] w \leq \sum_k \lambda_k t_k,$$

implying the validity of the implication

$$\lambda \geq 0, \quad P^T C P \preceq \sum_k \lambda_k T_k \Rightarrow \text{Opt}_*(C) \leq \phi_{\mathcal{T}}(\lambda) := \max_{t \in \mathcal{T}} \lambda^T t,$$

and thus—the first claim of the following

► **Theorem 2** ([21, Proposition 4.6]). *Given ellitope (3) and a matrix  $C \in \mathbf{S}^n$ , consider the quadratic maximization problem (4) along with its relaxation*

$$\text{Opt}(C) = \min_{\lambda} \left\{ \phi_{\mathcal{T}}(\lambda) : \lambda \geq 0, P^T C P \preceq \sum_k \lambda_k T_k \right\} \quad (5)$$

*The problem is computationally tractable and solvable, and  $\text{Opt}(C)$  is an efficiently computable upper bound on  $\text{Opt}_*(C)$ . This upper bound is reasonably tight:*

$$\text{Opt}_*(C) \leq \text{Opt}(C) \leq 3 \ln(\sqrt{3}K) \text{Opt}_*(C).$$

To the best of our knowledge, the first result of this type was established in [33] for  $\mathcal{X}$  which is an intersection of  $K$  concentric elliptic cylinders/ellipsoids; in this case, (4) becomes a special case of quadratically constrained quadratic optimization problem, and (5) is the standard Shor’s semidefinite relaxation (see, e.g., [6, Section 4.3]) of this problem. In [33] it is shown that the ratio  $\text{Opt}(C)/\text{Opt}_*(C)$  indeed can be as large as  $O(\ln(K))$ , even when all  $T_k = a_k a_k^T$  are of rank 1 and  $\mathcal{X}$  is the polytope  $\{x : |a_k^T x| \leq 1, k \leq K\}$ .

### 3 Bounding operator norms

As stated in Introduction, one of the subjects of this paper is tight efficiently computable upper-bounding of the operator norm

$$\|A\|_{\mathcal{B},\mathcal{X}} = \max_x \{\|Ax\|_{\mathcal{B}} : \|x\|_{\mathcal{X}} \leq 1\}$$

of a linear mapping  $x \rightarrow Ax : \mathbf{R}^n \rightarrow \mathbf{R}^m$  induced by norms  $\|\cdot\|_{\mathcal{X}}$  and  $\|\cdot\|_{\mathcal{B}}$  on the argument and the destination spaces, with  $\|\cdot\|_{\mathcal{U}}$  standing for the norm with unit ball  $\mathcal{U}$ . Our approach works for the case when  $\mathcal{X}$  and the polar  $\mathcal{B}_*$  of  $\mathcal{B}$  are ellitopes with nonempty interiors:

$$\begin{aligned} \mathcal{X} &= P\mathcal{W} = \{x \in \mathbf{R}^n : \exists w \in \mathcal{W} : x = Pw\}, \text{ int } \mathcal{X} \neq \emptyset, \\ \mathcal{W} &= \{w \in \mathbf{R}^p : \exists t \in \mathcal{T} : w^T T_k w \leq t_k, k \leq K\}; \\ \mathcal{B}_* &:= \{v \in \mathbf{R}^m : v^T y \leq 1 \ \forall y \in \mathcal{B}\} = Q\mathcal{Z} = \{y \in \mathbf{R}^m : \exists z \in \mathcal{Z} : y = Qz\}, \text{ int } \mathcal{B}_* \neq \emptyset, \\ \mathcal{Z} &= \{z \in \mathbf{R}^q : \exists r \in \mathcal{R} : z^T R_\ell z \leq r_\ell, \ell \leq L\} \end{aligned} \quad (6)$$

with  $T_k, \mathcal{T}, R_\ell, \mathcal{R}$  as required in the definition of a basic ellitope.

Under the assumptions just introduced,  $\|A\|_{\mathcal{B},\mathcal{X}}$  is the maximum of a quadratic form over a basic ellitope  $\mathcal{Z} \times \mathcal{W}$ :

$$\begin{aligned} \|A\|_{\mathcal{B},\mathcal{X}} &= \max_{x \in \mathcal{X}} \|Ax\|_{\mathcal{B}} = \max_{y \in \mathcal{B}_*, x \in \mathcal{X}} y^T Ax = \max_{w \in \mathcal{W}, z \in \mathcal{Z}} z^T Q^T APw \\ &= \frac{1}{2} \max_{[z;w] \in \mathcal{Z} \times \mathcal{W}} [z;w]^T \left[ \frac{Q^T AP}{P^T A^T Q} \right] [z;w]. \end{aligned}$$

In this case relaxation (5) provides efficiently computable upper bound on  $\|A\|_{\mathcal{B},\mathcal{X}}$ . Immediate computation taking into account the direct product structure of the ellitope  $\mathcal{Z} \times \mathcal{W}$  and bilinearity of the quadratic form we are maximizing over this ellitope shows that this bound is

$$\text{Opt}(A) = \min_{\lambda, v} \left\{ \phi_{\mathcal{T}}(\lambda) + \phi_{\mathcal{R}}(v) : \lambda \geq 0, v \geq 0, \left[ \frac{\sum_{\ell} v_{\ell} R_{\ell}}{\frac{1}{2} P^T A^T Q} \mid \frac{\frac{1}{2} Q^T AP}{\sum_k \lambda_k T_k} \right] \succeq 0 \right\}. \quad (7)$$

Note that  $\text{Opt}(A)$  clearly is a convex function of  $A$ , and Theorem 2 implies that

$$\|A\|_{\mathcal{B},\mathcal{X}} \leq \text{Opt}(A) \leq 3 \ln(\sqrt{3}[K+L]) \|A\|_{\mathcal{B},\mathcal{X}}.$$

Our main goal is to demonstrate that the latter bound can be refined.

► **Theorem 3.** *In the case of (6) one has*

$$\|A\|_{\mathcal{B},\mathcal{X}} \leq \text{Opt}(A) \leq \varsigma(K, L) \|A\|_{\mathcal{B},\mathcal{X}}, \quad \varsigma(K, L) = \begin{cases} 3\sqrt{\ln(4K)\ln(4L)}, & \max[K, L] > 1 \\ 1, & K = L = 1. \end{cases} \quad (8)$$

► **Remark 4.** Results of [34, 35] imply that in some cases the tightness factor  $\varsigma$  in (8) can be improved to an *absolute* constant. Specifically,

1. In the case of (6) with diagonal matrices  $T_k$  and  $R_\ell$ , it follows from [35, Theorem 13.2.1] that one can take  $\varsigma = \frac{\pi}{4-2\pi} \approx 3.660$
2. When  $\|\cdot\|_{\mathcal{X}} = \|\cdot\|_p, \|\cdot\|_{\mathcal{B}} = \|\cdot\|_r$  with  $\infty \geq p \geq 2, 1 \leq r \leq 2$  (this is a special case of 1)), Nesterov [34, 35] proved that the upper bound

$$\frac{1}{2} \min_{\lambda, \mu} \left\{ \|\lambda\|_{\frac{p}{p-2}} + \|\mu\|_{\frac{r}{2-r}} : \left[ \frac{\text{Diag}\{\mu\}}{A^T} \mid \frac{A}{\text{Diag}\{\lambda\}} \right] \succeq 0 \right\} \quad (9)$$

on  $\|A\|_{p \rightarrow r} := \max_{\|x\|_p \leq 1} \|Ax\|_r$  (this bound coincides with  $\text{Opt}(A)$  when  $\mathcal{X}$  is the ellitope  $\{x : \|x\|_p \leq 1\}$ , and  $\mathcal{B}_*$  is the ellitope  $\{v : \|v\|_{\frac{r}{r-1}} \leq 1\}$ ) is tight within (even better than in 1))

- factor  $\frac{\pi}{2\sqrt{3}-2\pi/3} \approx 2.2936$  in the entire range  $p \in [2, \infty], r \in [1, 2]$ ,
- factor  $\sqrt{\pi/2} \approx 1.2533$  when  $p = 2$  and  $r \in [1, 2]$ .<sup>3</sup>

<sup>3</sup> Using the identity  $\|A\|_{\mathcal{B},\mathcal{X}} = \|A^T\|_{\mathcal{X}_*,\mathcal{B}_*}$ , where  $\mathcal{X}_*$  is the polar of  $\mathcal{X}$  (as is immediately seen, this identity is respected by our bounding scheme), we see that  $\text{Opt}(A)$  is within  $\sqrt{\pi/2}$  from  $\|A\|_{p \rightarrow r}$  when  $p \geq 2$  and  $r = 2$ .

Needless to say, when  $p = r = 2$ , the tightness factor is 1. In addition, it is shown in [41] that in the range  $\infty \geq p \geq 2$ ,  $1 \leq r \leq 2$  bound (9) is *exactly equal* to the corresponding norm of  $A$  for *entrywise nonnegative matrices*.

Note that there is a simple case when  $\text{Opt}(A) = \|A\|_{\mathcal{B}, \mathcal{X}}$ —the one where  $A$  is a row vector,  $\mathcal{B} = [-1, 1] \subset \mathbf{R}$ , and, therefore,

$$\|A\|_{\mathcal{B}, \mathcal{X}} = \max_{x \in \mathcal{X}} Ax.$$

Our bounding is intelligent enough to recognize this situation. Indeed, in the case in question (7) reads

$$\text{Opt}(A) = \min_{\lambda, v} \left\{ \phi_{\mathcal{T}}(\lambda) + v : \lambda \geq 0, \left[ \frac{v}{\frac{1}{2}P^T A^T} \mid \frac{\frac{1}{2}AP}{\sum_k \lambda_k T_k} \right] \succeq 0 \right\}$$

while, by Lagrange duality,

$$\begin{aligned} \max_{x \in \mathcal{X}} Ax &= \max_{w, t} \left\{ APw : w^T T_k w \leq t_k, k \leq K, t \in \mathcal{T} \right\} \\ &= \min_{\lambda \geq 0} \max_{t \in \mathcal{T}, w} \left\{ APw + \sum_k \lambda_k t_k - w^T \left[ \sum_k \lambda_k T_k \right] w \right\} \\ &= \min_{\lambda \geq 0} \max_w \left\{ \phi_{\mathcal{T}}(\lambda) + APw - w^T \left[ \sum_k \lambda_k T_k \right] w \right\} \\ &= \min_{\lambda \geq 0, v} \left\{ \phi_{\mathcal{T}}(\lambda) + v : v - APw + w^T \left[ \sum_k \lambda_k T_k \right] w \geq 0 \forall w \right\} \\ &= \min_{\lambda, v} \left\{ \phi_{\mathcal{T}}(\lambda) + v : \lambda \geq 0, \left[ \frac{v}{\frac{1}{2}P^T A^T} \mid \frac{\frac{1}{2}AP}{\sum_k \lambda_k T_k} \right] \succeq 0 \right\} = \text{Opt}(A). \end{aligned}$$

To put this immediate observation into a proper perspective, see Section 3.2.

The just outlined results are stronger than what in the case in question is stated by Theorem 3. This being said, it can be proved that in the full scope of the latter theorem, logarithmic growth of the tightness factor with  $K, L$  is unavoidable.

### 3.1 On the scope of Theorem 3

The scope of Theorem 3—the set of the matrix norms to which the theorem applies—is restricted to the case when the norm in the argument space is *simple ellitopic norm*, meaning that its unit ball is an ellitope, and the norm on the image space is a *simple co-ellitopic norm*, meaning that the polar of its unit ball is an ellitope. Clearly, simple co-ellitopic norms (s.co.e.n.'s) are exactly the conjugates of simple ellitopic norms (s.e.n.'s). These classes of norms allow for certain “calculus” stating that some standard operations with norms preserve their ellitopic/co-ellitopic type.

*Basic calculus of simple ellitopic norms* is as follows.

- E.1.** (raw materials) When  $p \in [2, \infty]$ ,  $\|\cdot\|_p$  is a s.e.n. on  $\mathbf{R}^n$ ,
- E.2.** (taking finite maxima) When  $\|\cdot\|_{(k)}$ ,  $k \leq K$ , are s.e.n.'s on  $\mathbf{R}^n$ , so is their maximum.
- E.3.** (restriction to a linear subspace) When  $\|\cdot\|$  is a s.e.n. on  $\mathbf{R}^n$  and  $y \rightarrow Ay : \mathbf{R}^{n'} \rightarrow \mathbf{R}^n$  is a linear embedding,  $\|y\|' := \|Ay\|$  is a s.e.n. on  $\mathbf{R}^{n'}$
- E.4.** (passing to factor-norm) When  $\|\cdot\|$  is a s.e.n. on  $\mathbf{R}^n$  and  $x \mapsto Ax : \mathbf{R}^n \rightarrow \mathbf{R}^{n'}$  is an onto mapping, the factor-norm  $\|y\|' = \min_x \{\|x\| : Ax = y\}$  is s.e.n. on  $\mathbf{R}^{n'}$
- E.5.** (“aggregation”) Let  $\|\cdot\|_{(k)}$  be s.e.n. on  $\mathbf{R}^{n_k}$ ,  $k \leq K$ , and let  $\mathcal{A}$  be a monotone convex compact set with a nonempty interior in  $\mathbf{R}_+^K$ . Then the norm on  $\mathbf{R}^{n_1} \times \cdots \times \mathbf{R}^{n_K}$  with the unit ball

$$\mathcal{X} = \{[x_1; \dots; x_K] \in \mathbf{R}^{n_1} \times \cdots \times \mathbf{R}^{n_K} : \exists \alpha \in \mathcal{A}, \|x_k\|_{(k)} \leq \sqrt{\alpha_k}, k \leq K\}$$

is s.e.n. For instance, when  $p_k \in [2, \infty]$  and  $p \in [2, \infty]$ , the norm on  $\mathbf{R}^{n_1} \times \cdots \times \mathbf{R}^{n_K}$  given by  $\|[x_1; \dots; x_K]\| = \|[ \|x_i\|_{p_i}; \dots; \|x_K\|_{p_K} ]\|_p$  is s.e.n.

All these rules are immediate consequences of “calculus of ellitopes” [21, Section 4.6].



Basic calculus of simple co-elliptic norms is as follows.

**cE.1.** [raw materials] When  $r \in [1, 2]$ ,  $\|\cdot\|_r$  is a s.co-e.n. on  $\mathbf{R}^n$  (cf. E.1)

**cE.2.** [taking sums] When  $\|\cdot\|_{(k)}$ ,  $k \leq K$ , are s.co-e.n.'s on  $\mathbf{R}^n$ , so is their sum.

Indeed, the unit ball  $B$  of the sum of norms with polars  $B_i^*$  of the unit balls is

$$B = \left\{ x : \sum_i \max_{y_i \in B_i^*} y_i^T x \leq 1 \right\} = \left\{ x : \max_{y=[y_1; \dots; y_K] \in B_1^* \times \dots \times B_K^*} x^T [y_1 + \dots + y_K] \leq 1 \right\},$$

that is, the polar  $B^*$  of  $B$  is the image of  $B_1^* \times \dots \times B_K^*$  under a linear mapping. When all  $B_k^*$  are ellitopes, so is their direct product, and therefore—its linear image  $B^*$ . Thus, the polar of  $B$  is an ellitope, as claimed.

**cE.3.** [restriction to a linear subspace] When  $\|\cdot\|$  is a s.co-e.n. on  $\mathbf{R}^n$  and  $y \rightarrow Ay : \mathbf{R}^{n'} \rightarrow \mathbf{R}^n$  is a linear embedding,  $\|y\|' := \|Ay\|$  is a s.co-e.n. on  $\mathbf{R}^{n'}$

Indeed, assuming that the polar  $B^*$  of the unit ball of  $\|\cdot\|$  is an ellitope, we have  $\|y\|' = \max_{z \in B^*} z^T Ay$ . That is, the polar of the unit ball of  $\|\cdot\|'$  is the linear image  $A^T B^*$  of  $B^*$ , which is an ellitope along with  $B^*$ .

**cE.4.** [passing to factor-norm] When  $\|\cdot\|$  is a s.co-e.n. on  $\mathbf{R}^n$  and  $x \mapsto Ax : \mathbf{R}^n \rightarrow \mathbf{R}^{n'}$  is an onto mapping, the factor-norm  $\|y\|' = \min_x \{\|x\| : Ax = y\}$  is s.co-e.n. on  $\mathbf{R}^{n'}$ .

Indeed, when the polar  $B^*$  of the unit ball of  $\|\cdot\|$  is an ellitope, when denoting by  $A^\dagger$  the pseudoinverse of the onto mapping  $A$ , one has

$$\|y\|' = \min_{\delta \in \text{Ker } A} \|A^\dagger y + \delta\| = \min_{\delta \in \text{Ker } A} \max_{z \in B^*} z^T [A^\dagger y + \delta] = \max_{z \in B^* \cap \text{Im } A^T} [[A^\dagger]^T z]^T y.$$

Thus, the polar of the unit ball of  $\|\cdot\|'$  is a linear image of the intersection of ellitope  $B^*$  with a linear subspace, and as such is an ellitope.

**cE.5.** [“aggregation”] Let  $\|\cdot\|_{(k)}$  be s.co-e.n. on  $\mathbf{R}^{n_k}$ ,  $k \leq K$ , and let  $\mathcal{A}$  be a monotone convex compact set with a nonempty interior in  $\mathbf{R}_+^K$ . Then the norm on  $\mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_K}$  given by

$$\|x_1; \dots; x_K\| = \max_{\beta \in \mathcal{A}} \sum_k \sqrt{\beta_k} \|x_k\|_{(k)}$$

is s.co-e.n. For instance, when  $r_k \in [1, 2]$  and  $r \in [1, 2]$ , the norm on  $\mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_K}$  given by  $\|x_1; \dots; x_K\| = \|[\|x_1\|_{r_1}; \dots; \|x_K\|_{r_K}]\|_r$  is s.co-e.n.

Indeed, let  $\|\cdot\|_{(k)}^*$  be the s.e.n.'s conjugate to  $\|\cdot\|_{(k)}$ . Setting  $\mathcal{A}^{1/2} = \{[\alpha_1; \dots; \alpha_K] \geq 0 : [\alpha_1^2; \alpha_2^2; \dots; \alpha_K^2] \in \mathcal{A}\}$ , we get a convex compact monotone subset of  $\mathbf{R}_+^K$  such that the unit ball  $\mathcal{B}$  of  $\|\cdot\|$  is  $\mathcal{B} = \{[x_1; \dots; x_K] : \phi_{\mathcal{A}^{1/2}}(\|x_1\|_{(1)}; \dots; \|x_K\|_{(K)}) \leq 1\}$ . Hence, as is immediately seen, the polar  $\mathcal{B}_*$  of  $\mathcal{B}$  is

$$\begin{aligned} \mathcal{B}_* &= \left\{ [y_1; \dots; y_K] : \sum_k \zeta_k \|y_k\|_{(k)}^* \leq 1 \ \forall (\zeta \geq 0 : \sum_k \alpha_k \zeta_k \leq 1 \ \forall \alpha \in \mathcal{A}^{1/2}) \right\} \\ &= \left\{ [y_1; \dots; y_K] : \exists \alpha \in \mathcal{A}^{1/2} : \|y_k\|_{(k)}^* \leq \alpha_k, \ k \leq K \right\}, \end{aligned}$$

that is,  $\|\cdot\|_*$  is s.e.n. by E.5.

### 3.2 An extension

The above results can be straightforwardly extended from the case when  $\mathcal{B}_*$  and  $\mathcal{X}$  are ellitopes onto a more general case. Specifically, assume that

**A.**  $\mathcal{X} \subset \mathbf{R}^n$  is a set with nonempty interior represented as the convex hull of a finite union of ellitopes, or, which is the same,

$$\begin{aligned} \mathcal{X} &= \text{Conv} \left\{ \bigcup_{i=1}^I P_i \mathcal{X}_i \right\} = \left\{ x = \sum_{i=1}^I \lambda_i P_i x_i : x_i \in \mathcal{X}_i, \lambda_i \geq 0, \sum_i \lambda_i = 1 \right\} \\ &= \left\{ x = \sum_i P_i x_i : \sum_i \|x_i\|_{\mathcal{X}_i} \leq 1 \right\}, \end{aligned} \tag{10}$$

where  $\mathcal{X}_i \subset \mathbf{R}^{n_i}$  are basic ellitopes and  $\|\cdot\|_{\mathcal{X}_i}$  are s.e.n. on  $\mathbf{R}^{n_i}$  with unit balls  $\mathcal{X}_i$ .



Under Assumption A,  $\mathcal{X}$  is a convex compact symmetric w.r.t. the origin subset of  $\mathbf{R}^n$  with  $0 \in \text{int } \mathcal{X}$ ; as such,  $\mathcal{X}$  is the unit ball of a norm  $\|\cdot\|_{\mathcal{X}}$ . In the sequel we refer to the norms of this structure as to *elliptic norms*. Clearly, every simple elliptic norm is elliptic, e.g., the block  $\ell_{\infty}$  norm

$$\|[x_1; \dots; x_I]\| = \max_{i \leq I} \|x_i\|_{p_i} \quad [p_i \in [2, \infty] \ \forall i]$$

on the space  $\mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_I}$  is s.e.n. (by E.1 and E.5). In fact, the family of elliptic norms is much wider than the family of s.e.n.'s. For example,

**E.1.** When  $\|x_i\|_{(i)}$  are elliptic norms on  $\mathbf{R}^{n_i}$ ,  $i \leq I$ , the associated block  $\ell_1/\|\cdot\|_{(\cdot)}$  norm

$$\|[x_1; \dots; x_I]\| = \sum_{i=1}^I \|x_i\|_{(i)} \quad (11)$$

on  $\mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_I}$  is elliptic.

Indeed, the unit ball  $\mathcal{X}_i$  of  $\|\cdot\|_{(i)}$  is a convex subset of  $\mathbf{R}^{n_i}$  of the form

$$\text{Conv} \left\{ \bigcup_{\nu=1}^{I_i} P_{i\nu} \mathcal{X}_{i\nu} \right\}$$

with basic ellitopes  $\mathcal{X}_{i\nu}$ . Specifying linear mappings  $P_i$  from  $\mathbf{R}^{n_i}$  to  $\mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_I}$  as the natural embeddings

$$[P_i x_i]_s = \begin{cases} 0 \in \mathbf{R}^{n_s}, & s \neq i \\ x_i, & s = i, \end{cases}$$

the unit ball  $\mathcal{X}$  of norm (11) clearly is  $\text{Conv} \left\{ \bigcup_{i \leq I, \nu_i \leq I_i} P_i P_{i\nu_i} \mathcal{X}_{i\nu_i} \right\}$ . Because, in addition, this set has a nonempty interior, (11) is an elliptic norm.

Note that the property to be elliptic is inherited when passing to factor-norms (cf. E.4):

**E.2.** When  $\|\cdot\|$  is an elliptic norm and  $y \mapsto Ay : \mathbf{R}^n \rightarrow \mathbf{R}^{n'}$  is an onto mapping, the factor-norm  $\|x\|' = \min_y \{\|y\| : Ay = x\}$  on  $\mathbf{R}^{n'}$  induced by  $\|\cdot\|$  and  $A$  is elliptic.

Indeed, if the unit ball  $\mathcal{X}$  of  $\|\cdot\|$  is given by (10) then the unit ball  $\mathcal{X}'$  of  $\|\cdot\|'$  is the convex compact set with a nonempty interior given by

$$\mathcal{X}' = A\mathcal{X} = \text{Conv} \left\{ \bigcup_{i=1}^I [AP_i] \mathcal{X}_i \right\}.$$

**E.3.** Let  $\|\cdot\|_{(\chi)}$  be an elliptic norm on  $\mathbf{R}^{n_{\chi}}$ ,  $\chi = 1, 2$ . Then the norm  $\|[x_1; x_2]\| = \max[\|x_1\|_{(1)}, \|x_2\|_{(2)}]$  on  $\mathbf{R}^{n_1} \times \mathbf{R}^{n_2}$  is elliptic.

Indeed, if the unit ball of  $\|\cdot\|_{(\chi)}$  is  $\text{Conv} \left\{ \bigcup_{i=1}^{I_{\chi}} P_{i,\chi} \mathcal{X}_{i,\chi} \right\}$ , then the unit ball of  $\|\cdot\|$  is

$$\text{Conv} \left\{ \bigcup_{i_1 \leq I_1, i_2 \leq I_2} \text{Diag} \{ P_{i_1,1}, P_{i_2,2} \} [\mathcal{X}_{i_1,1} \times \mathcal{X}_{i_2,2}] \right\},$$

and  $\mathcal{X}_{i_1,1} \times \mathcal{X}_{i_2,2}$  are basic ellitopes along with  $\mathcal{X}_{i_1,1}$ ,  $\mathcal{X}_{i_2,2}$ .

By E.3, if  $\|\cdot\|_{(i)}$  are elliptic norms on  $\mathbf{R}^{n_i}$ ,  $i \leq I$ , then the norm  $\|[x_1; \dots; x_I]\| = \max_{i \leq I} \|x_i\|_{(i)}$  on  $\mathcal{R}^{n_1 + \dots + n_I}$  is elliptic as well. Note, however, that the number of ellitopes involved in the description of this norm is the product, over  $i \leq I$ , of the numbers of ellitopes in the description of norms  $\|\cdot\|_{(i)}$  and thus may explode exponentially fast as  $I$  grows.

Assume, next, that

**B.**  $\mathcal{B} \subset \mathbf{R}^m$  is a set with nonempty interior which is *the polar* of a set of the structure described in A:

$$\mathcal{B} = \left\{ v \in \mathbf{R}^m : \max_{y \in \mathcal{B}_*} v^T y \leq 1 \right\}, \quad \mathcal{B}_* = \left\{ y = \sum_{j=1}^J \mu_j Q_j z_j, z_j \in \mathcal{Z}_j, \mu_j \geq 0, \sum_j \mu_j = 1 \right\} \quad (12)$$

where  $\mathcal{Z}_j \subset \mathbf{R}^{m_j}$  are basic ellitopes and  $Q_j$ ,  $\mathcal{Z}_j$  are such that  $\mathcal{B}_*$  has a nonempty interior.

Under Assumption B,  $\mathcal{B}$  is a convex compact symmetric w.r.t. the origin subset of  $\mathbf{R}^n$  with  $0 \in \text{int } \mathcal{B}$ ; as such,  $\mathcal{B}$  is the unit ball of a norm  $\|\cdot\|_{\mathcal{B}}$ . In the sequel, we refer to norms of this structure as *co-elliptic*. Clearly, the conjugate of an elliptic norm is co-elliptic, and vice versa.

Note that in the case of (12) we have

$$\begin{aligned} \|u\|_{\mathcal{B}} &= \max_{y \in \mathcal{B}^*} u^T y = \max_{\{z_j, \mu_j\}} \left\{ \sum_j \mu_j u^T Q_j z_j : z_j \in \mathcal{Z}_j, \mu_j \geq 0 \ \forall j, \sum_j \mu_j = 1 \right\} \\ &= \max_j \max_{z_j \in \mathcal{Z}_j} u^T Q_j z_j = \max_{j \leq J} \|Q_j^T u\|_{\mathcal{Z}_j^*} \end{aligned} \quad (13)$$

where  $\mathcal{Z}_j^*$  is the polar of  $\mathcal{Z}_j$ .

Of course, every simple co-elliptic norm is co-elliptic. In fact, the family of co-elliptic norms is much wider than the family of simple co-elliptic norms due to the following observations:

**cE.1.** Maximum of finitely many co-elliptic norms is co-elliptic.

Indeed, if  $\|\cdot\|_{(k)}$ ,  $k \leq K$ , are co-elliptic norms on  $\mathbf{R}^n$ , their conjugates  $\|\cdot\|_{(k)}^*$  are elliptic, implying by E.1 that the norm  $\|[y_1; \dots; y_K]\| = \sum_k \|y_k\|_{(k)}^*$  on  $\mathbf{R}^{Kn}$  is elliptic, which by E.2 implies that the factor-norm

$$\|z\|_* = \min_{\{y_k\}} \left\{ \sum_k \|y_k\|_{(k)}^* : \sum_k y_k = z \right\}$$

is elliptic. The unit ball of the latter norm is the convex compact set

$$\mathcal{B}_* = \left\{ z = \sum_k y_k : \sum_k \|y_k\|_{(k)}^* \leq 1 \right\},$$

and the polar of this set is

$$\begin{aligned} \mathcal{B} &= \left\{ x : \max_y \left\{ \left[ \sum_k y_k \right]^T x : \sum_k \|y_k\|_{(k)}^* \leq 1 \right\} \leq 1 \right\} \\ &= \left\{ x : \max_{\lambda, y} \left\{ \left[ \sum_k y_k \right]^T x : \|y_k\|_{(k)}^* \leq \lambda_k, \sum_k \lambda_k \leq 1 \right\} \leq 1 \right\} \\ &= \left\{ x : \max_{\lambda} \left\{ \sum_k \left[ \max_{y_k} \{x^T y_k : \|y_k\|_{(k)}^* \leq \lambda_k\} \right], \sum_k \lambda_k \leq 1 \right\} \leq 1 \right\} \\ &= \left\{ x : \max_{\lambda \geq 0} \left\{ \sum_k \lambda_k \|x\|_{(k)} : \sum_k \lambda_k \leq 1 \right\} \leq 1 \right\} = \left\{ x : \max_k \|x\|_{(k)} \leq 1 \right\}. \end{aligned}$$

Thus, the norm  $\max_k \|x\|_{(k)}$  is conjugate to the elliptic norm  $\|\cdot\|_*$  and as such is co-elliptic.

A closely related statement is

**cE.2.**  $\ell_\infty$ -aggregation

$$\|[x_1; \dots; x_K]\| = \max_{k \leq K} \|x_k\|_{(k)} \quad (14)$$

of co-elliptic norms  $\|\cdot\|_{(k)}$  on  $\mathbf{R}^{n_k}$  is co-elliptic.

Indeed, as we have seen when justifying cE.1, if  $\|\cdot\|_{(k)}^*$  are elliptic norms conjugate to  $\|\cdot\|_{(k)}$ , the norm

$$\|[y_1; \dots; y_K]\|_* = \sum_k \|y_k\|_{(k)}^*$$

is elliptic; clearly, norm (14) is conjugate to this elliptic norm.

The second observation is as follows.

**cE.3.** The restriction of a co-elliptic norm onto a linear subspace is co-elliptic.

Indeed, we should verify that if  $x \mapsto Ax$  is an embedding of  $\mathbf{R}^{n'}$  into  $\mathbf{R}^n$  and  $\|\cdot\|$  is a co-elliptic norm on  $\mathbf{R}^n$  then the norm  $\|x\|' = \|Ax\|$  is co-elliptic. This is immediate—by the standard properties of norms,

under the circumstances, the norm conjugate to  $\|x\|'$  is the factor-norm  $\min_y \{\|y\|_* : A^T y = x\}$  induced by the conjugate to  $\|\cdot\|$  norm  $\|\cdot\|_*$  on  $\mathbf{R}^n$ . This conjugate is an elliptic norm on  $\mathbf{R}^n$ , and it remains to use E.2.

**cE.4.** The sum of two co-elliptic norms on  $\mathbf{R}^n$  is co-elliptic.

Indeed, if  $\|\cdot\|_{(\chi)}$ ,  $\chi = 1, 2$ , are co-elliptic norms on  $\mathbf{R}^n$ , and  $\|\cdot\|_{(\chi)}^*$  are their conjugates, then the norm  $\|[x_1; x_2]\|_* = \max[\|x_1\|_{(1)}^*, \|x_2\|_{(2)}^*]$  is elliptic norm on  $\mathbf{R}^{2n}$  by E.3, so that its conjugate, which is  $\|[x_1; x_2]\|_+ = \|x_1\|_{(1)} + \|x_2\|_{(2)}$ , is co-elliptic. By cE.3, the restriction of the latter norm on the subspace  $\{[x_1; x_2] : x_1 = x_2\} = [I_n; I_n]\mathbf{R}^n$  also is co-elliptic, and this restriction is nothing but the norm  $\|x\| = \|x\|_{(1)} + \|x\|_{(2)}$ .

### Simple observation

Let  $\|\cdot\|_{\mathcal{X}}$  and  $\|\cdot\|_{\mathcal{B}}$  be norms with  $\mathcal{X}$  given by (10) and  $\mathcal{B}$  given by (12). Then the operator norm of  $A \in \mathbf{R}^{m \times n}$  induced by the norms  $\|\cdot\|_{\mathcal{X}}$  and  $\|\cdot\|_{\mathcal{B}}$  on the argument and image spaces can be computed as follows, see (13):

$$\begin{aligned} \|A\|_{\mathcal{B}, \mathcal{X}} &= \max_x \{ \|Ax\|_{\mathcal{B}} : \|x\|_{\mathcal{X}} \leq 1 \} = \max_x \left\{ \max_j \|Q_j^T Ax\|_{\mathcal{Z}_j^*} : \|x\|_{\mathcal{X}} \leq 1 \right\} \\ &= \max_x \left\{ \max_j \max_{z_j} \{ z_j^T Q_j^T Ax : z_j \in \mathcal{Z}_j \} : \|x\|_{\mathcal{X}} \leq 1 \right\} \\ &= \max_j \left\{ \max_{x, z_j} z_j^T Q_j^T Ax : z_j \in \mathcal{Z}_j, x \in \mathcal{X} \right\} \\ &= \max_j \left\{ \max_{x, z_j} z_j^T Q_j^T Ax : z_j \in \mathcal{Z}_j, x \in \text{Conv}\{\cup_i P_i \mathcal{X}_i\} \right\} \\ &= \max_j \left\{ \max_i \left[ \max_{z_j \in \mathcal{Z}_j, x_i \in \mathcal{X}_i} z_j^T Q_j^T A P_i x_i \right] \right\} = \max_{i, j} \|Q_j^T A P_i\|_{ij} \end{aligned} \quad (15)$$

where

$$\|Q_j^T A P_i\|_{ij} = \max_{z_j \in \mathcal{Z}_j, x_i \in \mathcal{X}_i} z_j^T [Q_j^T A P_i] x_i = \|Q_j^T A P_i\|_{\mathcal{Z}_j^*, \mathcal{X}_i}. \quad (16)$$

Note that by the same token  $\max_i \|Q_j^T A P_i\|_{ij} = \|Q_j^T A\|_{\mathcal{Z}_j^*, \mathcal{X}}$  and  $\max_j \|Q_j^T A P_i\|_{ij} = \|A P_i\|_{\mathcal{B}, \mathcal{X}_i}$ , so that in the case of (10), (12) it holds

$$\|A\|_{\mathcal{B}, \mathcal{X}} = \max_j \|Q_j^T A\|_{\mathcal{Z}_j^*, \mathcal{X}} = \max_i \|A P_i\|_{\mathcal{B}, \mathcal{X}_i}.$$

As we know from Theorem 3, we can upper-bound  $\|Q_j^T A P_i\|_{ij}$  by  $\Phi_{ij}(Q_j^T A P_i)$  with convex and efficiently computable function  $\Phi_{ij}(\cdot)$ , the bound being tight within the factor  $\varsigma(K_i, L_j) \leq 3\sqrt{\ln(4K_i) \ln(4L_j)}$ , where  $K_i$  and  $L_j$  are the elliptic sizes of  $\mathcal{X}_i$  and  $\mathcal{Z}_j$ . As a result, the efficiently computable convex function

$$\Phi(A) = \max_{i, j} \Phi_{ij}(Q_j^T A P_i)$$

is an upper bound on  $\|A\|_{\mathcal{B}, \mathcal{X}}$  tight within the factor  $3\sqrt{\ln(4 \max_i K_i) \ln(4 \max_j L_j)}$ .

In some simple situations the above tightness factor can be improved. For example, when

$$\mathcal{X}_i = \{x_i : \|x_i\|_{p_i} \leq 1\}, \quad \mathcal{Z}_j = \{z_j : \|z_j\|_{q_j} \leq 1\}, \quad p_i \geq 2, \quad q_j \geq 2,$$

by Nesterov's results (cf. Remark 4), the tightness factor is an absolute constant (e.g., 1 in the trivial case where  $p_i = q_j = 2$  for all  $i, j$ ).

## 3.3 Applications

### 3.3.1 Least norm projector synthesis

Consider the *projection problem* as follows: we are given a linear subspace  $\mathcal{F}$  of linear space  $\mathcal{E} = \mathbf{R}^n$  and a norm  $\theta(\cdot)$  on  $\mathcal{E}$ ; our goal is to find a linear projector  $H$  of  $\mathcal{E}$  onto  $\mathcal{F}$ —a linear map  $x \mapsto Hx : \mathcal{E} \rightarrow \mathcal{F}$  with  $Hx = x$  for all  $x \in \mathcal{F}$ —which deviates the least from the identity mapping  $\text{Id}$  in the norm

$$\|\cdot\|_{\theta \rightarrow \theta} : \|A\|_{\theta \rightarrow \theta} = \max_{x \in \mathcal{E}} \{\theta(Ax) : \theta(x) \leq 1\}.$$

Consider the case when the norm in question is the block  $\ell_\infty/\ell_2$  norm

$$\theta(x) = \max_{k \leq K} \|G_k x\|_2 \quad \left[ x \mapsto G_k x : \mathcal{E} \rightarrow \mathbf{R}^{\nu_k}, \bigcap_k \text{Ker } G_k = \{0\} \right].$$

What makes the projection problem potentially difficult is the *block*  $\ell_\infty$  structure of  $\theta$ ; were  $\nu_k = 1$  for all  $k$ ,  $\|\cdot\|_{\theta \rightarrow \theta}$  would have polyhedral epigraph, and minimization of  $\|\text{Id} - H\|_{\theta \rightarrow \theta}$  would be a Linear Programming problem (note that property of  $H$  to project onto  $\mathcal{F}$  reduces to a system of linear equalities on  $H$ ).<sup>4</sup> In contrast, in the general  $\ell_\infty/\ell_2$  case as described above, the problem is NP hard. At the same time, the problem is within the scope of our machinery: the unit ball of  $\theta$  is the ellipsope

$$\mathcal{X} = \{x \in \mathbf{R}^n : x^T G_k^T G_k x \leq 1, k \leq K\},$$

and therefore  $\theta$  is a simple ellipsope norm. At the same time, we have

$$\theta(x) = \|Gx\|_{\infty/2}, \quad Gx = [G_1 x; G_2 x; \dots; G_K x], \quad \|[y_1; \dots; y_k]\|_{\infty/2} = \max_k \|y_k\|_2$$

As we know,  $\|\cdot\|_{\infty/2}$  is co-ellipsope (see cE.2 in Section 3.2) and this property is preserved under restriction of a norm on a linear subspace (cE.3), and it remains to recall that  $G$  is an embedding. The bottom line is that we can process the projection problem as explained in Section 3.2. It is immediately seen that the corresponding recipe, under the circumstances, boils down to the following:

We select a linear basis  $\{g_i : i \leq n\}$  in  $\mathcal{E}$  in such a way that the first  $m = \dim \mathcal{F}$  of these vectors form a basis of  $\mathcal{F}$ ; in the sequel, we identify vectors from  $\mathcal{E}$  with collections of their coordinates in this basis, and linear mappings from  $\mathcal{E}$  to  $\mathcal{E}$  with their matrices in this basis. Note that the (matrices of) projectors of  $\mathcal{E}$  onto  $\mathcal{F}$  are exactly block-matrices  $\left[ \begin{array}{c|c} I_m & P \\ \hline & \end{array} \right]$  with  $m \times (n - m)$  blocks  $P$ . Applying Theorem 3, we arrive at the efficiently solvable convex optimization problem

$$\text{Opt} = \min_{P, \{\mu_k, \lambda^k : k \leq K\}} \left\{ \max_k \left[ \mu_k + \sum_{j=1}^K \lambda_j^k \right] : P \in \mathbf{R}^{m \times (n-m)}, \lambda^k \geq 0, \right. \\ \left. \left[ \begin{array}{c|c} \mu_k I_{\nu_k} & \frac{1}{2} G_k \left[ \begin{array}{c|c} P & \\ \hline & -I_{n-m} \end{array} \right] \\ \hline \frac{1}{2} \left[ \begin{array}{c|c} P^T & -I_{n-m} \end{array} \right] G_k^T & \sum_j \lambda_j^k G_j^T G_j \end{array} \right] \succeq 0, k \leq K \right\} \quad (17)$$

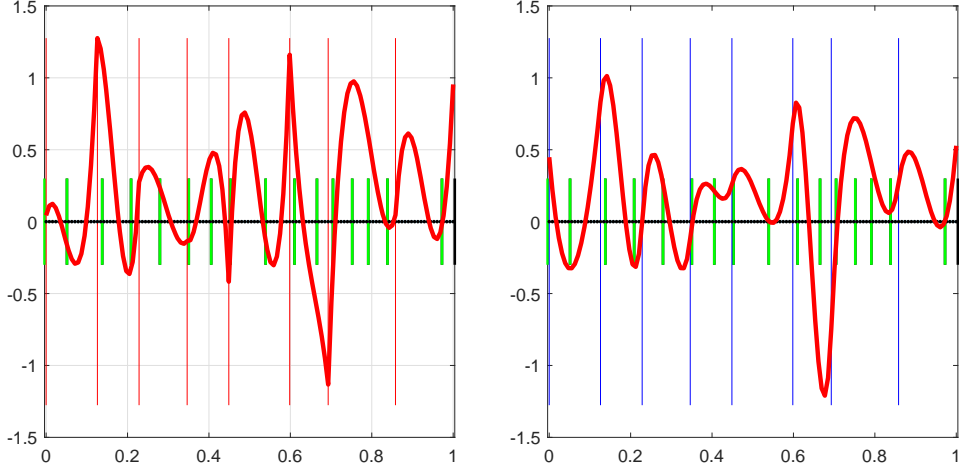
which is a safe tractable approximation of the problem of interest—the  $P$ -component of a feasible solution to the problem specifies projector of  $\mathcal{E}$  onto  $\mathcal{F}$  with the value of  $\|\cdot\|_{\theta \rightarrow \theta}$  not exceeding the value of the objective at this solution. This approximation is tight within the factor  $O(1)\sqrt{\ln(4K)}$ , meaning that Opt is at most by this factor greater than the actual optimal value in the projection problem. In addition, when  $\nu_k = 1 \forall k$ , the tightness factor is exactly 1.

### 3.3.2 Illustration: projecting splines

Consider a partition of  $[0, 1]$  into  $M$  “large” segments, which are further partitioned into total of  $N$  “small” segments. Let also  $\Gamma$  be equidistant grid on  $[0, 1]$  with  $L$  points. Given nonnegative integers  $\mu_L \geq \nu_L$ ,  $\mu_S \geq \mu_L$ , and  $\nu_S \leq \nu_L$ , let us define  $\mathcal{F}$  as the linear space of restrictions on  $\Gamma$  of splines which are polynomials of degree at most  $\mu_L$  in every large segment, with all derivatives of order  $\leq \nu_L$  continuous on the entire  $[0, 1]$ . We define  $\mathcal{E}$  as the linear space of restrictions on  $\Gamma$  of splines which are polynomials of degree of order  $\leq \mu_S$  in every small segment and have continuous on  $[0, 1]$  derivatives of order  $\leq \nu_S$ . With the above inequalities between  $\mu$ 's and  $\nu$ 's,  $\mathcal{F}$  is a subspace in  $\mathcal{E}$ . Now let  $\Delta_1, \dots, \Delta_K$  be partitioning of  $\Gamma$  into  $K$  consecutive segments, and let  $\theta$  be the  $\ell_\infty/\ell_2$  norm on  $\mathcal{E}$  given by

$$\theta(x) = \max_{k \leq K} \sqrt{\sum_{i \in \Delta_k} x_i^2},$$

$x_i$  being the value of spline  $x \in \mathcal{E}$  at the  $i$ -th point of  $\Gamma$ .



■ **Figure 1** Spline  $x$  from  $\mathcal{E}$  (left plot) and its projection  $Hx$  on  $\mathcal{F}$  (right plot).

In Figure 1 we present a sample pair of a spline from  $\mathcal{E}$  and its projection onto  $\mathcal{F}$ .

In this experiment,  $|\Gamma| = 128$ , there are eight identical large and small segments (separated by red/blue vertical lines on the plots), and  $K = 16$  (on the plots, 16 segments  $\Delta_k$  are separated from each other by green vertical lines). Splines from  $\mathcal{E}$  are continuous on  $[0, 1]$  and are polynomials of degree 3 on large/small segments, and  $\mathcal{F}$  is cut off  $\mathcal{E}$  by additional requirement for the spline to be continuously differentiable on  $[0, 1]$ . Solving (17) yields  $H$  with  $\|\text{Id} - H\|_{\theta \rightarrow \theta} \leq \text{Opt} \approx 1.255$  and  $H$  is “essentially different” from the  $\|\cdot\|_2$ -orthogonal projection<sup>5</sup>  $\bar{H}$  of  $\mathcal{E}$  onto  $\mathcal{F}$ —the spectral norm of  $H - \bar{H}$  is  $\approx 0.69$ , and the upper bound on  $\|\text{Id} - \bar{H}\|_{\theta \rightarrow \theta}$ , as given by our machinery, is  $\approx 1.527$ . In fact both upper bounds  $\approx 1.255$  on  $\|\text{Id} - H\|_{\theta \rightarrow \theta}$  and  $\approx 1.527$  on  $\|\text{Id} - \bar{H}\|_{\theta \rightarrow \theta}$  happen to coincide within four significant digits with the quantities themselves.<sup>6</sup>

### 3.3.3 Synthesis of linear controller with peak-to-peak design specifications

The situation we are about to address is as follows. We control a discrete time linear system

$$\begin{aligned} x_{t+1} &= A_t x_t + B_t u_t + D_t d_t, & x_0 &= z, \\ y_t &= C_t x_t + E_t d_t, & 0 \leq t < N, \end{aligned}$$

where  $x_t \in \mathbf{R}^{n_x}$ ,  $u_t \in \mathbf{R}^{n_u}$ ,  $d_t \in \mathbf{R}^{n_d}$ , and  $y_t \in \mathbf{R}^{n_y}$  are, respectively, states, controls, external disturbances, and observable outputs. When augmented with non-anticipating affine controller

$$u_t = g_t + \sum_{\tau=0}^t G_\tau^t y_{t-\tau}$$

the closed loop system specifies affine mappings

$$\begin{aligned} (d^N := [d_0; d_1; \dots; d_{N-1}], z) &\mapsto x^N := [x_1; \dots; x_N] = X_d^N d^N + X_z^N z + X^N \\ (d^N, z) &\mapsto u^N := [u_0; \dots; u_{N-1}] = U_d^N d^N + U_z^N z + U^N, \\ (d^N, z) &\mapsto y^N := [y_0; \dots; y_{N-1}] = Y_d^N d^N + Y_z^N z + Y^N. \end{aligned}$$

With “smart parameterizations” of the controller—passing from  $\{g_t, G_\tau^t, 0 \leq t < N, 0 \leq \tau \leq t\}$  to the parameters of the *affine purified-output-based controller*, matrices  $X_d^N, \dots, Y^N$  become *affine* functions of the vector  $\chi$  of controller’s parameters; this vector runs through certain finite-dimensional linear space  $\mathcal{C}$  equipped with filtration  $\mathcal{C}_0 \subset \mathcal{C}_1 \subset \dots \subset \mathcal{C}_{N-1} = \mathcal{C}$  by linear subspaces, with  $\mathcal{C}_d$  comprised of “controllers with memory  $d$ ”. We refer the reader to [22] for details of the controller construction.

<sup>4</sup> Allowing for a slight abuse of notation, we denote with  $H$  the matrix of the linear mapping  $H$ .

<sup>5</sup> Recall that vectors from  $\mathcal{E}$  are restrictions of functions on  $[0, 1]$  onto equidistant grid in this segment and as such  $\mathcal{E}$  is equipped with “canonical” Euclidean structure.

<sup>6</sup> One can easily build a numerical lower bound on  $\|A\|_{\mathcal{B}, \mathcal{X}}$  by alternating maximization of the bilinear function  $u^T A x$  over  $u \in \mathcal{B}_*$  and  $x \in \mathcal{X}$ ; in the reported experiment, these lower bounds were within the indicated accuracy with the upper bounds yielded by our machinery.

When designing a controller, one of natural design specifications (traditionally considered as not so easy to handle, cf., e.g., [1, 4, 5, 10, 15] and reference therein) are bounds on “peak-to-peak” gains. The *disturbance-to-state* gain is nothing but the norm of the matrix  $X^N$  induced by the norm

$$\|d^N\|_{\infty/p} = \max_{0 \leq t < N} \|d_t\|_p$$

on the space of sequences  $d^N$  of disturbances and the norm

$$\|x^N\|_{\infty/r} = \max_{1 \leq t \leq N} \|x_t\|_r$$

on the space of sequences of states; disturbance-to-control and disturbance-to-output peak-to-peak gains are defined similarly. When  $\infty \geq p \geq 2$  and  $1 \leq r \leq 2$ , we can enforce the desired bound on the peak-to-peak gain (which can be difficult to handle, since the corresponding norm of  $X_d^N$  is, in general, difficult to compute) by bounding from above the efficiently computable upper bound, yielded by our machinery, on the gain. As a result, we get an efficiently tractable convex constraint on the parameters of the controller which safely (and tightly within the factor  $\sqrt{\pi/2}$ , see the concluding comments in Section 3.2) approximates the design specification in question.

Note that our machinery remains applicable when  $\|\cdot\|_p$  and  $\|\cdot\|_r$  are replaced with, respectively, an s.e.n.  $\|\cdot\|_{(d)}$  and a s.co-e.n.  $\|\cdot\|_{(x)}$ , and also when the design specifications impose bound on the “restricted” peak-to-peak gains, e.g., on the peak-to-peak disturbance-to-state gain when the disturbances  $d^N$  are restricted to reside in a given linear subspace of the “complete disturbance space”  $\mathbf{R}^{n_d N}$ .

*The first numerical experiment* we are about to present<sup>7</sup> deals with minimizing disturbance-to-state  $\infty/2$  peak-to-peak gain (i.e.,  $p = r = 2$ ) when controlling linearized and discretized in time motion of Boeing 747; the model we use originates from [9], see also Section 4.3.2 below. We omit irrelevant for our purposes details (which can be found in [22]), here it suffices to mention that the model is time-invariant (matrices  $A_t \equiv A, \dots, E_t \equiv E$ ) with  $n_x = 4$  and  $n_u = n_d = n_y = 2$ . Applying our machinery on time horizon  $N = 256$  to build a purified-output-based linear controller with memory depth (whatever it means) 16, we end up with controller with disturbance-to-state peak-to-peak gain  $\approx 1.02$ . To put this result into proper perspective, note that the matrix  $A$  of the model in question is only marginally stable (the corresponding spectral radius is 0.9995). As a result, although trivial—identically zero—control results in uniformly bounded in  $N$  peak-to-peak gain, this gain ( $\approx 12$ ) is more than 10 times larger than the gain of the computed controller. Sample trajectories of the system with and without control are presented in Figure 2.

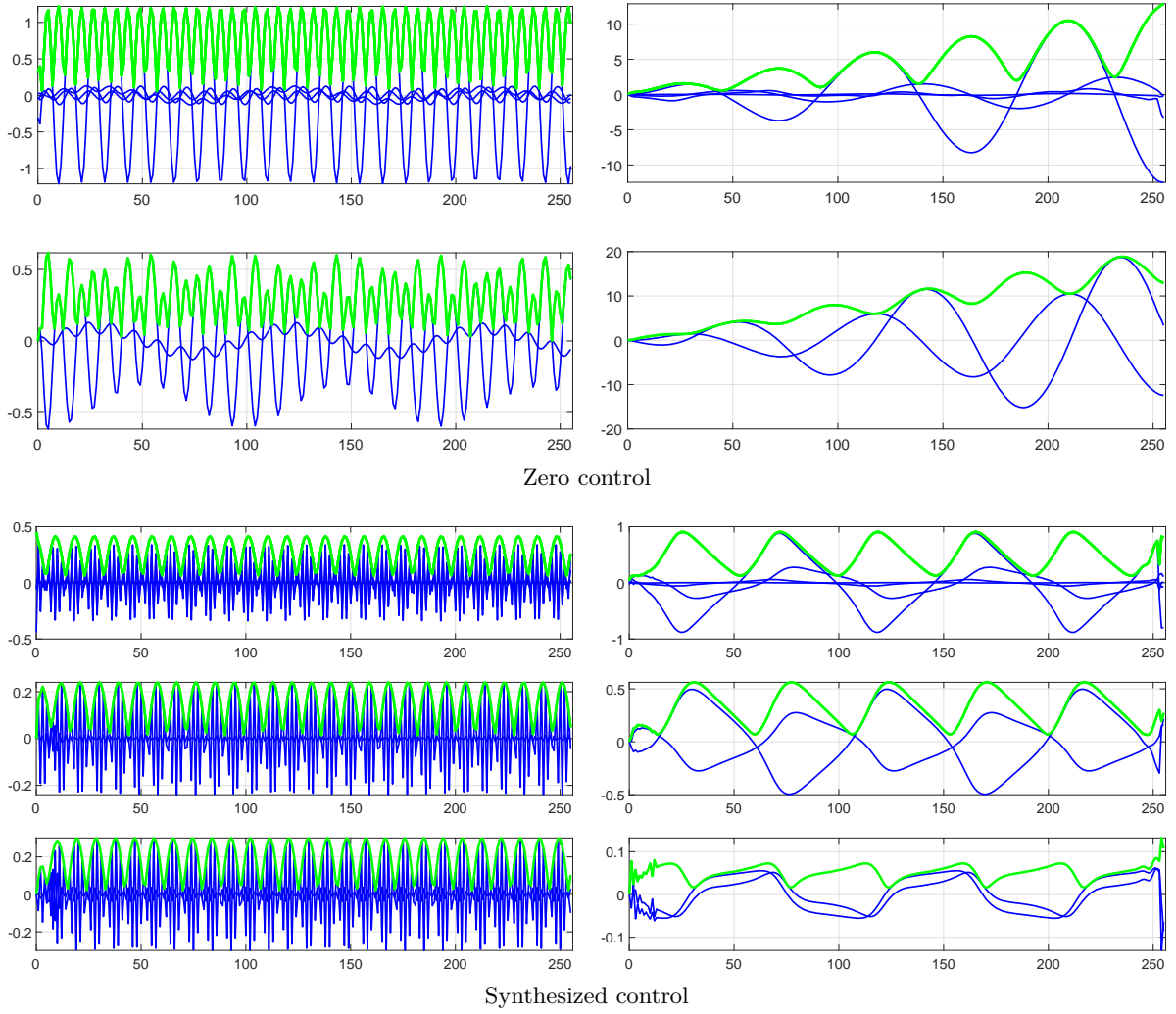
In the reported experiments,  $\|d_t\|_2 \equiv 1$  for all  $t$ . “Bad” disturbance is selected to result in large peak-to-peak gain with vanishing control; in this case  $\max_t \|x_t\|_2$  turns to be  $\approx 12$ , while with the control yielded by our synthesis, the same disturbances result in  $\max_t \|x_t\|_2 \approx 0.9$ , which is close to the upper bound on the gain ( $\approx 1.02$ ) guaranteed by our synthesis.

*Our second experiment* deals with toy dynamical system (3-dimensional states, scalar controls, disturbances, and outputs) considered for illustrative purposes in [11]; what we are interested in, are two scalar “performance outputs”  $z_t^{(1)}, z_t^{(2)}, t = 0, 1, \dots$ , given by

$$z_t^{(\chi)} = a_\chi^x x_t + a_\chi^y y_t + a_\chi^u u_t + a_\chi^d d_t, \quad t = 0, 1, \dots, \chi = 1, 2$$

where  $a_\chi^x, \dots, a_\chi^d$  are given, and  $x_t, y_t, u_t, d_t$  are, respectively, states, outputs, controls, and disturbances at time  $t$  (for detailed description, see [11, Section 5]). Following [11], we consider two infinite-horizon gains:  $\gamma^{(1)}$ —the norm of the induced by the linear controller in question mapping  $Z^{(1)}$  from the space of bounded sequences  $\{d_t, t \geq 0\}$  equipped with  $\ell_\infty$  norm to the space of sequences  $\{z_t^{(1)}, t \geq 0\}$  equipped with the same norm, and  $\gamma^{(2)}$ —the norm of mapping  $Z^{(2)}$  from the space of square-summable sequences  $\{d_t, t \geq 0\}$  to the space of sequences  $\{z_t^{(2)}, t \geq 0\}$ , both spaces equipped with  $\ell_2$  norm. The goal is to build linear controller resulting in as small as possible value of  $\gamma^{(1)}$  given an upper bound on  $\gamma^{(2)}$ . To process the problem, we select the “training horizon”  $T < \infty$  (we used  $T = 120$ ) and memory parameter  $m$  ( $m = 40$  in our experiments), and then optimize over the parameters of purified-output-based (POB) controller of memory depth  $m$  the “finite-horizon” approximations  $\gamma_T^{(\chi)}$  of  $\gamma^{(\chi)}$ ,  $\chi = 1, 2$ , defined as the respective operator norms of  $T \times T$  angular blocks in  $Z^{(\chi)}$ . After the control is synthesized, we estimate the “actual” values of  $\gamma^{(\chi)}$ ’s by computing the respective norms  $\gamma_{T'}^{(\chi)}$  of  $T' \times T'$

<sup>7</sup> MATLAB code for this experiment utilizing CVX [17] and MOSEK solver [2] is available at <https://github.com/ai1-fr/approximation-of-matrix-norms/tree/main/boeing>.



■ **Figure 2** In blue, from top to bottom: state ( $n_x = 4$ ), output ( $n_y = 2$ ) and control ( $n_u = 2$ , on the synthesized control plots) trajectories of the controlled plant. In the left pane: random harmonic oscillation disturbance, in the right pane: “bad disturbance”. In green:  $\|\cdot\|_2$ -norms of states, outputs and controls, respectively.

angular submatrices of the resulting matrices  $Z_\chi$  for  $T' \gg T$ .<sup>8</sup> Authors of [11] report on performance of 4 linear controllers obtained by 4 different synthesis methods. In order to make the comparison more informative, we used our machinery to design 4 POB controllers with upper bounds on  $\gamma^{(2)}$  close to the corresponding values given by designs in [11]. The benchmark results from [11] and our results are summarized in Table 1.

■ **Table 1** Value  $i$  in the first column corresponds to the row number in Table 1 of [11]; for each experiment we report the estimate of the gain obtained by our synthesis and (in “[.]”) the corresponding value from [11, Table 1].

$i$	$\gamma^{(1)}$	$\gamma^{(2)}$
1	24.29 [23.07]	21.35 [32.60]
2	73.93 [83.86]	10.52 [9.862]
3	33.22 [32.03]	14.97 [15.00]
4	33.22 [33.49]	14.97 [14.95]

Our tentative conclusion is that, as far as the reported data are concerned, our synthesis compares reasonably well to those discussed in [11].

<sup>8</sup> Indeed, there is clear numerical evidence of rapid stabilization of  $\gamma_{T'}^{(x)}$ 's as  $T'$  grows, and we have absolutely no doubt that the approximations  $\gamma_{4096}^{(x)}$  of  $\gamma^{(x)}$  we present below are accurate up to 4 significant digits.



## 4 Bounding robust norms of uncertain matrices

### 4.1 Motivation

Consider the following problem which arises, e.g., in Robust Control:

Given box-type uncertainty set

$$\mathcal{A}[\rho] = \left\{ A = \sum_{s=1}^S z_s A_s : \|z\|_\infty \leq \rho \right\}$$

in the space of  $m \times n$  matrices, upper-bound the quantity

$$\text{Opt}_*(\rho) = \max_{A \in \mathcal{A}[\rho]} |A|,$$

where  $|\cdot|$  stands for the spectral norm of a matrix.

This problem can be immediately reduced to the Matrix Cube problem (cf. [7], see also [6, Section 3.4.3.1]): associating with  $m \times n$  matrix  $A$  symmetric  $(m+n) \times (m+n)$  matrix

$$\mathcal{L}[A] = \left[ \begin{array}{c|c} \frac{1}{2}A^T & \frac{1}{2}A \\ \hline & \end{array} \right],$$

we have  $|A| \leq R$  if and only if  $RI_{m+n} - 2\mathcal{L}[A] \succeq 0$ . Therefore, the inequality

$$\text{Opt}_*(\rho) \leq R \tag{18}$$

is equivalent to

$$RI_{m+n} + 2 \sum_{s=1}^S z_s \mathcal{L}[A_s] \succeq 0 \quad \forall (z : \|z\|_\infty \leq \rho).$$

According to the results of [7], reproduced in [6, Theorem 3.4.7], an efficiently verifiable sufficient condition for the validity of the latter semi-infinite Linear Matrix Inequality (LMI) is the solvability of the parametric system of LMIs

$$RI_{m+n} - \rho \sum_{s=1}^S U_s \succeq 0, \quad U_s \succeq \pm 2\mathcal{L}[A_s], \quad 1 \leq s \leq S, \tag{18} \quad \mathcal{R}[R, \rho]$$

in matrix variables  $U_s$ , and this sufficient condition is tight within factor  $\vartheta(\cdot)$  depending solely of the maximum of ranks  $2\text{rank}(A_s)$  of the ‘‘edge matrices’’  $\mathcal{L}[A_s]$ . Specifically, when setting  $\mu = \max_{1 \leq s \leq S} \text{rank}(A_s)$ , we obtain:

- (18) does take place when  $\mathcal{R}[R, \rho]$  is feasible, and
- when  $\mathcal{R}[R, \rho]$  is infeasible, one has  $\text{Opt}(\vartheta(2\mu)\rho) > R$ , where  $\vartheta(\cdot)$  is a universal function specified in [7] (cf. [6, display (3.4.39)] and (21) below) such that

$$\vartheta(1) = 1, \vartheta(2) = \pi/2, \vartheta(3) = 1.7348\dots, \vartheta(4) = 2 \text{ \& } \vartheta(k) \leq \frac{1}{2}\pi\sqrt{k}, \quad k \geq 1. \tag{19}$$

The goal of this section is to extend this result onto more general matrix norms considered in Section 3.

### 4.2 Problem setting and main result

Let ellitopes  $\mathcal{X} \subset \mathbf{R}^n$ ,  $\mathcal{B}_* \subset \mathbf{R}^m$  with nonempty interior and basic ellitopes  $\mathcal{W}$ ,  $\mathcal{Z}$  be given by (6), let  $\mathcal{B}$  be the polar of  $\mathcal{B}_*$ , and let  $A_s \in \mathbf{R}^{m \times n}$ ,  $1 \leq s \leq S$ . These data define the uncertain matrix with box uncertainty

$$\mathcal{A} = \left\{ A = \sum_s \epsilon_s A_s : \|\epsilon\|_\infty \leq 1 \right\} \subset \mathbf{R}^{m \times n}$$

and the quantity

$$\|\mathcal{A}\|_{\mathcal{B}, \mathcal{X}} = \max_{A \in \mathcal{A}} \|A\|_{\mathcal{B}, \mathcal{X}}$$

which we refer to as *robust*  $\|\cdot\|_{\mathcal{B},\mathcal{X}}$ -norm of uncertain matrix  $\mathcal{A}$ . Note that this norm is difficult to compute already in the case of “general position” symmetric matrices  $A_s$  of rank 2. Our goal is to conceive a computationally efficient upper-bounding of the robust norm.

Let us consider the quantity

$$\varkappa(J) = \begin{cases} 1, & J = 1, \\ \frac{5}{2}\sqrt{\ln(2J)}, & J > 1, \end{cases} \quad (20)$$

and function  $\vartheta$  of the positive integer argument

$$\vartheta(k) = \left[ \min_{\alpha} \left\{ (2\pi)^{-k/2} \int |\alpha_1 u_1^2 + \dots + \alpha_k u_k^2| e^{-u^T u/2} du, \alpha \in \mathbf{R}^k, \|\alpha\|_1 = 1 \right\} \right]^{-1}; \quad (21)$$

note that  $\vartheta(k)$  satisfies (19) [7]. Let also

$$\text{Opt} = \min_{\substack{\lambda \geq 0, v \geq 0, \\ \{G_s, H_s\}}} \left\{ \phi_{\mathcal{T}}(\lambda) + \phi_{\mathcal{R}}(v) : \left[ \begin{array}{c|c} G_s & \frac{1}{2}Q^T A_s P \\ \hline \frac{1}{2}P^T A_s^T Q & H_s \end{array} \right] \succeq 0, s \leq S \right. \\ \left. \sum_s G_s \preceq \sum_{\ell} v_{\ell} R_{\ell}, \sum_s H_s \preceq \sum_k \lambda_k T_k \right\}. \quad (22)$$

► **Proposition 5.** *In the situation of this section, assuming that ranks of all  $A_s$  are  $\leq \kappa$ , the efficiently computable quantity Opt as given by (22) is a reasonably tight upper bound on the robust norm  $\|\mathcal{A}\|_{\mathcal{B},\mathcal{X}}$  of uncertain matrix  $\mathcal{A}$ , specifically,*

$$\|\mathcal{A}\|_{\mathcal{B},\mathcal{X}} \leq \text{Opt} \leq \varkappa(K)\varkappa(L)\vartheta(2\kappa)\|\mathcal{A}\|_{\mathcal{B},\mathcal{X}} \quad (23)$$

where  $K$  and  $L$  are given by (6).

► **Remark 6.** Assume that matrices  $A_s = A_s[\chi]$  are affine in some vector  $\chi$  of control parameters. In this case, the quantity  $\|\mathcal{A}\|_{\mathcal{B},\mathcal{X}}$  and its efficiently computable upper bound Opt become functions  $\text{Opt}_*(\chi) = \|\mathcal{A}\|_{\mathcal{B},\mathcal{X}}$  and  $\text{Opt}(\chi)$  of  $\chi$ , and it is immediately seen that both functions are convex. As a result, we can handle, to some extent, the problem of minimizing over  $\chi$  the robust  $\|\cdot\|$ -norm of uncertain matrix

$$\mathcal{A}[\chi] = \left\{ A = \sum_s \epsilon_s A_s[\chi] : \|\epsilon\|_{\infty} \leq 1 \right\}.$$

More precisely, we can minimize over  $\chi$  efficiently computable convex upper bound  $\text{Opt}(\chi)$  on the robust norm  $\text{Opt}_*(\chi)$  of  $\mathcal{A}[\chi]$ , the bound being reasonably tight provided that the ranks of matrices  $A_s[\chi]$  are small for all  $\chi$  in question.

► **Remark 7.** Note that the quantity

$$\text{Opt} = \min_{\substack{\lambda \geq 0, v \geq 0, \\ \{G_s, H_s\}}} \left\{ \phi_{\mathcal{T}}(\lambda) + \phi_{\mathcal{R}}(v) : \left[ \begin{array}{c|c} G_s & \frac{1}{2}Q^T A_s P \\ \hline \frac{1}{2}P^T A_s^T Q & H_s \end{array} \right] \succeq 0, s \leq S \right. \\ \left. \sum_s G_s \preceq \sum_{\ell} v_{\ell} R_{\ell}, \sum_s H_s \preceq \sum_k \lambda_k T_k \right\}$$

as given by (22) admits another representation which may sometimes be more convenient. Specifically, excluding trivial case  $\text{Opt} = 0$  which takes place if and only if  $Q^T A_s P = 0$  for all  $s$ , one has

$$\frac{1}{\text{Opt}} = \max_{\rho, \{G_s, H_s\}, \lambda, v} \left\{ \rho : \left[ \begin{array}{c|c} G_s & \rho Q^T A_s P \\ \hline \rho P^T A_s^T Q & H_s \end{array} \right] \succeq 0, s \leq S \right\}. \quad (24)$$

Indeed, the optimization problem specifying Opt clearly is solvable; let  $\lambda, v, \{G_s, H_s\}$  be its optimal solution. Looking at the problem, we see, first, that  $\text{Opt} > 0$  implies  $\lambda \neq 0$  and  $v \neq 0$ , and thus  $\phi_{\mathcal{R}}(v) > 0$  and  $\phi_{\mathcal{T}}(\lambda) > 0$ . Furthermore, whenever  $\theta > 0$ , the collection  $\theta^{-1}\lambda, \theta v, \{\theta G_s, \theta^{-1}H_s\}$  is a feasible solution with the value of the objective  $\theta\phi_{\mathcal{R}}(v) + \theta^{-1}\phi_{\mathcal{T}}(\lambda)$ . Since the solution we have started with is optimal, we have

$$\theta\phi_{\mathcal{R}}(v) + \theta^{-1}\phi_{\mathcal{T}}(\lambda) \geq \phi_{\mathcal{R}}(v) + \phi_{\mathcal{T}}(\lambda) = \text{Opt}.$$

This inequality holds true for all  $\theta > 0$ , which with positive  $\phi_{\mathcal{R}}(\lambda)$  and  $\phi_{\mathcal{T}}(\lambda)$  is possible if and only if  $\phi_{\mathcal{R}}(v) = \phi_{\mathcal{T}}(\lambda) = \text{Opt}/2$ . It follows that setting

$$\bar{\lambda} = 2\lambda/\text{Opt}, \quad \bar{v} = 2v/\text{Opt}, \quad \bar{G}_s = 2G_s/\text{Opt}, \quad \bar{H}_s = 2H_s/\text{Opt}, \quad \rho = 1/\text{Opt},$$

we get a feasible solution to (24) with the value of the objective  $1/\text{Opt}$ , implying that the left hand side in (24) is  $\leq$  the right hand side. On the other hand, the optimization problem in (24) clearly is solvable. If  $\rho, \lambda, v, \{G_s, H_s\}$  is an optimal solution to (24) then

$$\bar{G}_s = G_s/(2\rho), \quad \bar{H}_s = H_s/(2\rho), \quad \bar{\lambda} = \lambda/(2\rho), \quad \bar{v} = v/(2\rho)$$

clearly form a feasible solution to the problem specifying  $\text{Opt}$ , and the value of the objective of the latter problem at this solution is  $\leq 1/\rho$ . Thus,  $\text{Opt} \leq 1/\rho$ ,  $\rho$  being the optimal value of the optimization problem in (24), so that the left hand side in (24) is  $\geq$  the right hand side.

### 4.2.1 An extension

Similarly to what was done in Section 3.2, the above results can be straightforwardly extended to the case when  $\|\cdot\|_{\mathcal{X}}$  is elliptic, and  $\|\cdot\|_{\mathcal{B}}$  is co-elliptic norm. Specifically, for an uncertain matrix

$$\mathcal{A} = \left\{ \sum_s \epsilon_s A_s : \|\epsilon\|_{\infty} \leq 1 \right\},$$

the robust norm of  $\mathcal{A}$  in the case of (10), (12) is

$$\max_{i,j} \left\| \left\{ \sum_s \epsilon_s Q_j^T A_s P_i : \|\epsilon\|_{\infty} \leq 1 \right\} \right\|_{\mathcal{Z}_j^*, \mathcal{X}_i}$$

where  $\mathcal{X}_i$  and polars  $\mathcal{Z}_j^*$  of  $\mathcal{Z}_j$  are ellitopes, and we know how to efficiently upper-bound the robust norms  $\|\{\sum_s \epsilon_s Q_j^T A_s P_i : \|\epsilon\|_{\infty} \leq 1\}\|_{\mathcal{Z}_j^*, \mathcal{X}_i}$  and how tight such bounds are.

### 4.2.2 Putting things together

So far, we have considered separately computationally efficient bounding of operator norms of matrices and robust norms of uncertain matrices with box uncertainty. In applications to follow, we will be interested in a “mixed” setting, where we want to upper-bound the robust norm

$$\|\mathcal{U}\|_{\mathcal{B}, \mathcal{X}} = \max_{A \in \mathcal{U}} \|A\|_{\mathcal{B}, \mathcal{X}}$$

of uncertain matrix

$$\mathcal{U} = A_{\text{nom}} + \mathcal{A}, \quad \mathcal{A} = \left\{ \sum_{s=1}^S \epsilon_s A_s : \|\epsilon\|_{\infty} \leq 1 \right\}. \quad (25)$$

The corresponding blend of our preceding results is as follows:

► **Proposition 8.** *Let  $\mathcal{X} \subset \mathbf{R}^n$ ,  $\mathcal{B}, \mathcal{B}_* \subset \mathbf{R}^m$  be given by (10), (12), with basic ellitopes*

$$\begin{aligned} \mathcal{X}_i &= \{x_i \in \mathbf{R}^{\nu_i} : \exists t^i \in \mathcal{T}^i : x_i^T T_{ki} x_i \leq t_k^i, 1 \leq k \leq K_i\}, \quad i \leq I \\ \mathcal{Z}_j &= \{z_j \in \mathbf{R}^{\mu_j} : \exists s^j \in \mathcal{R}^j : z_j^T R_{\ell j} z_j \leq s_{\ell}^j, 1 \leq \ell \leq L_j\}, \quad j \leq J \end{aligned}$$

Then the efficiently computable quantity

$$\text{Opt}[\mathcal{U}] = \max_{i \leq I, j \leq J} \text{Opt}_{ij}[\mathcal{U}],$$

where

$$\text{Opt}_{ij}[\mathcal{U}] = \min_{\substack{\lambda^{ij}, v^{ij}, \bar{G}^{ijs}, \bar{H}^{ij} \\ 1 \leq i \leq I, 1 \leq j \leq J, 1 \leq s \leq S}} \left\{ \phi_{\mathcal{T}^i}(\lambda^{ij}) + \phi_{\mathcal{R}^j}(v^{ij}) : \begin{array}{l} \lambda^{ij} \geq 0, v^{ij} \geq 0 \\ \sum_{s=1}^S G^{ijs} + \bar{G}^{ij} \leq \sum_{\ell=1}^{L_j} v_\ell^{ij} R_{\ell j}, \\ \sum_{s=1}^S H^{ijs} + \bar{H}^{ij} \leq \sum_{k=1}^{K_i} \lambda_k^{ij} T_{ki} \\ \left[ \begin{array}{c|c} G^{ijs} & \frac{1}{2}[Q_j^T A_s P_i] \\ \hline \frac{1}{2}[Q_j^T A_s P_i]^T & H^{ijs} \end{array} \right] \succeq 0, s \leq S \\ \left[ \begin{array}{c|c} \bar{G}^{ij} & \frac{1}{2}[Q_j^T A_{\text{nom}} P_i] \\ \hline \frac{1}{2}[Q_j^T A_{\text{nom}} P_i]^T & \bar{H}^{ij} \end{array} \right] \succeq 0 \end{array} \right\} \quad (26)$$

is an efficiently computable convex in  $(A_{\text{nom}}, A_1, \dots, A_S)$  upper bound on  $\|\mathcal{U}\|_{\mathcal{B}, \mathcal{X}}$ . This upper bound is reasonably tight, specifically, setting

$$\mathcal{U}_{ij} = Q_j^T A_{\text{nom}} P_i + \left\{ \sum_{s=1}^S \epsilon_s [Q_j^T A_s P_i] : \|\epsilon\|_\infty \leq 1 \right\},$$

we have

$$\|\mathcal{U}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i} \leq \text{Opt}_{ij}[\mathcal{U}] \leq [\zeta(K_i, L_j) + \varkappa(K_i) \varkappa(L_j) \vartheta(2\kappa)] \|\mathcal{U}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i},$$

and

$$\begin{aligned} \|\mathcal{U}\|_{\mathcal{B}, \mathcal{X}} &= \max_{i \leq I, j \leq J} \|\mathcal{U}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i} \leq \text{Opt}[\mathcal{U}] = \max_{i \leq I, j \leq J} \text{Opt}_{ij}[\mathcal{U}] \\ &\leq \left[ \max_{i \leq I, j \leq J} [\zeta(K_i, L_j) + \varkappa(K_i) \varkappa(L_j) \vartheta(2\kappa)] \right] \|\mathcal{U}\|_{\mathcal{B}, \mathcal{X}} \end{aligned} \quad (27)$$

where  $\kappa$  is the maximum of ranks of  $A_s$ ,  $1 \leq s \leq S$ , and  $\zeta(K, L)$  and  $\varkappa(\cdot), \vartheta(\cdot)$  are as defined in Theorem 3 and Proposition 5.

Note that “extreme cases” ( $A_s = 0$  for all  $s$ , on one hand, and  $A_{\text{nom}} = 0$ , on the other) of Proposition 8 recover Theorem 3 and Proposition 5, and even their “advanced” versions with simple ellitopic/co-ellitopic norms extended to ellitopic/co-ellitopic ones.

### 4.3 Application to robust signal recovery

Consider the standard Signal Processing problem as follows. *Given noisy observations*

$$\omega = Ax + \xi, \quad \xi \sim \mathcal{N}(0, I_m) \quad (28)$$

of unknown signal  $x$  known to belong to a given signal set  $\mathcal{X} \subset \mathbf{R}^n$ , we want to recover  $Bx \in \mathbf{R}^\nu$ . Here  $A \in \mathbf{R}^{m \times n}$  and  $B \in \mathbf{R}^{\nu \times n}$  are given matrices. We consider linear recovery  $\hat{x} = \hat{x}_H(\omega) := H^T \omega$ ,  $H \in \mathbf{R}^{m \times \nu}$  and quantify the performance of a candidate estimate  $\hat{x}_H$  by its *worst-case risk*

$$\text{Risk}_{\|\cdot\|_{\mathcal{B}}}[\hat{x}_H | \mathcal{X}] = \sup_{x \in \mathcal{X}} \mathbf{E}_{\xi \sim \mathcal{N}(0, I_m)} \{ \|Bx - \hat{x}_H(Ax + \xi)\|_{\mathcal{B}} \},$$

where  $\|\cdot\|_{\mathcal{B}}$  is a given norm on  $\mathbf{R}^\nu$ . There is an extensive literature dealing with the design and performance analysis of linear estimates. In particular, it is known [21, Proposition 4.16] that when  $\mathcal{X}$  is an ellitope of ellitopic size  $K$  and the polar  $\mathcal{B}_*$  of the unit ball  $\mathcal{B}$  of  $\|\cdot\|_{\mathcal{B}}$  is an ellitope of ellitopic size  $L$ , the linear estimate  $\hat{x}_{H_*}$  yielded by the optimal solution to an explicit efficiently solvable convex optimization problem is optimal within logarithmic in  $K, L$  factor:

$$\text{Risk}_{\|\cdot\|_{\mathcal{B}}}[\hat{x}_{H_*} | \mathcal{X}] \leq O(1) \sqrt{\ln(2K) \ln(2L)} \text{Risk}_{\text{Opt}}_{\|\cdot\|_{\mathcal{B}}}[\mathcal{X}];$$

here  $\text{Risk}_{\text{Opt}}_{\|\cdot\|_{\mathcal{B}}}[\mathcal{X}]$  is the minimax risk—the infimum of risks  $\text{Risk}_{\|\cdot\|_{\mathcal{B}}}[\hat{x} | \mathcal{X}]$  over all estimates  $\hat{x}$ , linear and nonlinear alike.

The result we have just cited, as well as most of known to us results on performance of linear estimates, deals with the case when the sensing matrix  $A$  is known in advance. Here we want to address the case when  $A$  is

subject to “uncertain-but-bounded” perturbations, specifically, is selected (by nature or by an adversary) from the uncertainty set

$$\mathcal{U} = A_{\text{nom}} + \mathcal{A}, \quad \mathcal{A} = \left\{ \sum_{s=1}^S \epsilon_s A_s : \|\epsilon\|_{\infty} \leq 1 \right\}.$$

This problem can be seen as a “noninterval” extension of the problem of solving systems of equations affected by interval uncertainty which has received significant attention in the literature, cf., e.g., [14, 18, 23, 31, 36, 37, 38] and references therein. Assuming that given perturbation in  $A$  and “true” signal  $x$ , the observation noise  $\xi$  is  $\mathcal{N}(0, I_m)$ , the worst-case risk of a linear estimate  $\hat{x}_H$  becomes

$$\begin{aligned} \text{Risk}_{\|\cdot\|_{\mathcal{B}}}^+[\hat{x}_H|\mathcal{X}] &:= \sup_{\substack{x \in \mathcal{X}, \epsilon: \\ \|\epsilon\|_{\infty} \leq 1}} \mathbf{E}_{\xi \sim \mathcal{N}(0, I_m)} \left\{ \left\| [B - H^T A_{\text{nom}}]x - \left[ \sum_s \epsilon_s H^T A_s \right]x - H^T \xi \right\|_{\mathcal{B}} \right\} \\ &\leq \|\mathcal{V}[H]\|_{\mathcal{B}, \mathcal{X}} + \mathbf{E}_{\xi \sim \mathcal{N}(0, I_m)} \{ \|H^T \xi\|_{\mathcal{B}} \} \end{aligned}$$

where

$$\mathcal{V}[H] = \left\{ [B - H^T A_{\text{nom}}] + \sum_{s=1}^S \epsilon_s H^T A_s : \|\epsilon\|_{\infty} \leq 1 \right\}$$

and  $\|\mathcal{V}[H]\|_{\mathcal{B}, \mathcal{X}} = \max_{V \in \mathcal{V}[H]} \|V\|_{\mathcal{B}, \mathcal{X}}$ . The simplest way to build a “presumably good” linear estimate is to minimize over  $H$  the sum of the (efficiently computable upper bound on the) robust norm of  $\mathcal{V}[H]$  and an efficiently computable upper bound on  $\Psi(H) := \mathbf{E}_{\xi \sim \mathcal{N}(0, I_m)} \{ \|H^T \xi\|_{\mathcal{B}} \}$ . Combining the results of Proposition 8 with the upper bound on  $\Psi(H)$  from [21, Lemma 4.11], in the case of  $\mathcal{X} = P\mathcal{X}_1$ ,  $\mathcal{B}_* = Q\mathcal{Z}_1$ ,  $I = J = 1$ , we arrive at the efficiently solvable convex optimization problem

$$\begin{aligned} \text{Opt} = \min_{\substack{H, \lambda, v, \mu, \\ \Theta, G^s, H^s, G, H}} & \left\{ \phi_{\mathcal{T}}(\lambda) + \phi_{\mathcal{R}}(v) + \phi_{\mathcal{R}}(\mu) + \text{Tr}(\Theta) : \right. \\ & \left. \begin{aligned} & \lambda \geq 0, v \geq 0, \mu \geq 0, G + \sum_s G^s \preceq \sum_{\ell} v_{\ell} R_{\ell}, H + \sum_s H^s \preceq \sum_k \lambda_k T_k \\ & \left[ \begin{array}{c|c} \Theta & \frac{1}{2} H Q \\ \hline \frac{1}{2} Q^T H^T & \sum_{\ell} \mu_{\ell} R_{\ell} \end{array} \right] \succeq 0, \left[ \begin{array}{c|c} G^s & \frac{1}{2} Q^T H^T A_s P \\ \hline \frac{1}{2} P^T A_s^T H Q & H^s \end{array} \right] \succeq 0, s \leq S \\ & \left[ \begin{array}{c|c} G & \frac{1}{2} Q^T [B - H^T A_{\text{nom}}] P \\ \hline \frac{1}{2} P^T [B^T - A_{\text{nom}}^T H] Q & H \end{array} \right] \succeq 0 \end{aligned} \right\} \quad (29) \end{aligned}$$

(we use the notation from Proposition 8 with  $\nu$  in the role of  $m$ ). For every feasible solution to this problem, the value of the objective at the solution is an upper bound on  $\text{Risk}_{\|\cdot\|_{\mathcal{B}}}^+[\hat{x}_H|\mathcal{X}]$ ,  $H$  being the  $H$ -component of the solution in question. Moreover, from Proposition 8 combined with [21, Lemma 4.11] it follows that the function  $\text{Opt}[H]$  obtained by partial minimization of the objective in (29) over all decision variables except  $H$  is a tight, within factor  $O(1)\sqrt{\ln(2K)\ln(2L)\vartheta(2\kappa)}$ , upper bound on  $\text{Risk}_{\|\cdot\|_{\mathcal{B}}}^+[\hat{x}_H|\mathcal{X}]$ ; here  $\kappa = \min[m, \nu, \max_s \text{rank}(A_s)]$ . In particular, linear estimate  $\hat{x}_{H^*}$  yielded by an optimal solution to (29) is optimal within the above factor, in terms of its risk  $\text{Risk}_{\|\cdot\|_{\mathcal{B}}}^+[\cdot|\mathcal{X}]$ , among all linear estimates. Finally, when there is no uncertainty ( $A_s = 0$  for all  $s$ ),  $\hat{x}_{H^*}$  is exactly the near-minimax-optimal estimate from [21, Proposition 4.16].

### 4.3.1 System identification: the problem

Consider situation as follows: a linear time-invariant dynamical system with states  $u_t \in \mathbf{R}^d$  and inputs  $r_t \in \mathbf{R}^h$  evolves according to

$$u_{t+1} = X[u_t; r_t]. \quad (30)$$

We are given noisy observations  $\bar{u}_t$  of the states and of the inputs on time horizon  $0 \leq t \leq N$ :

$$\bar{u}_{ti} = u_{ti} - \xi_{ti}, \quad 0 \leq t \leq N, 1 \leq i \leq d; \quad \bar{u}_{ti} = r_{t, i-d} - \xi_{ti}, \quad 0 \leq t < N, d+1 \leq i \leq d+h; \quad (31)$$

we also have at our disposal upper bounds on the magnitudes of observation errors:

$$|\xi_{tj}| \leq \bar{\xi}_{tj}$$

with known  $\bar{\xi}$ 's. In addition, we have partial a priori knowledge of  $X$  expressed by a system of linear equations on the entries of  $X$ . Our goal is to recover the image  $X^+$  of  $X$  under a given linear mapping.

Observe that the considered setting is rather different from the ‘‘classical’’ setting of linear system identification problem, cf. [3, 16, 27, 29, 40], in which it is assumed that the states of the system are observed without errors, and the errors in observations of inputs are corrupted by random zero mean noise. The situation in which perturbations in the observation of the state of the system are *uncertain-but-bounded* (e.g., belong to an ellipsoid) is the subject of the significant literature (see, e.g., [8, 12, 13, 19, 24, 25, 26, 28, 30, 32, 39, 43] and references therein). The ‘‘generic’’ approach to the problem we develop below, to the best of our knowledge, differs significantly from those proposed so far, and, we believe, can be considered as a meaningful contribution to the this line of research.

Assigning the entries of  $X$  serial indices, denoting by  $\iota(i, j)$  the index of  $X_{ij}$  and setting  $x_{\iota(i, j)}^* = X_{ij}$ , we get  $n$ -dimensional vector  $x^*$ ,  $n = d(d + h)$ , known to satisfy the system of linear equations

$$Px = p \tag{32}$$

( $P \in \mathbf{R}^{\nu \times n}$  has linearly independent rows), expressing our a priori knowledge of the actual entries of  $X$ . Dynamic equations read

$$\bar{u}_{t+1, i} = \sum_{j=1}^{d+h} \bar{u}_{tj} x_{\iota(i, j)}^* - \xi_{t+1, i} + \sum_{j=1}^{d+h} \xi_{tj} x_{\iota(i, j)}^*, \quad 1 \leq i \leq d, \quad 0 \leq t \leq N - 1, \tag{!ti}$$

which we rewrite as a system of linear equations on  $x^*$  of the form

$$Qx - \sum_{s=1}^S \zeta_s Q_s x = q - \sum_{s=1}^S \zeta_s q_s$$

where  $S = N(d + h) + d$  is the total count of observation errors  $\xi_{tj}$ ,  $\zeta_1, \dots, \zeta_S$  are these errors written down in certain order,  $Q$  and  $Q_s$  are observable  $m \times n$  matrices,  $m = dN$ , and  $q, q_s$  are observable  $m$ -dimensional vectors. Note that each matrix  $Q_s$  has at most  $d + 1$  nonzero rows. Indeed, observation error  $\xi_{tj}$  with  $j \leq d$  participates only in equation (! $_{t-1, j}$ ) (this happens when  $t \geq 1$ ) and  $d$  equations (! $_{ti}$ ),  $1 \leq i \leq d$ , and observation error  $\xi_{tj}$  with  $j > d$  participates only in  $d$  equations (! $_{ti}$ ),  $1 \leq i \leq d$ . Setting

$$\mathcal{L} = \{x \in \mathbf{R}^n : Px = 0\}, \quad \bar{x} = (PP^T)^{-1}P^T p, \quad \Pi = I_n - P^T(PP^T)^{-1}P,$$

so that  $\Pi$  is an orthoprojector of  $\mathbf{R}^n$  onto  $\mathcal{L}$  and  $\bar{x}$  is the orthogonal projection of  $x^*$  onto the orthogonal complement of  $\mathcal{L}$ , we have

$$x^* = \bar{x} + \Delta^*$$

with  $\Delta^*$  satisfying the relations

$$\Delta^* \in \mathcal{L},$$

$$\exists (\epsilon^* \in \mathbf{R}^S, \|\epsilon^*\|_\infty \leq 1) : Q[\bar{x} + \Delta^*] - \left[ \sum_{s=1}^S \epsilon_s^* \underbrace{\bar{\zeta}_s Q_s}_{\bar{Q}_s} \right] [\bar{x} + \Delta^*] = q - \sum_{s=1}^S \epsilon_s^* \bar{\zeta}_s q_s.$$

Thus,  $x^* = \bar{x} + \Delta^*$ , where  $\Delta^*$  solves, for properly selected vector  $\epsilon = \epsilon^* \in \mathbf{R}^S$ ,  $\|\epsilon^*\|_\infty \leq 1$ , the system of linear equations

$$\begin{aligned} \left[ Q - \sum_s \epsilon_s \bar{Q}_s \right] \Delta &= \left[ \bar{q} + \sum_s \epsilon_s \bar{q}_s \right] \& \Pi \Delta = \Delta \\ \left[ \bar{q} = q - Q\bar{x}, \bar{q}_s = \bar{Q}_s \bar{x} - \bar{\zeta}_s q_s \right] \end{aligned} \tag{33}$$

in variables  $\Delta \in \mathbf{R}^n$ .

Recall that our goal is to recover from observation the image of  $X$  under a given linear mapping; this is the same as to recover

$$y^* = Bx^* = \underbrace{B\bar{x}}_{\bar{y}} + \underbrace{B\Delta^*}_{\delta^*}$$

for a given  $\nu \times n$  matrix  $B$ . Let us quantify the recovery error by the norm  $\|\cdot\|_B$  on  $\mathbf{R}^\nu$ .

### 4.3.2 Robust linear recovery

Given  $m \times n$  matrix  $E$  and  $m \times \nu$  matrix  $H$ , let us recover

■  $\Delta^*$  by the vector

$$\hat{\Delta} := \Pi E^T \bar{q} = \left[ \Pi E^T Q - \sum_s \epsilon_s^* \Pi E^T \bar{Q}_s \right] \Delta^* - \sum_s \epsilon_s^* \Pi E^T \bar{q}_s$$

and  $x^*$ —by the vector  $\bar{x} + \hat{\Delta}$ ,

■  $\delta^*$  by the vector

$$\hat{\delta} := H^T \bar{q} = \left[ H^T Q - \sum_s \epsilon_s^* H^T \bar{Q}_s \right] \Delta^* - \sum_s \epsilon_s^* H^T \bar{q}_s$$

and  $y^*$ —by  $\bar{y} + \hat{\delta}$ .

### Performance analysis

By (33) we have

$$\bar{q} = \left[ Q - \sum_s \epsilon_s^* \bar{Q}_s \right] \Delta^* - \sum_s \epsilon_s^* \bar{q}_s.$$

Thus,  $\hat{\Delta} \in \mathcal{L}$ ,  $\Delta^* \in \mathcal{L}$ , and

$$\begin{aligned} \hat{\Delta} - \Delta^* &= \left[ \Pi E^T Q - I_n - \sum_s \epsilon_s^* \Pi E^T \bar{Q}_s \right] \Delta^* - \left[ \sum_s \epsilon_s^* \Pi E^T \bar{q}_s \right] \\ &= \left[ \Pi [E^T Q - I_n] \Pi - \sum_s \epsilon_s^* \Pi E^T \bar{Q}_s \Pi \right] \Delta^* - \left[ \sum_s \epsilon_s^* \Pi E^T \bar{q}_s \right] \end{aligned} \quad (34)$$

where the concluding equality is due to  $\Delta^* = \Pi \Delta^*$  and  $\Pi^2 = \Pi$ . Besides this,

$$\begin{aligned} \hat{\delta} - \delta^* &= \left[ H^T Q - B - \sum_s \epsilon_s^* H^T \bar{Q}_s \right] \Delta^* - \left[ \sum_s \epsilon_s^* H^T \bar{q}_s \right] \\ &= \left[ [H^T Q - B] \Pi - \sum_s \epsilon_s^* H^T \bar{Q}_s \Pi \right] \Delta^* - \left[ \sum_s \epsilon_s^* H^T \bar{q}_s \right] \end{aligned} \quad (35)$$

Now let  $\mathcal{X}$  be the unit ball of a norm  $\|\cdot\|_{\mathcal{X}}$  on  $\mathbf{R}^n$ ; assume that this norm is both elliptic and co-elliptic. Let

$$\begin{aligned} \mathcal{W}_0[E] &= \left\{ \sum_s \epsilon_s \Pi E^T \bar{q}_s : \|\epsilon\|_\infty \leq 1 \right\} \subset \mathbf{R}^n, \\ \mathcal{W}[E] &= \left\{ \Pi [E^T Q - I_n] \Pi - \sum_s \epsilon_s \Pi E^T \bar{Q}_s \Pi : \|\epsilon\|_\infty \leq 1 \right\}, \end{aligned}$$

and let  $\Upsilon_0[E]$  and  $\Upsilon[E]$  be the efficiently computable convex in  $E$  upper bounds, given by our machinery, on the robust norms

$$\|\mathcal{W}_0[E]\|_{\mathcal{X}, [-1,1]} = \max_w \{ \|w\|_{\mathcal{X}} : w \in \mathcal{W}_0[E] \}, \quad \|\mathcal{W}[E]\|_{\mathcal{X}, \mathcal{X}} = \max_W \{ \|W\|_{\mathcal{X}, \mathcal{X}} : W \in \mathcal{W}[E] \}$$



of the uncertain  $n \times 1$  matrix  $\mathcal{W}_0[E]$  and uncertain  $n \times n$  matrix  $\mathcal{W}[E]$ . By (34) we have

$$\|\widehat{\Delta} - \Delta^*\|_{\mathcal{X}} \leq \Upsilon[E]\|\Delta^*\|_{\mathcal{X}} + \Upsilon_0[E]. \quad (36)$$

Assume from now on that  $\|\cdot\|_{\mathcal{B}}$  is a co-elliptic norm, let

$$\mathcal{V}_0[H] = \left\{ \sum_s \epsilon_s H^T \bar{q}_s : \|\epsilon\|_{\infty} \leq 1 \right\} \subset \mathbf{R}^{n\nu}, \quad \mathcal{V}[H] = \left\{ [H^T Q - B]\Pi - \sum_s \epsilon_s H^T \bar{Q}_s \Pi : \|\epsilon\|_{\infty} \leq 1 \right\}$$

and let  $\bar{\Upsilon}_0[H]$ ,  $\bar{\Upsilon}[H]$  be the efficiently computable convex in  $H$  upper bounds, given by our machinery, on the robust norms

$$\|\mathcal{V}_0[H]\|_{\mathcal{B},[-1,1]} = \max_w \{\|w\|_{\mathcal{B}} : w \in \mathcal{V}_0[H]\}, \quad \|\mathcal{V}[H]\|_{\mathcal{B},\mathcal{X}} = \max_W \{\|W\|_{\mathcal{B},\mathcal{X}} : W \in \mathcal{V}[H]\}$$

of the uncertain  $\nu \times 1$  matrix  $\mathcal{V}_0[H]$  and uncertain  $\nu \times n$  matrix  $\mathcal{V}[H]$ . By (35) we have

$$\|\widehat{\delta} - \delta^*\|_{\mathcal{B}} \leq \bar{\Upsilon}[H]\|\Delta^*\|_{\mathcal{X}} + \bar{\Upsilon}_0[H]. \quad (37)$$

Assume now that  $E$  is such that  $\Upsilon[E] < 1$ . Then

$$\|\Delta^*\|_{\mathcal{X}} \leq \|\widehat{\Delta} - \Delta^*\|_{\mathcal{X}} + \|\widehat{\Delta}\|_{\mathcal{X}} \leq \Upsilon[E]\|\Delta^*\|_{\mathcal{X}} + [\|\widehat{\Delta}\|_{\mathcal{X}} + \Upsilon_0[E]]$$

whence

$$\|\Delta^*\|_{\mathcal{X}} \leq \frac{\|\widehat{\Delta}\|_{\mathcal{X}} + \Upsilon_0[E]}{1 - \Upsilon[E]}. \quad (38)$$

As a result,

$$\begin{aligned} \|\widehat{x} - x^*\|_{\mathcal{X}} &= \|\widehat{\Delta} - \Delta^*\|_{\mathcal{X}} \leq \frac{\Upsilon[E]}{1 - \Upsilon[E]} [\|\widehat{\Delta}\|_{\mathcal{X}} + \Upsilon_0[E]] + \Upsilon_0[E] \\ &= \frac{\Upsilon[E]}{1 - \Upsilon[E]} [\|\Pi E^T \bar{q}\|_{\mathcal{X}} + \Upsilon_0[E]] + \Upsilon_0[E] \end{aligned} \quad (39a)$$

$$\begin{aligned} \|\widehat{y} - y^*\|_{\mathcal{B}} &= \|\widehat{\delta} - \delta^*\|_{\mathcal{B}} \leq \frac{\bar{\Upsilon}[H]}{1 - \Upsilon[E]} [\|\widehat{\Delta}\|_{\mathcal{X}} + \Upsilon_0[E]] + \bar{\Upsilon}_0[H] \\ &= \frac{\bar{\Upsilon}[H]}{1 - \Upsilon[E]} [\|\Pi E^T \bar{q}\|_{\mathcal{X}} + \Upsilon_0[E]] + \bar{\Upsilon}_0[H] \end{aligned} \quad (39b)$$

## Synthesis of the linear estimate

Recall that the problem of minimizing  $\Upsilon[E]$  w.r.t.  $E$  is efficiently solvable. If we are lucky to have  $\Upsilon_* := \inf_E \Upsilon[E] < 1$ , we can optimize, to some extent, our estimate  $H^T \bar{q}$  of  $y^* = Bx^*$  in  $H$ . To this end, let us select  $E$  which “nearly minimizes” the quantity

$$\Gamma = \frac{1}{1 - \Upsilon[E]} [\|\Pi E^T \bar{q}\|_{\mathcal{X}} + \Upsilon_0[E]]$$

over  $E$  under the constraint  $\Upsilon(E) < 1$ ; after  $E$  is selected, we specify  $H$  by minimizing the resulting right hand side of (39b), that is,  $\Gamma \bar{\Upsilon}(H) + \bar{\Upsilon}_0[H]$  in  $H$ .

“Near-minimization” of  $\Gamma$  over  $E$  can be carried out as follows. Let us select somehow  $\beta < 1$  close to 1 (e.g.,  $\beta = 0.9$  or  $\beta = 0.99$ ) and set  $\Upsilon_i = (1 - \beta^i) + \beta^i \Upsilon_*$ ,  $i = 0, 1, 2, \dots$ , so that  $\frac{\beta^i}{1 - \Upsilon} \leq \frac{1}{1 - \Upsilon_*}$  is equivalent to  $\Upsilon \leq \Upsilon_i$ . We solve one by one feasible convex optimization problems

$$\text{Opt}_i = \frac{1}{1 - \Upsilon_i} \min_E \left\{ \|\Pi E^T \bar{q}\|_{\mathcal{X}} + \Upsilon_0[E] : \Upsilon[E] \leq \Upsilon_i \right\}, \quad i = 0, 1, \dots, \quad (P_i)$$

we run this process until the quantities  $\text{Opt}_i$  start to grow, and specify  $\Gamma$  as the smallest of  $\text{Opt}_i$  we have generated.

Let us write explicitly the problem  $(P_i)$  in the situation where

$$\mathcal{X} = \text{Conv} \left\{ \bigcup_{k \leq K} P_k \mathcal{B}_{n_k} \right\} \quad (40)$$

where  $\mathcal{B}_m$  is the unit  $\|\cdot\|_2$ -ball in  $\mathbf{R}^m$  and  $P_k \in \mathbf{R}^{n \times n_k}$ . As we know, in this case

$$\begin{aligned} \|x\|_{\mathcal{X}} &= \min_{x_k \in \mathbf{R}^{n_k}, k \leq K} \left\{ \sum_k \|x_k\|_2 : \sum_k P_k x_k = x \right\}, \\ \|x\|_{\mathcal{X}_*} &= \max_{k \leq K} \|P_k^T x\|_2, \\ \|A\|_{\mathcal{X}, \mathcal{X}} &= \max_{k \leq K} \|A P_k\|_{\mathcal{X}, \mathcal{B}_{n_k}}. \end{aligned} \quad (41)$$

Exploiting the fact that in our present situation  $\mathcal{X}_* = \{x : \|P_k^T x\|_2 \leq 1, k \leq K\}$  is an ellitope,  $(P_i)$  may be rewritten as follows (cf. (26) in Proposition 8):

$$\begin{aligned} \|\Pi E^T \bar{q}\|_{\mathcal{X}} &= \min_{\{x_k, k \leq K\}} \left\{ \sum_k \|x_k\|_2 : \sum_k P_k x_k = \Pi E^T \bar{q} \right\}; \\ \Upsilon_0[E] &= \min_{\substack{\{G_s \in \mathbf{R}^n, H_s \in \mathbf{R}, s \leq S\}, \\ v \in \mathbf{R}^K, \lambda \in \mathbf{R}}} \left\{ \frac{1}{2} \left[ \sum_k v_k + \lambda \right] : \begin{array}{l} v \geq 0, \sum_s G_s \preceq \sum_k v_k P_k P_k^T, \sum_s H_s \leq \lambda \\ \left[ \begin{array}{c|c} G_s & \Pi E^T \bar{q}_s \\ \hline \bar{q}_s^T E \Pi & H_s \end{array} \right] \succeq 0 \end{array} \right\} \\ &= \min_{\substack{\{H_s \in \mathbf{R}, s \leq S\}, \\ v \in \mathbf{R}^K, \lambda \in \mathbf{R}}} \left\{ \frac{1}{2} \left[ \sum_k v_k + \lambda \right] : \begin{array}{l} v \geq 0, \sum_s H_s \leq \lambda \\ \sum_s H_s^{-1} [\Pi E^T \bar{q}_s] [\Pi E^T \bar{q}_s]^T \preceq \sum_k v_k P_k P_k^T \end{array} \right\} \\ &= \min_{v \in \mathbf{R}^K, \mu \in \mathbf{R}^S} \left\{ \frac{1}{2} \left[ \sum_k v_k + \sum_s \mu_s \right] : v \geq 0, \left[ \begin{array}{c|c} \sum_k v_k P_k P_k^T & \Pi E^T [\bar{q}_1, \dots, \bar{q}_S] \\ \hline [\bar{q}_1, \dots, \bar{q}_S]^T E \Pi & \text{Diag}\{\mu\} \end{array} \right] \succeq 0 \right\}; \\ \Upsilon[E] &= \min_{\substack{\{G_s^k, H_s^k, \bar{G}^k, \bar{H}^k : k \leq K, s \leq S\}, \\ \{v^k, \lambda^k : k \leq K\}}} \left\{ \frac{1}{2} \max_{k \leq K} \left[ \sum_{j=1}^K v_j^k + \lambda^k \right] : \right. \\ &\quad \left. \begin{array}{l} v^k \geq 0, \bar{G}^k + \sum_s G_s^k \leq \sum_{j=1}^K v_j^k P_j P_j^T, \bar{H}^k + \sum_s H_s^k \leq \lambda^k I_{n_k} \\ \left[ \begin{array}{c|c} \bar{G}^k & \Pi[E^T Q - I_n] \Pi P_k \\ \hline P_k^T \Pi [Q^T E - I_n] \Pi & \bar{H}^k \end{array} \right] \succeq 0, \\ \left[ \begin{array}{c|c} G_s^k & \Pi E^T \bar{Q}_s \Pi P_k \\ \hline P_k^T \Pi \bar{Q}_s^T E \Pi & H_s^k \end{array} \right] \succeq 0, k \leq K, s \leq S \end{array} \right\}, \end{aligned}$$

and

$$\begin{aligned} \text{Opt}_i &= \frac{1}{1 - \Upsilon_i} \min_{\substack{E, \{x_k, k \leq K\}, v \in \mathbf{R}^K, \mu \in \mathbf{R}^S, \\ \{G_s^k, H_s^k, \bar{G}^k, \bar{H}^k, k \leq K, s \leq S\}, \\ \{v^k \in \mathbf{R}^K, \lambda^k, k \leq K\}}} \left\{ \sum_k \|x_k\|_2 + \frac{1}{2} \left[ \sum_k v_k + \sum_s \mu_s \right] : \right. \\ &\quad \left. \begin{array}{l} \sum_k P_k x_k = \Pi E^T \bar{q} \\ v \geq 0, \left[ \begin{array}{c|c} \sum_k v_k P_k P_k^T & \Pi E^T [\bar{q}_1, \dots, \bar{q}_S] \\ \hline [\bar{q}_1, \dots, \bar{q}_S]^T E \Pi & \text{Diag}\{\mu\} \end{array} \right] \succeq 0 \\ v^k \geq 0, \bar{G}^k + \sum_s G_s^k \leq \sum_{j=1}^K v_j^k P_j P_j^T, \bar{H}^k + \sum_s H_s^k \leq \lambda^k I_{n_k}, k \leq K, \\ \left[ \begin{array}{c|c} \bar{G}^k & \Pi[E^T Q - I_n] \Pi P_k \\ \hline P_k^T \Pi [Q^T E - I_n] \Pi & \bar{H}^k \end{array} \right] \succeq 0 \\ \left[ \begin{array}{c|c} G_s^k & \Pi E^T \bar{Q}_s \Pi P_k \\ \hline P_k^T \Pi \bar{Q}_s^T E \Pi & H_s^k \end{array} \right] \succeq 0, k \leq K, s \leq S \\ \frac{1}{2} \max_{k \leq K} [\sum_j v_j^k + \lambda^k] \leq \Upsilon_i \end{array} \right\}. \end{aligned}$$

► **Remark 9.** Rationale behind restricting ourselves to  $\mathcal{X}$  as in (40) is as follows. Recall that the norm  $\|\cdot\|_{\mathcal{X}}$  we consider is assumed to be both ellitopic and co-ellitopic. There are only two known to us generic situations in which the corresponding unit ball  $\mathcal{X}$  is both ellitopic and co-ellitopic at the same time, and (40) is one of them. The other nice situation which is “symmetric” to the first, is when  $\|\cdot\|_{\mathcal{X}}$  is the conjugate of the norm just defined, that is, norm of the form  $\max_{k \leq K} \|P_k^T x\|_2$ . In our context, this second case reduces to the first due to  $\|A\|_{\mathcal{X}, \mathcal{X}} = \|A^T\|_{\mathcal{X}_*, \mathcal{X}_*}$ .

## Numerical illustration

to follow deals with recovery of the parameters of the “Boeing 747” model used in Section 3.3.3, which in our present notation reads

$$u_{t+1} = \begin{bmatrix} 0.9957 & 0.0339 & -0.0211 & -0.3214 & 0.0140 & 0.9886 & 0.0043 & -0.0337 \\ 0.0076 & 0.4699 & 4.6604 & 0.0022 & -3.4373 & 1.6648 & -0.0079 & 0.5285 \\ 0.0168 & -0.0605 & 0.4038 & -0.0029 & -0.8219 & 0.4378 & -0.0167 & 0.0600 \\ 0.0091 & -0.0370 & 0.7194 & 0.9990 & -0.4735 & 0.2491 & -0.0091 & 0.0370 \end{bmatrix} [u_t; r_t]$$

where  $u_t \in \mathbf{R}^4$  are the states, and  $r_t \in \mathbf{R}^4$  are the inputs (“in reality” the first two entries in  $r_t$  are controls, and the last two—external disturbances). We observe  $u_t$ ’s for  $0 \leq t \leq N = 12$  and  $r_t$ ’s for  $0 \leq t < N$ ; in the resulting identification problem,  $m = 48$ ,  $n = 32$ ,  $S = 100$ , and  $\mathcal{L} = \mathbf{R}^n$  (whence  $\Pi = I_n$  and  $\bar{x} = 0$ ). Observations of states and inputs are corrupted by “relative  $\epsilon$ -noises”, so that an observable real  $r$  and its observation  $\bar{r}$  satisfy  $|r - \bar{r}| \leq \epsilon \max[|\bar{r}|, 1]$ . In an experiment, we select a noise level  $\epsilon \in [0.001, 0.01]$ , generate a sample trajectory of the system by selecting at random the initial state and the inputs, then add to the states and the inputs random  $\epsilon$ -errors, and apply to the resulting observations the above robust linear recovery with  $B = I_n$  and  $\mathcal{B} = \mathcal{X}$  being the unit  $\|\cdot\|_2$ -ball in  $\mathbf{R}^n$  to recover the parameters of the system. We compare this recovery with the simplest Least Squares estimation  $H_{LS}^T q = \operatorname{argmin}_x \|Qx - q\|_2^2$ ,  $H_{LS} = Q(Q^T Q)^{-1}$ .

The results of a series of 10 experiments are presented in Table 2.<sup>9</sup> In Figure 3, we present the trajectories of the actual and the recovered (experiment # 10,  $\epsilon = 0.01$ ) systems on time horizon  $1 \leq t \leq 49$  for random initial state and inputs (different from those used in the experiment).

■ **Table 2**  $\|\cdot\|_2$ -errors of recovery (first number in a cell) and ratios of upper bounds by (39b) on these errors to the actual errors (second number in a cell); median values for 10 simulations per each value of noise level  $\epsilon$ .

$\epsilon$	0.001	0.002	0.003	0.004	0.005
Least Squares	0.017/12.1	0.035/10.0	0.071/12.6	0.078/11.0	0.061/12.1
Linear Recovery	0.021/1.9	0.042/1.5	0.085/1.6	0.096/1.5	0.074/1.7
$\epsilon$	0.006	0.007	0.008	0.009	0.010
Least Squares	0.118/12.2	0.124/11.4	0.123/ 10.4	0.124/12.1	0.145/13.2
Linear Recovery	0.131/1.4	0.163/1.4	0.162/1.6	0.157/1.8	0.189/1.6

## A Proofs

### A.1 Proof of Theorem 3

The below proof follows that of Theorem 2 as given in [21, Section 4.8.2], utilizing at some point bilinearity of the quadratic form we want to upper-bound on  $\mathcal{Z} \times \mathcal{W}$ .

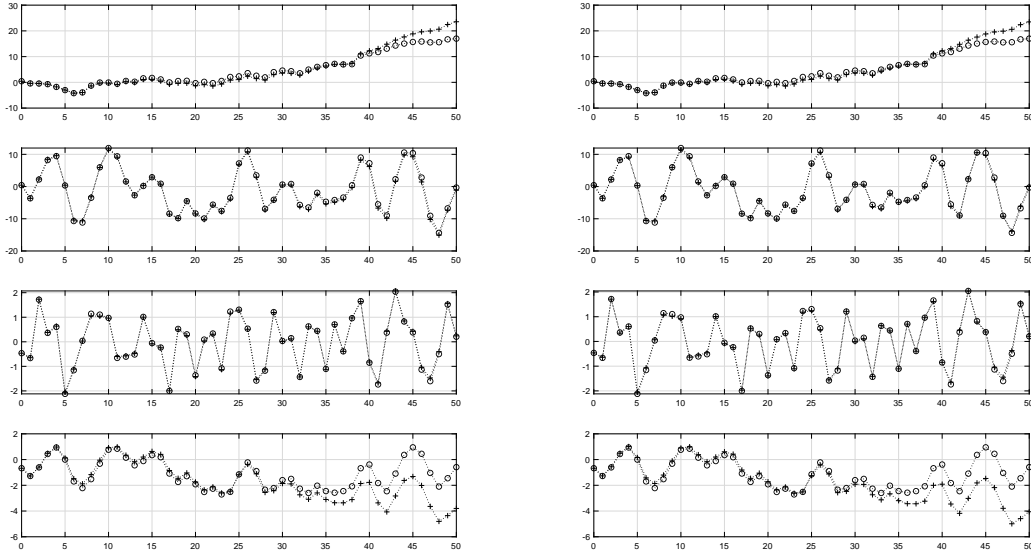
Let  $q, p$  be the dimensions of the embedding spaces of  $\mathcal{Z}$  and  $\mathcal{W}$ , and assume w.l.o.g. that  $q \leq p$ .<sup>10</sup>

1°. Let

$$\mathfrak{T} = \operatorname{cl}\{[t; \tau] : \tau > 0, t/\tau \in \mathcal{T}\} \quad \text{and} \quad \mathfrak{R} = \operatorname{cl}\{[r; \theta] : \theta > 0, r/\theta \in \mathcal{R}\}$$

<sup>9</sup> MATLAB code for this experiment is available at <https://github.com/ai1-fr/approximation-of-matrix-norms/tree/main/ident>.

<sup>10</sup> It is immediately seen that the norm bound (7) is intelligent enough to respect the identity  $\|A\|_{\mathcal{B}, \mathcal{X}} = \|A^T\|_{\mathcal{X}_*, \mathcal{B}_*}$  where  $\mathcal{Q}_*$  stands for the polar of a set  $\mathcal{Q}$ . As a result, to ensure  $q \leq p$ , we can pass, if necessary, from  $\mathcal{B}, \mathcal{X}$  and  $A$  to  $\mathcal{X}_*, \mathcal{B}_*$  and  $A^T$ .



■ **Figure 3** States of the actual (marked “o”) and the recovered (marked “+”) systems vs. time. Left: Least Squares recovery; right: Robust recovery. Maximum relative magnitude of observation errors 0.01. Actual values of  $\|\cdot\|_2$ -errors of recovery are 0.1372 for LS and 0.1279 for Robust recovery.

be the closed conic hulls of  $\mathcal{T}$  and  $\mathcal{R}$ , so that  $\mathfrak{T}$  and  $\mathfrak{R}$  are regular (closed, pointed and convex with nonempty interior) cones such that

$$\mathcal{T} = \{t : [t; 1] \in \mathfrak{T}\}, \quad \mathcal{R} = \{r : [r; 1] \in \mathfrak{R}\}.$$

As is immediately seen, the cones dual to  $\mathfrak{T}$ ,  $\mathfrak{R}$  are

$$\mathfrak{T}_* = \{[g; \tau] : \tau \geq \phi_{\mathcal{T}}(-g)\}, \quad \mathfrak{R}_* = \{[h; \theta] : \theta \geq \phi_{\mathcal{R}}(-h)\}.$$

In view of these observations, (7) is nothing but the conic problem

$$\text{Opt}(A) = \min_{\lambda, v, \tau, \theta} \left\{ \tau + \theta : \begin{array}{l} \lambda \geq 0, v \geq 0, [-\lambda; \tau] \in \mathfrak{T}_*, [-v; \theta] \in \mathfrak{R}_*, \\ \left[ \begin{array}{c|c} \sum_{\ell} v_{\ell} R_{\ell} & \frac{1}{2} Q^T A P \\ \hline \frac{1}{2} P^T A^T Q & \sum_k \lambda_k T_k \end{array} \right] \succeq 0 \end{array} \right\}.$$

It is easily seen that this problem is strictly feasible and bounded. By Conic Duality,

$$\begin{aligned} \text{Opt}(A) &= \max_{r, t, U, V, W} \left\{ \text{Tr}(W^T Q^T A P) : \begin{array}{l} t \in \mathcal{T}, r \in \mathcal{R} \\ \text{Tr}(R_{\ell} U) \leq r_{\ell} \quad \forall \ell, \text{Tr}(T_k V) \leq t_k \quad \forall k \\ \left[ \begin{array}{c|c} U & W \\ \hline W^T & V \end{array} \right] \succeq 0 \end{array} \right\} \\ &= \max_{r, t, U, V, Y} \left\{ \text{Tr}([U^{1/2} Y V^{1/2}]^T Q^T A P) : \begin{array}{l} r \in \mathcal{R}, t \in \mathcal{T}, U \succeq 0, V \succeq 0, Y^T Y \preceq I \\ \text{Tr}(R_{\ell} U) \leq r_{\ell} \quad \forall \ell, \text{Tr}(T_k V) \leq t_k \quad \forall k \end{array} \right\} \\ &= \max_{r, t, U, V} \left\{ \sum_{i=1}^q \sigma_i(U^{1/2} Q^T A P V^{1/2}) : \begin{array}{l} U \succeq 0, \text{Tr}(R_{\ell} U) \leq r_{\ell} \quad \forall \ell, r \in \mathcal{R} \\ V \succeq 0, \text{Tr}(T_k V) \leq t_k \quad \forall k, t \in \mathcal{T} \end{array} \right\} \end{aligned}$$

where  $\sigma_i(\cdot)$ ,  $i \leq q$ , are the singular values of  $q \times p$  matrix (recall that  $q \leq p$ ). At the last two steps of the above derivation, we have used the following well known facts

■  $\left[ \begin{array}{c|c} U & W \\ \hline W^T & V \end{array} \right] \succeq 0$  if and only if  $U \succeq 0, V \succeq 0$  and  $W = U^{1/2} Y V^{1/2}$  with  $Y^T Y \preceq I$ ,

and

■ the maximum of Frobenius inner products of a given matrix with matrices of spectral norm not exceeding 1 is the nuclear norm of the matrix—the sum of singular values.

**2°.** The concluding optimization problem in the above chain clearly is solvable; let  $U, V, r, t$  be the optimal solution, and let  $\sigma_i = \sigma_i(U^{1/2} Q^T A P V^{1/2})$ , and  $\sum_{i=1}^q \sigma_i e_i f_i^T$  be the singular value decomposition of  $U^{1/2} Q^T A P V^{1/2}$ , so

that

$$\begin{aligned} \text{Opt}(A) &= \sum_{\iota=1}^q \sigma_{\iota}, \\ U^{1/2} Q^T A P V^{1/2} &= \sum_{\iota=1}^q \sigma_{\iota} e_{\iota} f_{\iota}^T, \\ e_i^T e_j &= \begin{cases} 1, & i = j \\ 0, & i \neq j, \end{cases} \quad i, j \leq q \quad \& \quad f_i^T f_j = \begin{cases} 1, & i = j \\ 0, & i \neq j, \end{cases} \quad i, j \leq p. \end{aligned} \quad (42)$$

Let  $\epsilon_1, \dots, \epsilon_p$  be independent random variables taking values  $\pm 1$  with probabilities  $1/2$ , and let

$$\xi = \sum_{i=1}^q \epsilon_i e_i, \quad \eta = \sum_{j=1}^p \epsilon_j f_j.$$

Then in view of (42) it holds, identically in  $\epsilon_i = \pm 1$ ,  $1 \leq i \leq p$ :

$$\xi^T U^{1/2} Q^T A P V^{1/2} \eta = \sum_{i, \iota \leq q, j \leq p} [\epsilon_i \epsilon_j \sigma_{\iota} e_i^T e_{\iota} f_{\iota}^T f_j] = \sum_{\iota=1}^q \sigma_{\iota} = \text{Opt}(A). \quad (43)$$

On the other hand, setting  $E = [e_1, \dots, e_q]$ , we get an orthonormal  $q \times q$  matrix such that  $\xi = E \underline{\epsilon}$ , where  $\underline{\epsilon} = [\epsilon_1; \dots; \epsilon_q]$  is a Rademacher vector (i.e., random vector with independent entries taking values  $\pm 1$  with probabilities  $1/2$ ), and

$$\xi^T U^{1/2} R_{\ell} U^{1/2} \xi = \underline{\epsilon}^T \underbrace{[E^T U^{1/2} R_{\ell} U^{1/2} E]}_{\bar{R}_{\ell}} \underline{\epsilon}$$

By construction,  $\bar{R}_{\ell} \succeq 0$  and

$$\text{Tr}(\bar{R}_{\ell}) = \text{Tr}(U^{1/2} R_{\ell} U^{1/2}) = \text{Tr}(R_{\ell} U) \leq r_{\ell}.$$

For every  $\ell$  such that  $r_{\ell} > 0$  we have  $\text{Tr}(r_{\ell}^{-1} \bar{R}_{\ell}) \leq 1$ . Now let us use the following fact.

► **Lemma 10** ([21, Lemma 4.48]). *Let  $Q$  be positive semidefinite  $N \times N$  matrix with trace  $\leq 1$  and  $\zeta$  be  $N$ -dimensional Rademacher random vector. Then*

$$\mathbf{E} \left\{ \exp \left\{ \frac{1}{3} \zeta^T Q \zeta \right\} \right\} \leq \sqrt{3}.$$

By Lemma 10, whenever  $r_{\ell} > 0$  we have

$$\mathbf{E} \{ \exp \{ \xi^T [r_{\ell}^{-1} U^{1/2} R_{\ell} U^{1/2}] \xi / 3 \} \} = \mathbf{E} \{ \exp \{ \underline{\epsilon}^T [r_{\ell}^{-1} \bar{R}_{\ell}] \underline{\epsilon} / 3 \} \} \leq \sqrt{3}.$$

As a result, for every  $\ell$  such that  $r_{\ell} > 0$  we have

$$\text{Prob} \{ \xi^T U^{1/2} R_{\ell} U^{1/2} \xi > 3 \ln(4L) r_{\ell} \} < 1/(2L).$$

The latter relation holds true for those  $\ell$  for which  $r_{\ell} = 0$  as well, since for these  $\ell$  one has  $U^{1/2} R_{\ell} U^{1/2} = 0$  because trace of the latter positive semidefinite matrix is  $\leq r_{\ell}$ . Similar reasoning with  $\bar{\epsilon} = [\epsilon_1; \dots; \epsilon_p]$  in the role of  $\underline{\epsilon}$  and  $T_k$ ,  $t_k$  in the roles of  $R_{\ell}$ ,  $r_{\ell}$  demonstrates that for every  $k$  we have

$$\text{Prob} \{ \eta^T V^{1/2} T_k V^{1/2} \eta > 3 \ln(4K) t_k \} < 1/(2K).$$

Consequently, invoking (43), we conclude that there exists realization  $(\bar{\xi}, \bar{\eta})$  of  $(\xi, \eta)$  such that

$$\bar{\xi}^T U^{1/2} Q^T A P V^{1/2} \bar{\eta} = \text{Opt}(A),$$

and

$$\bar{\xi}^T U^{1/2} R_{\ell} U^{1/2} \bar{\xi} \leq 3 \ln(4L) r_{\ell} \quad \forall \ell, \quad \bar{\eta}^T V^{1/2} T_k V^{1/2} \bar{\eta} \leq 3 \ln(4K) t_k \quad \forall k.$$

Setting  $v = QU^{1/2}\bar{\xi}$ ,  $x = PV^{1/2}\bar{\eta}$  and invoking (6), we get  $\|x\|_{\mathcal{X}} \leq \sqrt{3\ln(4K)}$ ,  $\|v\|_{\mathcal{B}_*} \leq \sqrt{3\ln(4L)}$ , resulting in

$$\text{Opt}(A) = \bar{\xi}^T U^{1/2} Q^T A P V^{1/2} \bar{\eta} = v^T A x \leq \|x\|_{\mathcal{X}} \|v\|_{\mathcal{B}_*} \|A\|_{\mathcal{B}, \mathcal{X}},$$

that is,

$$\text{Opt}(A) \leq 3\sqrt{\ln(4K)\ln(4L)} \|A\|_{\mathcal{B}, \mathcal{X}},$$

as claimed.

3°. It remains to consider the case of  $K = L = 1$ . By evident scaling argument, the situation reduces to that where  $\mathcal{X} = P\{w : w^T T w \leq 1\}$  and  $\mathcal{B}_* = Q\{z : z^T S z \leq 1\}$ . In this case,

$$\begin{aligned} \|A\|_{\mathcal{B}, \mathcal{X}} &= \max_{\substack{z: z^T S z \leq 1, \\ w: w^T T w \leq 1}} z^T [Q^T A P] w = \max_{\substack{\zeta: \|\zeta\|_2 \leq 1 \\ \omega: \|\omega\|_2 \leq 1}} \omega^T [S^{-1/2} Q^T A P T^{-1/2}] \zeta \\ &= \min_{\nu} \left\{ \sqrt{\nu} : \left[ \frac{\nu I_q}{[S^{-1/2} Q^T A P T^{-1/2}]^T} \middle| \frac{[S^{-1/2} Q^T A P T^{-1/2}]}{I_p} \right] \succeq 0 \right\}. \end{aligned}$$

On the other hand,

$$\begin{aligned} \text{Opt}(A) &= \min_{\lambda, v} \left\{ \lambda + v : \left[ \frac{v S}{\frac{1}{2} [Q^T A P]^T} \middle| \frac{\frac{1}{2} [Q^T A P]}{\lambda T} \right] \succeq 0 \right\} \\ &= \min_{\lambda, v} \left\{ \frac{1}{2} [\lambda + v] : \left[ \frac{v I_q}{[S^{-1/2} Q^T A P T^{-1/2}]^T} \middle| \frac{[S^{-1/2} Q^T A P T^{-1/2}]}{\lambda I_p} \right] \succeq 0 \right\} \\ &= \min_{\lambda \geq 0, v \geq 0} \left\{ \frac{1}{2} [\lambda + v] : \left[ \frac{v I_q}{[S^{-1/2} Q^T A P T^{-1/2}]^T} \middle| \frac{[S^{-1/2} Q^T A P T^{-1/2}]}{I_p} \right] \succeq 0 \right\} \\ &= \min_{\lambda \geq 0, v \geq 0, \nu} \left\{ \frac{1}{2} [\lambda + v] : v \lambda \geq \nu, \left[ \frac{\nu I_q}{[S^{-1/2} Q^T A P T^{-1/2}]^T} \middle| \frac{[S^{-1/2} Q^T A P T^{-1/2}]}{I_p} \right] \succeq 0 \right\} \\ &= \min_{\nu} \left\{ \sqrt{\nu} : \left[ \frac{\nu I_q}{[S^{-1/2} Q^T A P T^{-1/2}]^T} \middle| \frac{[S^{-1/2} Q^T A P T^{-1/2}]}{I_p} \right] \succeq 0 \right\} = \|A\|_{\mathcal{B}, \mathcal{X}}. \end{aligned}$$

## A.2 Proof of Proposition 5

1°. Let  $\mathfrak{R}$ ,  $\mathfrak{T}$ ,  $\mathfrak{R}_*$  and  $\mathfrak{T}_*$  be as defined in item 1° of the proof of Theorem 3. Observe that

$$\begin{aligned} \text{Opt} &= \min_{\lambda, v, G_s, H_s, \alpha, \beta} \left\{ \alpha + \beta : \left[ \frac{G_s}{\frac{1}{2} P^T A_s^T Q} \middle| \frac{\frac{1}{2} Q^T A_s P}{H_s} \right] \succeq 0 \forall s \leq S, [-v; \alpha] \in \mathfrak{R}_*, [-\lambda; \beta] \in \mathfrak{T}_* \right\} \\ &= \max_{Y, X, W_s, r, t} \left\{ \sum_s \text{Tr}(W_s^T Q^T A_s P) : \left[ \frac{Y}{W_s^T} \middle| \frac{W_s}{X} \right] \succeq 0 \forall s \leq S, t \in \mathcal{T}, r \in \mathcal{R} \right. \\ &\quad \left. \text{Tr}(Y R_\ell) \leq r_\ell, \ell \leq L, \text{Tr}(X T_k) \leq t_k, k \leq K \right\} \\ &\quad [\text{by conic duality}] \\ &= \max_{Y, X, r, t} \left\{ \sum_s \|\sigma(Y^{1/2} Q^T A_s P X^{1/2})\|_1 : \begin{array}{l} Y \succeq 0, X \succeq 0, t \in \mathcal{T}, r \in \mathcal{R}, \\ \text{Tr}(Y R_\ell) \leq r_\ell, \ell \leq L, \text{Tr}(X T_k) \leq t_k, k \leq K \end{array} \right\} \end{aligned}$$

where  $\sigma(A)$  is the singular spectrum of  $A$ ; the last equality in the chain follows from the two simple observations (cf. the proof of Theorem 3):

- LMI  $\left[ \frac{P}{Q^T} \middle| \frac{Q}{R} \right] \succeq 0$  with  $p \times p$  matrix  $P$  and  $r \times r$  matrix  $R$  takes place if and only if  $P \succeq 0$ ,  $R \succeq 0$ , and  $Q = P^{1/2} Y R^{1/2}$  with  $p \times r$  matrix  $Y$  such that  $Y^T Y \preceq I_r$ , and
- for  $p \times r$  matrix  $A$ , one has  $\max_Y \{\text{Tr}(Y^T A) : Y \in \mathbf{R}^{p \times r}, Y^T Y \preceq I_r\} = \|\sigma(A)\|_1$

With  $\mathcal{L}[B] = \left[ \frac{P}{\frac{1}{2} B^T} \middle| \frac{\frac{1}{2} B}{X} \right]$ , the nonzero eigenvalues of  $2\mathcal{L}[B]$  are exactly plus and minus nonzero singular values of  $B$ , and we conclude that

$$\text{Opt} = \max_{Y, X, r, t} \left\{ \sum_s \|\lambda(\mathcal{L}[Y^{1/2} Q^T A_s P X^{1/2}])\|_1 : \begin{array}{l} Y \in \mathbf{S}_+^q, X \in \mathbf{S}_+^p, t \in \mathcal{T}, r \in \mathcal{R} \\ \text{Tr}(Y R_\ell) \leq r_\ell, \ell \leq L, \\ \text{Tr}(X T_k) \leq t_k, k \leq K \end{array} \right\}, \quad (44)$$

where  $\lambda(A)$  is the vector of eigenvalues of a symmetric matrix  $A$ .

Note that  $\text{Opt}$  as defined in (44) clearly is a convex function of  $[A_1, \dots, A_S]$ .

Observe that  $\|\mathcal{A}\|_{\mathcal{B}, \mathcal{X}} \leq \text{Opt}$ . Indeed, the problem specifying  $\text{Opt}$  clearly is solvable, and if  $\lambda \geq 0, v \geq 0, \{G_s, H_s\}$  is its optimal solution, we have for all  $z \in \mathcal{Z}, w \in \mathcal{W}, \epsilon_s = \pm 1$  :

$$\epsilon_s z^T Q^T A_s P w \leq z^T G_s z + w^T H_s w.$$

Thus,

$$\sum_s \epsilon_s z^T Q^T A_s P w \leq z^T \left[ \sum_\ell v_\ell R_\ell \right] z + w^T \left[ \sum_k \lambda_k T_k \right] w \leq \max_{r \in \mathcal{R}, t \in \mathcal{T}} [v^T r + \lambda^T t] \leq \phi_{\mathcal{R}}(v) + \phi_{\mathcal{T}}(\lambda) = \text{Opt}$$

for all  $w \in \mathcal{W}, z \in \mathcal{Z}$ , and all  $\epsilon_s = \pm 1$ , implying that  $\|\mathcal{A}\|_{\mathcal{B}, \mathcal{X}} \leq \text{Opt}$  (recall that  $P\mathcal{W} = \mathcal{X}$  and  $Q\mathcal{Z} = \mathcal{B}_*$ ).

2°. Now, let  $X \succeq 0, Y \succeq 0, t, r$  be such that  $t \in \mathcal{T}, r \in \mathcal{R}, \text{Tr}(Y R_\ell) \leq r_\ell, \ell \leq L, \text{Tr}(X T_k) \leq t_k, k \leq K$ , and

$$\text{Opt} = \sum_s \|\lambda(\mathcal{L}[Y^{1/2} Q^T A_s P X^{1/2}])\|_1.$$

By [7, Lemma 2.3] (cf. [6, Lemma 3.4.3]), if the ranks of all matrices  $A_s$  (and thus—matrices  $Q^T A_s P$ ) do not exceed a given  $\kappa$ , which we assume from now on, then for  $\omega \sim \mathcal{N}(0, I_{m+n})$  one has

$$\mathbf{E} \left\{ |\omega^T \mathcal{L}[Y^{1/2} Q^T A_s P X^{1/2}] \omega| \right\} \geq \|\lambda(\mathcal{L}[Y^{1/2} Q^T A_s P X^{1/2}])\|_1 / \vartheta(2\kappa),$$

where  $\vartheta(k)$  is defined in (21). It follows that for  $[\eta; \xi] \sim \mathcal{N}(0, \text{Diag}\{Y, X\})$ ,

$$\text{Opt} \leq \vartheta(2\kappa) \mathbf{E} \left\{ \sum_s |\omega^T \mathcal{L}[Y^{1/2} Q^T A_s P X^{1/2}] \omega| \right\} = \vartheta(2\kappa) \mathbf{E} \left\{ \sum_s |\eta^T Q^T A_s P \xi| \right\}.$$

Now, let  $\pi(\cdot)$  be the norm on  $\mathbf{R}^p$  with the unit ball  $\mathcal{W}$ , and  $\rho(\cdot)$  be the norm on  $\mathbf{R}^q$  with the unit ball  $\mathcal{Z}$ . Taking into account that  $\mathcal{X} = P\mathcal{W}$  and  $\mathcal{B}_* = Q\mathcal{Z}$  we conclude that

$$\forall (\eta \in \mathbf{R}^q, \xi \in \mathbf{R}^p) : \sum_s |\eta^T Q^T A_s P \xi| = \max_{\epsilon_s = \pm 1} \eta^T Q^T \left[ \sum_s \epsilon_s A_s \right] P \xi \leq \rho(\eta) \pi(\xi) \|\mathcal{A}\|_{\mathcal{B}, \mathcal{X}},$$

thus arriving at

$$\text{Opt} \leq \vartheta(2\kappa) \|\mathcal{A}\|_{\mathcal{B}, \mathcal{X}} \mathbf{E} \{ \rho(\eta) \pi(\xi) \} = \vartheta(2\kappa) \|\mathcal{A}\|_{\mathcal{B}, \mathcal{X}} \mathbf{E} \{ \pi(\xi) \} \mathbf{E} \{ \rho(\eta) \}. \quad (45)$$

3°. It remains to invoke

► **Lemma 11.** *Let*

$$\mathcal{V} = \{v \in \mathbf{R}^d : \exists r \in \mathcal{R} : v^T R_j v \leq r_j, 1 \leq j \leq J\} \subset \mathbf{R}^d$$

be a basic ellitope,  $W \succeq 0$  be symmetric  $d \times d$  matrix such that

$$\exists r \in \mathcal{R} : \text{Tr}(W R_j) \leq r_j, j \leq J,$$

and  $\omega \sim \mathcal{N}(0, W)$ . Denoting by  $\rho(\cdot)$  the norm on  $\mathbf{R}^d$  with the unit ball  $\mathcal{V}$ , we have

$$\mathbf{E} \{ \rho(\omega) \} \leq \varkappa(J)$$

where  $\varkappa(\cdot)$  is as in (22).

The statement of the proposition now follows from (45) by applying Lemma 11 to  $\mathcal{V} = \mathcal{W}, W = X$ , and to  $\mathcal{V} = \mathcal{Z}, W = Y$ .

4°. To complete the proof, it remains to prove Lemma 11. Let us start with the case of  $J = 1$ . Setting  $\bar{r} = \max\{r : r \in \mathcal{R}\}$  and  $R = R_1 / \bar{r}$ , we have  $\text{Tr}(WR) \leq 1$  and  $\rho(u) = \|R^{1/2} u\|_2$ . Setting  $\bar{W} = R^{1/2} W R^{1/2}$  and  $\bar{\omega} = R^{1/2} \omega$ , we get  $\bar{\omega} \sim \mathcal{N}(0, \bar{W}), \text{Tr}(\bar{W}) \leq 1$ , and

$$\mathbf{E} \{ \rho(\omega) \} = \mathbf{E} \{ \|\bar{\omega}\|_2 \} \leq \sqrt{\mathbf{E} \{ \bar{\omega}^T \bar{\omega} \}} = \sqrt{\text{Tr}(\bar{W})} \leq 1 = \varkappa(1).$$



Now let  $J > 1$ . Observe that if  $\Theta \succeq 0$  is a  $d \times d$  matrix with trace 1,  $0 \leq t < 1/2$ , and  $\zeta \sim \mathcal{N}(0, I_d)$  then by convexity of  $\mathbf{E}\{\exp\{t \sum_i \zeta_i^2 \lambda_i\}\}$  in  $\lambda$

$$\mathbf{E}\{\exp\{t \zeta^T \Theta \zeta\}\} = \mathbf{E}\left\{\exp\left\{t \sum_i \zeta_i^2 \lambda_i(\Theta)\right\}\right\} \leq \mathbf{E}_{\zeta \sim \mathcal{N}(0,1)}\{\exp\{t \zeta^2\}\} = (1 - 2t)^{-1/2}.$$

As a result,

$$\forall s \geq 0 : \text{Prob}\{\zeta^T \Theta \zeta \geq s^2\} \leq \frac{\exp\{-ts^2\}}{\sqrt{1-2t}}.$$

Under the premise of the lemma, let  $W \succeq 0$  be such that  $\text{Tr}(WR_j) \leq r_j$  for all  $j$ . For every  $j$  such that  $r_j > 0$ , setting  $\Theta_j = W^{1/2} R_j W^{1/2} / r_j$ , we get  $\Theta_j \succeq 0$ ,  $\text{Tr}(\Theta_j) \leq 1$ , so that by the above for all  $s > 0$  and  $0 \leq t < 1/2$

$$\text{Prob}_{\omega \sim \mathcal{N}(0,W)}\{\omega^T R_j \omega > s^2 r_j\} = \text{Prob}_{\zeta \sim \mathcal{N}(0,I_d)}\{\zeta^T \Theta_j \zeta > s^2\} \leq \frac{\exp\{-ts^2\}}{\sqrt{1-2t}}.$$

The resulting inequality clearly holds true for  $j$  with  $r_j = 0$  as well. Now, when  $\omega$  and  $s > 0$  are such that  $\omega^T R_j \omega \leq s^2 r_j$  for all  $j$ , we have  $\rho(\omega) \leq s$ . Combining our observations, we get

$$\text{Prob}_{\omega \sim \mathcal{N}(0,W)}\{\rho(\omega) > s\} \leq \min\left[1, J \frac{\exp\{-ts^2\}}{\sqrt{1-2t}}\right],$$

implying that

$$\mathbf{E}_{\omega \sim \mathcal{N}(0,W)}\{\rho(\omega)\} \leq \int_0^\infty \min\left[1, J \frac{\exp\{-ts^2\}}{\sqrt{1-2t}}\right] ds$$

Optimizing w.r.t.  $t$ , we arrive at

$$\mathbf{E}_{\omega \sim \mathcal{N}(0,W)}\{\rho(\omega)\} \leq \frac{5}{2} \sqrt{\ln(2J)} = \varkappa(J). \quad \blacktriangleleft$$

### A.3 Proof of Proposition 8

0°. Equalities in (27) follow from (15), (16). Consequently, all we need is to prove that for all  $i, j$  it holds

$$\|\mathcal{U}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i} \leq \text{Opt}_{ij}[\mathcal{U}] \leq \max[\varsigma(K_i, L_j) + \varkappa(K_i)\varkappa(L_j)\vartheta(2\kappa)] \|\mathcal{U}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i}. \quad (46)$$

1°. Let us fix  $i \leq I$ ,  $j \leq J$ , and let  $\bar{A}_0 = Q_j^T A_{\text{nom}} P_i$  and  $\bar{A}_s = Q_j^T A_s P_i$ ,  $1 \leq s \leq S$ . Setting

$$\bar{\mathcal{U}}_{ij} = \left\{ \sum_{s=0}^N \epsilon_s \bar{A}_s : \|\epsilon\|_\infty \leq 1 \right\},$$

we clearly have  $\|\mathcal{U}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i} = \|\bar{\mathcal{U}}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i}$ . Now, comparing (22) with  $\mathcal{Z}_j^*$  in the role of  $\mathcal{B}$  and  $\mathcal{X}_i$  in the role of  $\mathcal{X}$  with the definition of  $\text{Opt}_{ij}$  in (26), we see that  $\text{Opt}_{ij}$  is nothing but the upper bound, as given by Proposition 5, on  $\|\bar{\mathcal{U}}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i}$ , implying the left inequality in (46).

2°. Observe that the upper bound on  $\alpha := \|\bar{A}_0\|_{\mathcal{Z}_j^*, \mathcal{X}_i}$  as given by Theorem 3, is nothing but

$$\bar{\alpha} := \min_{\substack{\lambda', v', \\ G, H}} \left\{ \phi_{\mathcal{T}_i}(\lambda') + \phi_{\mathcal{R}_j}(v') : \begin{array}{l} \lambda' \geq 0, v' \geq 0, G \leq \sum_{\ell=1}^{L_j} v'_\ell R_{\ell j}, \\ H \leq \sum_{k=1}^{K_i} \lambda'_k T_{ki}, \left[ \begin{array}{c|c} G & \frac{1}{2} \bar{A}_0 \\ \frac{1}{2} \bar{A}_0^T & H \end{array} \right] \succeq 0 \end{array} \right\},$$

and by this Theorem,

$$\alpha \leq \bar{\alpha} \leq \varsigma(K_i, L_j) \alpha.$$

Next, the upper bound on  $\beta := \|\mathcal{A}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i}$ ,  $\mathcal{A}_{ij} = \{\sum_{s=1}^N \epsilon_s \bar{A}_s : \|\epsilon\|_\infty \leq 1\}$ , given by Proposition 5 is

$$\bar{\beta} := \min_{\substack{\lambda'', v'', \\ \{G^s, H^s, s \leq S\}}} \left\{ \phi_{\mathcal{T}_i}(\lambda'') + \phi_{\mathcal{R}_j}(v'') : \begin{array}{l} \lambda'' \geq 0, v'' \geq 0, \sum_s G^s \leq \sum_{\ell=1}^{L_j} v''_\ell R_{\ell j}, \\ \sum_s H^s \leq \sum_{k=1}^{K_i} \lambda''_k T_{ki}, \left[ \begin{array}{c|c} G^s & \frac{1}{2} \bar{A}_s \\ \frac{1}{2} \bar{A}_s^T & H^s \end{array} \right] \succeq 0 \end{array} \right\}.$$

and

$$\bar{\beta} \leq \vartheta(2\kappa)\varkappa(K_i)\varkappa(L_j)\|\mathcal{A}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}}$$

(since the ranks of matrices  $\bar{A}_s$ ,  $s \geq 1$ , do not exceed those of matrices  $A_s$ ).

Looking at (26), we see that if  $(\lambda', v', G, H)$ ,  $(\lambda'', v'', \{G^s, H^s\})$  are feasible solutions to the optimization problems specifying  $\bar{\alpha}$  and  $\bar{\beta}$ , then

$$\lambda^{ij} = \lambda' + \lambda'', v^{ij} = v' + v'', G^{ijs} = G^s, H^{ijs} = H^s, \bar{G}^{ij} = G, \bar{H}^{ij} = H$$

is a feasible solution to the problem specifying  $\text{Opt}_{ij}[\mathcal{U}]$ , and the value of the objective of the latter problem at this feasible solution is

$$\phi_{\mathcal{T}_i}(\lambda' + \lambda'') + \phi_{\mathcal{R}_j}(v' + v'') \leq \phi_{\mathcal{T}_i}(\lambda') + \phi_{\mathcal{T}_i}(\lambda'') + \phi_{\mathcal{R}_j}(v') + \phi_{\mathcal{R}_j}(v'').$$

We conclude that

$$\text{Opt}_{ij}[\mathcal{U}] \leq \bar{\alpha} + \bar{\beta} \leq \varsigma(K_i, L_j)\|\bar{A}_0\|_{\mathcal{Z}_j^*, \mathcal{X}_i} + \varkappa(K_i)\varkappa(L_j)\vartheta(2\kappa)\|\mathcal{A}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i},$$

and since by evident reasons one has  $\|\mathcal{U}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i} \geq \max[\|\bar{A}_0\|_{\mathcal{Z}_j^*, \mathcal{X}_i}, \|\mathcal{A}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i}]$ , we arrive at the right inequality in (46). ◀

## B Spectratopic case

### B.1 Spectratopes

A *basic spectratope* is a bounded set  $\mathcal{W}$  represented as

$$\mathcal{W} = \{w \in \mathbf{R}^p : \exists t \in \mathcal{T} : T_k^2[w] \preceq t_k I_{d_k}, 1 \leq k \leq K\} \quad (47)$$

where

$$T_k[w] = \sum_{i=1}^p w_i T_{ki}$$

is a linear mapping from  $\mathbf{R}^p$  to  $\mathbf{S}^{d_k}$  (so that  $T_{ki}$  are symmetric  $d_k \times d_k$  matrices), and  $\mathcal{T}$  is as in the definition of a basic ellitope.

A *spectratope* is a set  $\mathcal{X}$  represented as the linear image of a basic spectratope  $\mathcal{W}$ :

$$\mathcal{X} = P\mathcal{W} = \{x \in \mathbf{R}^n : \exists w \in \mathcal{W} : x = Pw\}, \quad \mathcal{W} = \left\{ w \in \mathbf{R}^p : \exists t \in \mathcal{T} : \left[ \sum_i w_i T_{ki} \right]^2 \preceq t_k I_{d_k}, k \leq K \right\}$$

We refer to  $D = \sum_{k=1}^K d_k$  as *spectratopic size* of  $\mathcal{W}$  and  $\mathcal{X}$ .

Same as ellitopes, spectratopes are convex compact sets symmetric w.r.t. the origin; a basic spectratope, in addition, has a nonempty interior.

#### B.1.1 Examples

First of all, *every ellitope is a spectratope*. Indeed, it suffices to consider the case when the ellitope  $\mathcal{W}$  in question is the basic ellitope (2). In this case, passing to eigenvalue decompositions of matrices  $T_k$ , we have

$$T_k = \sum_{i=1}^{\nu_k} e_{ki} e_{ki}^T, \quad \nu_k = \text{rank}(T_k),$$

whence

$$\begin{aligned} \mathcal{W} &= \{w : \exists t \in \mathcal{T} : w^T T_k w \leq t_k, k \leq K\} \\ &= \{w : \exists \bar{t} = \{t_{ki}, i \leq \nu_k, k \leq K\} \in \bar{\mathcal{T}} : [e_{ki}^T w]^2 \preceq t_{ki} I_p, 1 \leq i \leq \nu_k, k \leq K\}, \\ \bar{\mathcal{T}} &:= \left\{ t_{ki} \geq 0 : \left[ \sum_i t_{1i}; \dots; \sum_i t_{Ki} \right] \in \mathcal{T} \right\}. \end{aligned}$$

An example of a “genuine” basic spectratope is the unit  $|\cdot|$ -ball,  $|\cdot|$  being the spectral norm on  $\mathbf{R}^{p \times q}$ :

$$\{w \in \mathbf{R}^{p \times q} : |w| \leq 1\} = \left\{ w : \exists t \in [0, 1] : T^2[w] := \left[ \frac{w}{w^T} \middle| w \right]^2 \preceq t I_{p+q} \right\},$$

Same as ellitopes, spectratopes admit fully algorithmic “calculus”, and their family is closed with respect to basic operations preserving convexity and symmetry w.r.t. the origin, such as taking finite intersections, linear images, inverse images under linear embedding, direct products, arithmetic summation (see [21, Section 4.6] for details); what is missing, is taking convex hulls of finite unions.

### B.1.2 Bounding maximum of quadratic form over a spectratope

Given a linear mapping

$$R[w] = \sum_{i=1}^{\nu} w_i R_i : \mathbf{R}^{\nu} \rightarrow \mathbf{S}^d$$

so that  $R_i \in \mathbf{S}^d$ , we associate with it linear mappings

$$R^+[W] = \sum_{i,j=1}^{\nu} W_{ij} R_i R_j : \mathbf{S}^{\nu} \rightarrow \mathbf{S}^d, \quad R^{+,*}[\Lambda] = [\text{Tr}(\Lambda R_i R_j)]_{i,j \leq \nu} : \mathbf{S}^d \rightarrow \mathbf{S}^{\nu}.$$

Note that

$$R^+[ww^T] = R^2[w] \tag{48}$$

and

$$\text{Tr}(R^+[W]\Lambda) = \text{Tr}(WR^{+,*}[\Lambda]) \quad \forall (W \in \mathbf{S}^{\nu}, \Lambda \in \mathbf{S}^d). \tag{49}$$

Given a collection  $\Lambda = \{\Lambda_k, k \leq K\}$  of symmetric matrices (of, perhaps, different sizes), we set

$$\lambda[\Lambda] = [\text{Tr}(\Lambda_1); \dots; \text{Tr}(\Lambda_K)].$$

Finally, same as above, for a convex compact set  $\mathcal{T}$ ,

$$\phi_{\mathcal{T}}(\lambda) = \max_{t \in \mathcal{T}} \lambda^T t$$

is the support function of  $\mathcal{T}$ .

Given a spectratope

$$\mathcal{X} = P\mathcal{W}, \quad \mathcal{W} = \left\{ w \in \mathbf{R}^q : \exists t \in \mathcal{T} : T_k^2[w] := \left[ \sum_i w_i T_{ki} \right]^2 \preceq t_k I_{d_k}, k \leq K \right\},$$

an efficiently computable upper bound  $\text{Opt}(C)$  on the quantity

$$\text{Opt}_*(C) = \max_{x \in \mathcal{X}} x^T C x$$

can be built as follows. Assume that  $\Lambda = \{\Lambda_k \in \mathcal{S}_+^{d_k}, k \leq K\}$  is such that

$$P^T C P \preceq \sum_k T_k^{+,*}[\Lambda_k]. \tag{50}$$

When  $x \in \mathcal{X}$ , there exists  $w \in \mathbf{R}^q$  and  $t \in \mathcal{T}$  such that (see (48))

$$x = Pw \text{ \& } T_k^+[ww^T] = T_k^2[w] \preceq t_k I_{d_k}, k \leq K,$$

whence

$$\sum_k \text{Tr}(T_k^+[ww^T]\Lambda_k) \leq \sum_k t_k \text{Tr}(\Lambda_k) \leq \phi_{\mathcal{T}}(\lambda[\Lambda]). \tag{51}$$

On the other hand, by (49) we have

$$\text{Tr}(T_k^+[ww^T]\Lambda_k) = \text{Tr}(T_k^{+,*}[\Lambda_k][ww^T]) = w^T T_k^{+,*}[\Lambda_k]w, \quad (52)$$

so that

$$x^T Cx = w^T [P^T C P]w \underset{\text{[by (50)]}}{\leq} w^T \left[ \sum_k T_k^{+,*}[\Lambda_k] \right] w \leq \phi_{\mathcal{T}}(\lambda[\Lambda])$$

due to (51) and (52). As a result, the efficiently computable convex function

$$\text{Opt}(C) = \min_{\Lambda} \left\{ \phi_{\mathcal{T}}(\lambda[\Lambda]) : \Lambda = \{\Lambda_k \in \mathbf{S}_+^{d_k}, k \leq K\}, P^T C P \preceq \sum_k T_k^{+,*}[\Lambda_k] \right\}$$

is an upper bound on  $\text{Opt}(C)$ . It is known ([21, Proposition 4.8]) that this bound is reasonably tight:

$$\text{Opt}_*(C) \leq \text{Opt}(C) \leq 2 \ln(2D) \text{Opt}_*(C), \quad D = \sum_k d_k.$$

## B.2 Bounding operator norms, spectratopic case

Similarly to the ellitopic case, our current problem of interest is tight computationally efficient upper-bounding of the norm

$$\|A\|_{\mathcal{B}, \mathcal{X}} = \max_{x \in \mathcal{X}} \|Ax\|_{\mathcal{B}} = \max_{[y;x] \in \mathcal{B}_* \times \mathcal{X}} [y; x]^T \left[ \frac{\frac{1}{2}A}{\frac{1}{2}A^T} \right] [y; x]$$

in the case when  $\mathcal{X}$  and  $\mathcal{B}_*$  are spectratopes:

$$\begin{aligned} \mathcal{X} &= P\mathcal{W} = \{x \in \mathbf{R}^n : \exists w \in \mathcal{W} : x = Pw\}, \\ \mathcal{W} &= \{w \in \mathbf{R}^p : \exists t \in \mathcal{T} : T_k^2[w] \preceq t_k I_{d_k}, k \leq K\} \\ \mathcal{B} &= \{v \in \mathbf{R}^m : v^T y \leq 1 \ \forall y \in \mathcal{B}_*\}, \mathcal{B}_* = Q\mathcal{Z} = \{y \in \mathbf{R}^m : \exists z \in \mathcal{Z} : y = Qz\}, \\ \mathcal{Z} &= \{z \in \mathbf{R}^q : \exists r \in \mathcal{R} : R_\ell^2[z] \preceq s_\ell I_{g_\ell}, \ell \leq L\}. \end{aligned} \quad (53)$$

In this case the efficiently computable upper bound on  $\|A\|_{\mathcal{B}, \mathcal{X}}$  and its tightness are given by the following result (which is an improvement of the just cited result from [21]):

► **Theorem 12.** *In the case of (53) the efficiently computable convex function of  $A$  given by*

$$\text{Opt}(A) = \min_{\Lambda, \Upsilon} \left\{ \phi_{\mathcal{T}}(\lambda[\Lambda]) + \phi_{\mathcal{R}}(\lambda[\Upsilon]) : \begin{array}{l} \Lambda = \{\Lambda_k \in \mathbf{S}_+^{d_k}, k \leq K\}, \Upsilon = \{\Upsilon_\ell \in \mathbf{S}_+^{g_\ell}, \ell \leq L\} \\ \left[ \begin{array}{c|c} \sum_\ell R_\ell^{+,*}[\Upsilon_\ell] & \frac{1}{2}Q^T A P \\ \hline \frac{1}{2}P^T A^T Q & \sum_k T_k^{+,*}[\Lambda_k] \end{array} \right] \succeq 0 \end{array} \right\} \quad (54)$$

is a reasonably tight upper bound on  $\|A\|_{\mathcal{B}, \mathcal{X}}$ :

$$\begin{aligned} \|A\|_{\mathcal{B}, \mathcal{X}} &\leq \text{Opt}(A) \leq \bar{\varsigma} \left( \sum_{k=1}^K d_k \right) \bar{\varsigma} \left( \sum_{\ell=1}^L g_\ell \right) \|A\|_{\mathcal{B}, \mathcal{X}}, \\ \bar{\varsigma}(M) &= \sqrt{2 \ln(5M)} \end{aligned} \quad (55)$$

**Proof. 1°.** The left inequality in (55) is evident. Let us prove the right inequality. Let  $q, p$  be the dimensions of the embedding spaces of  $\mathcal{Z}$  and  $\mathcal{W}$ , and assume that  $q \leq p$ , which is w.l.o.g. for the same reasons as in the ellitopic case. Same as in the latter case, (54) is nothing but the conic problem

$$\text{Opt}(A) = \min_{\Lambda, \Upsilon, \tau, \theta} \left\{ \tau + \theta : \begin{array}{l} \Lambda = \{\Lambda_k \in \mathbf{S}_+^{d_k}, k \leq K\}, [-\lambda[\Lambda]; \tau] \in \mathfrak{T}_* \\ \Upsilon = \{\Upsilon_\ell \in \mathbf{S}_+^{g_\ell}, \ell \leq L\}, [-\lambda[\Upsilon]; \theta] \in \mathfrak{R}_* \\ \left[ \begin{array}{c|c} \sum_\ell R_\ell^{+,*}[\Upsilon_\ell] & \frac{1}{2}Q^T A P \\ \hline \frac{1}{2}P^T A^T Q & \sum_k T_k^{+,*}[\Lambda_k] \end{array} \right] \succeq 0 \end{array} \right\}$$

with the same cones  $\mathfrak{I}, \mathfrak{R}$  and their duals  $\mathfrak{I}_*, \mathfrak{R}_*$  as in the elliptic case. Same as in that case, the latter problem is strictly feasible and bounded, and by Conic Duality one has

$$\begin{aligned} \text{Opt}(A) &= \max_{r,t,U,V,W} \left\{ \text{Tr}(W^T Q^T A P) : \begin{array}{l} R_\ell^+[U] \leq r_\ell I_{g_\ell} \quad \forall \ell, T_k^+[V] \leq t_k I_{d_k} \quad \forall k \\ t \in \mathcal{T}, r \in \mathcal{R}, \left[ \begin{array}{c|c} U & W \\ \hline W^T & V \end{array} \right] \succeq 0 \end{array} \right\} \\ &= \max_{r,t,U,V,Y} \left\{ \text{Tr}([U^{1/2} Y V^{1/2}]^T Q^T A P) : \begin{array}{l} r \in \mathcal{R}, t \in \mathcal{T}, U \succeq 0, V \succeq 0, Y^T Y \preceq I \\ R_\ell^+[U] \leq r_\ell I_{g_\ell} \quad \forall \ell, T_k^+[V] \leq t_k I_{d_k} \quad \forall k \end{array} \right\} \\ &= \max_{r,t,U,V} \left\{ \sum_{i=1}^q \sigma_i(U^{1/2} Q^T A P V^{1/2}) : \begin{array}{l} R_\ell^+[U] \leq r_\ell I_{g_\ell} \quad \forall \ell, T^+[V] \leq t_k I_{d_k} \quad \forall k \\ U \succeq 0, V \succeq 0, r \in \mathcal{R}, t \in \mathcal{T} \end{array} \right\} \end{aligned}$$

(cf. item 1° in the “elliptic proof”).

2°. The concluding optimization problem in the above chain clearly is solvable; let  $U, V, r, t$  be the optimal solution, and  $\sum_{i=1}^q \sigma_i e_i f_i^T$  be the singular value decomposition of  $U^{1/2} Q^T A P V^{1/2}$ , so that

$$\begin{aligned} \text{Opt}(A) &= \sum_{i=1}^q \sigma_i \\ U^{1/2} Q^T A P V^{1/2} &= \sum_{i=1}^q \sigma_i e_i f_i^T \\ e_i^T e_j &= \begin{cases} 1, & i = j \\ 0, & i \neq j, \end{cases} \quad i, j \leq q \quad \& \quad f_i^T f_j = \begin{cases} 1, & i = j \\ 0, & i \neq j, \end{cases} \quad i, j \leq p. \end{aligned} \tag{56}$$

Let  $\epsilon_1, \dots, \epsilon_p$  be independent random variables taking values  $\pm 1$  with probabilities  $1/2$ , and let

$$\xi = \sum_{i=1}^q \epsilon_i e_i, \quad \eta = \sum_{j=1}^p \epsilon_j f_j.$$

Then, in view of (56) it holds, identically in  $\epsilon_i = \pm 1, 1 \leq i \leq p$ :

$$\xi^T U^{1/2} Q^T A P V^{1/2} \eta = \sum_{i, \ell \leq q, j \leq p} [\epsilon_i \epsilon_j \sigma_i e_i^T e_\ell f_\ell^T f_j] = \sum_{i=1}^q \sigma_i = \text{Opt}(A). \tag{57}$$

On the other hand, setting  $E = [e_1, \dots, e_q]$ , we get an orthonormal  $q \times q$  matrix such that  $\xi = E \underline{\epsilon}$ , where  $\underline{\epsilon} = [\epsilon_1; \dots; \epsilon_q]$  is a Rademacher vector. Now let  $\hat{\xi} = U^{1/2} \xi = U^{1/2} E \underline{\epsilon}$ . Observe that for every  $\ell \leq L$  and for properly selected matrices  $\bar{R}_{\ell i} \in \mathbf{S}^{g_\ell}$  we have

$$R_\ell[U^{1/2} E y] = \sum_{i=1}^q \bar{R}_{\ell i} y_i, \quad \forall y \in \mathbf{R}^q.$$

We have

$$\mathbf{E}\{\hat{\xi} \hat{\xi}^T\} = \mathbf{E}\{U^{1/2} E \underline{\epsilon} \underline{\epsilon}^T E^T U^{1/2}\} = U^{1/2} E \mathbf{E}\{\underline{\epsilon} \underline{\epsilon}^T\} E^T U^{1/2} = U^{1/2} E E^T U^{1/2} = U,$$

whence

$$\mathbf{E}\{R_\ell^2[\hat{\xi}]\} = R_\ell^+[\mathbf{E}\{\hat{\xi} \hat{\xi}^T\}] = R_\ell^+[U] \preceq r_\ell I_{g_\ell} \quad \forall \ell \leq L$$

(we have used (48)). On the other hand,

$$R_\ell[\hat{\xi}] = R_\ell[U^{1/2} E \underline{\epsilon}] = \sum_{i=1}^q \bar{R}_{\ell i} \epsilon_i,$$

so that  $\sum_{i=1}^q \bar{R}_{\ell i}^2 = \mathbf{E}\{R_\ell^2[\hat{\xi}]\}$ , and we end up with

$$\sum_{i=1}^q \bar{R}_{\ell i}^2 \preceq r_\ell I_{g_\ell}.$$

Applying the noncommutative Khintchine inequality<sup>11</sup> we conclude that

$$\forall s > 0 : \text{Prob}\{R_\ell^2[\widehat{\xi}] \preceq s^2 r_\ell I_{g_\ell}\} = 1 - \text{Prob}\left\{\left|\sum_i \bar{R}_{\ell i} \epsilon_i\right| > s\sqrt{r_\ell}\right\} \geq 1 - 2d_\ell \exp\left\{-\frac{1}{2}s^2\right\}.$$

As a result, when setting  $D = \sum_k d_k$  and  $s = \sqrt{2\ln(5D)}$  we get

$$\text{Prob}\{R_\ell^2[\widehat{\xi}] \preceq 2\ln(5D)r_\ell I_{d_\ell}, \ell \leq L\} > 1/2,$$

and

$$\text{Prob}\{\|QU^{1/2}\xi\|_{\mathcal{B}_*} \leq \sqrt{2\ln(5D)}\} \geq \text{Prob}\{R_\ell^2[\widehat{\xi}] \preceq 2\ln(5D)r_\ell I_{d_\ell}, \ell \leq L\} > 1/2.$$

By similar reasoning,

$$\text{Prob}\{\|PV^{1/2}\eta\|_{\mathcal{X}} \leq \sqrt{2\ln(5G)}\} > 1/2, \quad G = \sum_\ell g_\ell.$$

As a result, there exists realization  $(\bar{\xi}, \bar{\eta})$  of  $(\xi, \eta)$  such that

$$\|QU^{1/2}\bar{\xi}\|_{\mathcal{B}_*} \leq \sqrt{2\ln(5D)} \ \& \ \|PV^{1/2}\bar{\eta}\|_{\mathcal{X}} \leq \sqrt{2\ln(5G)}.$$

On the other hand, invoking (57),

$$\text{Opt}(A) = \bar{\xi}^T U^{1/2} Q^T A P V^{1/2} \bar{\eta} \leq \|QU^{1/2}\bar{\xi}\|_{\mathcal{B}_*} \|PV^{1/2}\bar{\eta}\|_{\mathcal{X}} \|A\|_{\mathcal{B}, \mathcal{X}}.$$

Combining our observations, we conclude that

$$\text{Opt}(A) \leq 2\sqrt{\ln(5D)\ln(5G)} \|A\|_{\mathcal{B}, \mathcal{X}}. \quad \blacktriangleleft$$

### B.3 Bounding robust norms of uncertain matrices, spectratopic case

Let spectratopes  $\mathcal{X} \subset \mathbf{R}^n$ ,  $\mathcal{B}_* \subset \mathbf{R}^m$  with nonempty interiors and the polar  $\mathcal{B}$  of  $\mathcal{B}_*$  be given by (53). Our goal is to conceive a computationally efficient upper-bounding of the robust norm

$$\|A\|_{\mathcal{B}, \mathcal{X}} = \max_{A \in \mathcal{A}} \|A\|_{\mathcal{B}, \mathcal{X}}$$

of uncertain matrix

$$\mathcal{A} = \left\{ \sum_s \epsilon_s A_s : \|\epsilon\|_\infty \leq 1 \right\} \subset \mathbf{R}^{m \times n}.$$

#### B.3.1 Processing the problem

Acting exactly as in the ellitopic case, with the results of Section B.1.2 in the role of their “ellitopic counterparts” from Section 2.2, we conclude that the efficiently computable quantity

$$\text{Opt} := \min_{\substack{\Lambda, \Upsilon, \\ \{G_s, H_s\}}} \left\{ \phi_{\mathcal{R}}(\lambda[\Upsilon]) + \phi_{\mathcal{T}}(\lambda[\Lambda]) : \begin{array}{l} \left[ \frac{G_s}{\frac{1}{2}P^T A_s^T Q} \mid \frac{\frac{1}{2}Q^T A_s P}{H_s} \right] \succeq 0, s \leq S \\ \Upsilon = \{\Upsilon_\ell \in \mathbf{S}_+^{g_\ell}, \ell \leq L\}, \sum_s G_s \preceq \sum_\ell R_\ell^{+,*}[\Upsilon_\ell] \\ \Lambda = \{\Lambda_k \in \mathbf{S}_+^{d_k}, k \leq K\}, \sum_s H_s \preceq \sum_k T_k^{+,*}[\Lambda_k] \end{array} \right\} \quad (58)$$

—the “spectratopic analog” of (22)—is an upper bound on  $\|A\|_{\mathcal{B}, \mathcal{X}}$  such that for properly selected matrices  $X \in \mathbf{S}_+^p$ ,  $Y \in \mathbf{S}_+^q$  and  $r \in \mathcal{R}$ ,  $t \in \mathcal{T}$  one has

$$R_\ell^+[Y] \preceq r_\ell I_{g_\ell}, \ell \leq L, \ \& \ T_k^+[X] \preceq t_k I_{d_k}, k \leq K,$$

and for the norms  $\pi(\cdot)$  and  $\rho(\cdot)$  with unit balls  $\mathcal{W}$  and  $\mathcal{Z}$ , respectively, and  $[\eta; \xi] \sim \mathcal{N}(0, \text{Diag}\{Y, X\})$ ,

$$\text{Opt} \leq \vartheta(2\kappa) \|A\|_{\mathcal{B}, \mathcal{X}} \mathbf{E}\{\rho(\eta)\pi(\xi)\} = \vartheta(2\kappa) \|A\|_{\mathcal{B}, \mathcal{X}} \mathbf{E}\{\pi(\xi)\} \mathbf{E}\{\rho(\eta)\} \quad (59)$$

where  $\kappa$  is the maximum of ranks of  $A_s$  and  $\vartheta(\cdot)$  is given by (21) (cf. (45)).

We have the following spectratopic analog of Lemma 11.

<sup>11</sup> Noncommutative Khintchine’s inequality due to Lust-Piquard, Pisier, and Buchholz, see [42, Theorem 4.6.1], states that if  $Q_i \in \mathbf{S}^n$ ,  $1 \leq i \leq I$ , and  $\xi_i$ ,  $i = 1, \dots, I$ , are independent Rademacher or  $\mathcal{N}(0, 1)$  random variables, then for all  $t \geq 0$  one has

$$\text{Prob}\left\{\left|\sum_{i=1}^I \xi_i Q_i\right| \geq t\right\} \leq 2n \exp\left\{-\frac{t^2}{2\left|\sum_{i=1}^I Q_i^2\right|}\right\}$$

where  $|\cdot|$  is the spectral norm.

► **Lemma 13.** *Let*

$$\mathcal{V} = \{v \in \mathbf{R}^d : \exists r \in \mathcal{R} : R_j^2[v] \preceq r_j I_{\nu_j}, 1 \leq j \leq J\} \subset \mathbf{R}^d$$

be a basic spectratope,  $W \succeq 0$  be symmetric  $d \times d$  matrix such that

$$\exists r \in \mathcal{R} : R_j^+[W] \preceq r_j I_{\nu_j}, j \leq J,$$

and  $\omega \sim \mathcal{N}(0, W)$ . Denoting by  $\gamma(\cdot)$  the norm on  $\mathbf{R}^d$  with the unit ball  $\mathcal{V}$ , we have

$$\mathbf{E}\{\rho(\omega)\} \leq \bar{\varkappa} \left( \sum_j \nu_j \right), \quad \bar{\varkappa}(F) = 2\sqrt{2\ln(2F)}. \quad (60)$$

**Proof.** Let  $\zeta \sim \mathcal{N}(0, I_d)$ . When setting

$$\bar{R}_j[z] = R_j[W^{1/2}z] = \sum_{i=1}^d \bar{R}_{ji} z_i, \quad [\bar{R}_{ji} \in \mathbf{S}^{\nu_j}, j \leq J],$$

we have

$$\sum_i \bar{R}_{ji}^2 = \mathbf{E}\{\bar{R}_{ji}^2[\zeta]\} = \mathbf{E}\{R_j^2[W^{1/2}\zeta]\} = \mathbf{E}\{R_j^+[W^{1/2}\zeta\zeta^T W^{1/2}]\} = R_j^+[W] \preceq r_j I_{\nu_j}.$$

Hence for every  $s > 0$

$$\begin{aligned} \text{Prob}\{R_j^2[\omega] \preceq s^2 r_j I_{\nu_j}\} &= \text{Prob}\{\bar{R}_{ji}^2[\zeta] \preceq s^2 r_j I_{\nu_j}\} = 1 - \text{Prob}\left\{\left|\sum_i \zeta_i \bar{R}_{ji}\right| > s\sqrt{r_j}\right\} \\ &\quad [\text{as above, } |\cdot| \text{ is the spectral norm}] \\ &\geq 1 - 2\nu_j \exp\{-s^2/2\}, \end{aligned}$$

with the concluding  $\geq$  given by  $\sum_i \bar{R}_{ji}^2 \preceq r_j I_{\nu_j}$  combined with the noncommutative Khintchine inequality. As a result,

$$\text{Prob}\{\gamma(\omega) > s\} \leq 1 - \text{Prob}\{\exists j : R_j^2[\omega] \preceq s^2 r_j I_{\nu_j}\} \leq \left[ \sum_j 2\nu_j \right] \exp\{-s^2/2\}.$$

Therefore, when setting  $F = \sum_j \nu_j$  we obtain

$$\mathbf{E}\{\gamma(\omega)\} \leq \int_0^\infty \min[1, 2F \exp\{-\gamma^2/2\}] d\gamma \leq 2\sqrt{2\ln(2F)}. \quad \blacktriangleleft$$

Applying the lemma to  $\mathcal{V} = \mathcal{W}$ ,  $W = X$ , and to  $\mathcal{V} = \mathcal{Z}$ ,  $W = Y$ , we get from (59) the following analog of Proposition 5:

► **Proposition 14.** *In the situation described in the beginning of this section, assuming that ranks of all  $A_s$  are  $\leq \kappa$ , the efficiently computable quantity  $\text{Opt}$  as given by (58) is a reasonably tight upper bound on the robust norm  $\|A\|_{\mathcal{B}, \mathcal{X}}$  of uncertain matrix  $A$ , specifically,*

$$\|A\|_{\mathcal{B}, \mathcal{X}} \leq \text{Opt} \leq \bar{\varkappa} \left( \sum_k d_k \right) \bar{\varkappa} \left( \sum_\ell g_\ell \right) \vartheta(2\kappa) \|A\|_{\mathcal{B}, \mathcal{X}} \quad (61)$$

where  $\bar{\varkappa}(\cdot)$  is given by (60) and  $\vartheta(\cdot)$ , as given by (21), satisfies

$$\vartheta(1) = 1, \quad \vartheta(2) = \frac{\pi}{2}, \quad \vartheta(4) = 2, \quad \vartheta(k) \leq \pi\sqrt{k}/2.$$

### B.3.2 Putting things together

Results of Proposition 14 (and, as a byproduct, of Theorem 12) can be extended, in exactly the same fashion as in the elliptic case, to the situation where  $\mathcal{X}$  and the polar  $\mathcal{B}_*$  of  $\mathcal{B}$  are convex hulls of finite unions of spectratopes rather than plain spectratopes, and the uncertain matrix in question is not centered, resulting in the following spectratopic analogy of Proposition 8:

► **Theorem 15.** *Let  $\mathcal{U} = \{A_{\text{nom}} + \sum_{s=1}^S \epsilon_s A_s : \|\epsilon\|_\infty \leq 1\}$  be an uncertain  $m \times n$  matrix,  $\mathcal{X} \subset \mathbf{R}^n$ ,  $\mathcal{B}, \mathcal{B}_* \subset \mathbf{R}^m$  be given by*

$$\mathcal{X} = \text{Conv} \left\{ \bigcup_{i=1}^I P_i \mathcal{X}_i \right\} = \left\{ x = \sum_{i=1}^I \lambda_i P_i x_i : x_i \in \mathcal{X}_i, \lambda_i \geq 0, \sum_i \lambda_i = 1 \right\},$$

$$\mathcal{B} = \left\{ v \in \mathbf{R}^m : \max_{y \in \mathcal{B}_*} v^T y \leq 1 \right\}, \mathcal{B}_* = \text{Conv} \left\{ \bigcup_{j=1}^J Q_j \mathcal{Z}_j \right\} = \left\{ y = \sum_{j=1}^J \mu_j Q_j z_j, z_j \in \mathcal{Z}_j, \mu_j \geq 0, \sum_j \mu_j = 1 \right\},$$

with basic spectratopes

$$\mathcal{X}_i = \{x_i \in \mathbf{R}^{\nu_i} : \exists t^i \in \mathcal{T}^i : T_{ki}^2[x_i]^2 \preceq t_k^i I_{d_{ki}}, 1 \leq k \leq K_i\}, T_{ki}[x] = \sum_{\iota=1}^{\nu_i} x_\iota T_{k\iota i}, i \leq I$$

$$\mathcal{Z}_j = \{z_j \in \mathbf{R}^{\mu_j} : \exists r^j \in \mathcal{R}^j : R_{\ell j}^2[z_j] \preceq r_\ell^j I_{g_{\ell j}}, 1 \leq \ell \leq L_j\}, R_{\ell j}[z] = \sum_{\iota=1}^{\mu_j} z_\iota R_{\ell j \iota}, j \leq J$$

Then the efficiently computable quantity

$$\text{Opt}[\mathcal{U}] = \max_{i \leq I, j \leq J} \text{Opt}_{ij}[\mathcal{U}],$$

where

$$\text{Opt}_{ij}[\mathcal{U}] = \min_{\substack{\Lambda^{ij}, \Upsilon^{ij}, G^{ijs}, H^{ijs} \\ \bar{G}^{ij}, \bar{H}^{ij} \\ 1 \leq i \leq I, 1 \leq j \leq J, 1 \leq s \leq S}} \left\{ \phi_{\mathcal{T}^i}(\lambda[\Upsilon^{ij}]) + \phi_{\mathcal{R}^j}(\lambda[\Lambda^{ij}]) : \right.$$

$$\left. \begin{array}{l} \Lambda^{ij} = \{\Lambda_k^{ij} \succeq 0, k \leq K_i\}, \Upsilon^{ij} = \{\Upsilon_\ell^{ij} \succeq 0, \ell \leq L_j\} \\ \sum_{s=1}^S H^{ijs} + \bar{H}^{ij} \preceq \sum_{k=1}^{K_i} T_{ki}^{+,*}[\Lambda_k^{ij}] \\ \sum_{s=1}^S G^{ijs} + \bar{G}^{ij} \preceq \sum_{\ell=1}^{L_j} R_{\ell j}^{+,*}[\Upsilon_\ell^{ij}], \\ \left[ \begin{array}{c|c} G^{ijs} & \frac{1}{2}[Q_j^T A_s P_i] \\ \hline \frac{1}{2}[Q_j^T A_s P_i]^T & H^{ijs} \end{array} \right] \succeq 0, s \leq S \\ \left[ \begin{array}{c|c} \bar{G}^{ij} & \frac{1}{2}[Q_j^T A_{\text{nom}} P_i] \\ \hline \frac{1}{2}[Q_j^T A_{\text{nom}} P_i]^T & \bar{H}^{ij} \end{array} \right] \succeq 0 \end{array} \right\}, \quad i \leq I, j \leq J,$$

is an efficiently computable convex in  $(A_{\text{nom}}, A_1, \dots, A_S)$  upper bound on  $\|\mathcal{U}\|_{\mathcal{B}, \mathcal{X}}$ . This upper bound is reasonably tight, specifically, setting

$$\mathcal{U}_{ij} = Q_j^T A_{\text{nom}} P_i + \left\{ \sum_{s=1}^S \epsilon_s [Q_j^T A_s P_i] : \|\epsilon\|_\infty \leq 1 \right\},$$

we have

$$\|\mathcal{U}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i} \leq \text{Opt}_{ij}[\mathcal{U}] \leq [\bar{\zeta}(D_i) \bar{\zeta}(G_j) + \bar{\varkappa}(D_i) \bar{\varkappa}(G_j) \vartheta(2\kappa)] \|\mathcal{U}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i},$$

$$D_i = \sum_{k=1}^{K_i} d_{ki}, \quad G_j = \sum_{\ell=1}^{L_j} g_{\ell j},$$



and

$$\begin{aligned} \|\mathcal{U}\|_{\mathcal{B},\mathcal{X}} &= \max_{i \leq I, j \leq J} \|\mathcal{U}_{ij}\|_{\mathcal{Z}_j^*, \mathcal{X}_i} \leq \text{Opt}[\mathcal{U}] = \max_{i \leq I, j \leq J} \text{Opt}_{ij}[\mathcal{U}] \\ &\leq \left[ \max_{i \leq I, j \leq J} [\bar{\varsigma}(D_i) \bar{\varsigma}(G_j) + \bar{\varkappa}(D_i) \bar{\varkappa}(G_j) \vartheta(2\kappa)] \right] \|\mathcal{U}\|_{\mathcal{B},\mathcal{X}} \end{aligned}$$

where  $\kappa$  is the maximum of ranks of  $A_s$ ,  $1 \leq s \leq S$ ,  $\bar{\varsigma}(\cdot)$  and  $\bar{\varkappa}(\cdot)$  are as defined in (55) and (60), and  $\vartheta(\cdot)$  is defined by (21) and satisfies (19).

---

## References

- 1 John Abedor, Krishan Nagpal, and Kameshwar Poolla. A linear matrix inequality approach to peak-to-peak gain minimization. *Int. J. Robust Nonlinear Control*, 6(9-10):899–927, 1996.
- 2 Erling D. Anderson. *The MOSEK optimization toolbox for MATLAB manual. Version 8.0*, 2015. <https://docs.mosek.com/latest/toolbox/index.html>.
- 3 Karl Johan Åström and Pieter Eykhoff. System identification—a survey. *Automatica*, 7(2):123–162, 1971.
- 4 Johannes Aubrecht and Petros G. Voulgaris. Minimization of the peak-to-peak gain in periodic systems under full state feedback. *J. Dyn. Syst. Meas. Control*, 123(1):10–20, 2001.
- 5 Venkataramanan Balakrishnan and Stephen Boyd. On computing the worst-case peak gain of linear systems. *Syst. Control Lett.*, 19(4):265–269, 1992.
- 6 Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. Society for Industrial and Applied Mathematics, 2001.
- 7 Aharon Ben-Tal and Arkadi Nemirovski. On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty. *SIAM J. Optim.*, 12(3):811–833, 2002.
- 8 Dimetri Bertsekas and Ian Rhodes. Recursive state estimation for a set-membership description of uncertainty. *IEEE Trans. Autom. Control*, 16(2):117–128, 1971.
- 9 Stephen Boyd. EE263 Lecture Notes, 2007. <https://web.stanford.edu/class/archive/ee/ee263/ee263.1082/lectures/aircraft2.pdf>.
- 10 Stephen Boyd and John Doyle. Comparison of peak and RMS gains for discrete-time systems. *Syst. Control Lett.*, 9(1):1–6, 1987.
- 11 Juanyu Bu and Mario Sznaiar. A linear matrix inequality approach to synthesizing low-order suboptimal mixed  $\ell_1/H_p$  controllers. *Automatica*, 36(7):957–963, 2000.
- 12 Marco Casini, Andrea Garulli, and Antonio Vicino. Feasible parameter set approximation for linear models with bounded uncertain regressors. *IEEE Trans. Autom. Control*, 59(11):2910–2920, 2014.
- 13 Vito Cerone. Feasible parameter set for linear models with bounded errors in all variables. *Automatica*, 29(6):1551–1555, 1993.
- 14 J. E. Cope and Bert W. Rust. Bounds on solutions of linear systems with inaccurate data. *SIAM J. Numer. Anal.*, 16(6):950–963, 1979.
- 15 Ignacio J. Diaz-Bobillo and Munther A. Dahleh. Minimization of the maximum peak-to-peak gain: The general multiblock problem. *IEEE Trans. Autom. Control*, 38(10):1459–1482, 1993.
- 16 Pieter Eykhoff. *Trends and progress in system identification: IFAC Series for Graduates, Research Workers & Practising Engineers*. Pergamon Press, 1981.
- 17 Michael Grant and Stephen Boyd. *The CVX Users Guide. Release 2.1*, 2014. <http://web.cvxr.com/cvx/doc/>.
- 18 Nicholas J Higham. *Accuracy and stability of numerical algorithms*. Society for Industrial and Applied Mathematics, 2002.
- 19 Luc Jaulin, Michel Kieffer, Olivier Didrit, and Éric Walter. Interval analysis. In *Applied interval analysis*, pages 11–43. Springer, 2001.
- 20 Anatoli Juditsky and Arkadi Nemirovski. Near-optimality of linear recovery from indirect observations. *Math. Stat. Learn.*, 1(2):101–110, 2018.
- 21 Anatoli Juditsky and Arkadi Nemirovski. *Statistical Inference via Convex Optimization*. Princeton University Press, 2020.
- 22 Georgios Kotsalis, Guanghui Lan, and Arkadi Nemirovski. Convex Optimization for Finite-Horizon Robust Covariance Control of Linear Stochastic Systems. *SIAM J. Control Optim.*, 59(1):296–319, 2021.
- 23 Vladik Kreinovich, Anatoly V. Lakeyev, and Sergey I. Noskov. Optimal solution of interval linear systems is intractable (NP-hard). *Interval Comput.*, 1:6–14, 1993.
- 24 Aleksandr B. Kurzhanskiĭ. *Control and observation under uncertainty*. Nauka, 1977.
- 25 Aleksandr B. Kurzhanskiĭ. Identification problem—theory of guaranteed estimates. *Automation and Remote Control*, 52(4):447–465, 1991.

- 26 Aleksandr B. Kurzhanskiĭ and István Vályi. *Ellipsoidal calculus for estimation and control*. Systems and Control: Foundations and Applications. Birkhäuser, 1997.
- 27 Lennart Ljung. *System identification: Theory for the user*. Prentice Hall, 2nd edition, 1997.
- 28 Aleksandr Ivanovič Matasov. *Estimators for uncertain dynamic systems*, volume 458. Springer, 1998.
- 29 Raman K. Mehra and Dimitri G. Lainiotis. *System identification advances and case studies*. Academic Press Inc., 1977.
- 30 Mario Milanese, John Norton, Hélène Piet-Lahanier, and Éric Walter. *Bounding approaches to system identification*. Springer, 2013.
- 31 Sergey A. Nazin and Boris T. Polyak. Interval parameter estimation under model uncertainty. *Math. Comput. Model. Dyn. Syst.*, 11(2):225–237, 2005.
- 32 Sergey A. Nazin and Boris T. Polyak. Ellipsoid-based parametric estimation in the linear multidimensional systems with uncertain model description. *Automation and Remote Control*, 68(6):993–1005, 2007.
- 33 Arkadi Nemirovski, Cornelis Roos, and Tamás Terlaky. On maximization of quadratic form over intersection of ellipsoids with common center. *Math. Program.*, 86(3):463–473, 1999.
- 34 Yurii Nesterov. Semidefinite relaxation and nonconvex quadratic optimization. *Optim. Methods Softw.*, 9(1-3):141–160, 1998.
- 35 Yurii Nesterov. Global quadratic optimization via conic relaxation. In R Saigal, H Wolkowicz, and L Vandenberghe, editors, *Handbook on Semidefinite Programming*, pages 363–387. Kluwer Academic Publishers, 2000.
- 36 Arnold Neumaier and Arnold Neumaier. *Interval methods for systems of equations*. Number 37. Cambridge University Press, 1990.
- 37 Werner Oettli and William Prager. Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides. *Numer. Math.*, 6(1):405–409, 1964.
- 38 Boris T. Polyak. Robust linear algebra and robust aperiodicity. In A. Rantzer and C. Byrnes, editors, *Directions in Mathematical Systems Theory and Optimization*, pages 249–260. Springer, 2003.
- 39 Fred C. Schweppe. *Uncertain dynamic systems*. Prentice Hall, 1973.
- 40 Torsten Söderström and Petre Stoica. *System identification*. Pearson Education Ltd, 1994.
- 41 Daureen Steinberg. Computation of matrix norms with applications to robust optimization. Master’s thesis, M.Sc. thesis in OR and System Analysis, Technion-Israel Institute of Technology, 2005. <https://www2.isye.gatech.edu/~nemirovs/Daureen.pdf>.
- 42 Joel A. Tropp. An introduction to matrix concentration inequalities. *Found. Trends Mach. Learn.*, 8(1-2):1–230, 2015.
- 43 Éric Walter. Special issue on parameter identification with error bounds. *Math. Comput. Simul.*, 32(447):607, 1990.