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

Robust signal recovery under uncertain-but-bounded perturbations in observation matrix

Yannis Bekri · Anatoli Juditsky · Arkadi Nemirovski

the date of receipt and acceptance should be inserted later

Abstract In this paper our focus is on analysis and design of linear and polyhedral signal recoveries robust with respect to the deterministic uncertainty in the observation matrix. This can be seen as a "deterministic counterpart" of the work [1] where the case of random uncertainty was studied. We investigate the performance of estimates robust w.r.t. deterministic norm-bounded matrix uncertainty, derive efficiently computable bounds for the estimation risk and discuss the construction of "presumably good" estimates.

2020 Mathematics Subject Classification: 62G05, 62G10, 90C90
Keywords: statistical linear inverse problems, robust estimation, observation matrix uncertainty

1 Introduction

In this paper we consider the estimation problem as follows. We are given an observation $\omega \in \mathbf{R}^m$,

$$\omega = A[\eta]x + \xi \tag{1}$$

Yannis Bekri
LJK, Université Grenoble Alpes
Campus de Saint-Martin-d'Hères, 38401 France
yannis.bekri@univ-grenoble-alpes.fr

Anatoli Juditsky, Corresponding author
LJK, Université Grenoble Alpes
Campus de Saint-Martin-d'Hères, 38401 France
anatoli.juditsky@univ-grenoble-alpes.fr

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

where $\xi \in \mathbf{R}^m$ is zero mean random noise and $A[\eta] \in \mathbf{R}^{m\times n}$ is affine in *perturbation* $\eta \in \mathbf{R}^q$ sensing matrix. We assume that unknown signal x belongs to a given convex set $\mathcal{X} \subset \mathbf{R}^n$ and that perturbation vector $\eta \in \mathbf{R}^q$ in (1) ("uncertainty", for short) is deterministic and runs through a given *uncertainty set* $\mathcal{U} \subset \mathbf{R}^q$,

$$A[\eta] = A + D[\eta], \quad D[\eta] = \sum\nolimits_{\alpha=1}^{q} \eta_\alpha A_\alpha \in \mathbf{R}^{m\times n} \tag{2}$$

where $A, A_1, ..., A_q$ are given matrices. Our objective is to build an estimate of the linear image $w = Bx$ of x, $B \in \mathbf{R}^{\nu\times n}$, which is robust with respect to uncertainty η.

Estimation from observations (1) under uncertain-but-bounded perturbation of observation matrix can be seen as an extension of the problem of solving systems of equations affected by uncertainty which has received significant attention in the literature (cf., e.g., [10,11,15,19,21,22,23] and references therein). It is also closely related to the problem of system identification in the case when observations of system's states are subjected to uncertain-but-bounded perturbation [6,8,9,12,16,17,18,20,24].

As applied to the estimation problem above, *linear estimate* $\widehat{w}^H_{\rm lin}(\omega)$ of w is of the form $\widehat{w}^H_{\rm lin}(\omega) = H^T\omega$ where *contrast matrix* $H \in \mathbf{R}^{m\times\nu}$ is the estimate parameter. A polyhedral estimate $\widehat{w}^H_{\rm poly}(\omega)$ is specified by a *contrast matrix* $H \in \mathbf{R}^{m\times M}$ according to

$$\omega \mapsto \widehat{x}^H(\omega) \in \mathop{\rm Argmin}_{x\in\mathcal{X}} \left\{ \|H^T(\omega - Ax)\|_\infty \right\}, \ \ \widehat{w}^H_{\rm poly}(\omega) := B\widehat{x}(\omega). \tag{3}$$

In this paper, our goal is to investigate design of robust linear and polyhedral estimates under deterministic uncertainty in the sensing matrix. We extend the approach developed in [1] where robust estimates of both types were analyzed in the case of stochastic uncertainty η to the situation in which η is uncertain-but-bounded. The approach reduces to deriving a tight efficiently computable upper bound on the estimate's risk and then "building" the estimate by minimizing this bound with respect to the estimate parameter H.
Our contributions are as follows.

- We analyse the ϵ-risk (the maximum, over signals from $\mathcal{X}$, of the radii of $(1-\epsilon)$-confidence $\|\cdot\|$-balls) in Section 2 and build presumably good linear estimates in the case of *structured norm-bounded uncertainty* (cf. [2, Chapter 7] and references therein), thus extending the corresponding results of [12].
  In our analysis, similarly to [1], we assume that the signal set $\mathcal{X}$ is an ellitope [13,14],[1] and the norm $\|\cdot\|$ quantifying the recovery error is the maximum of a finite collection of Euclidean norms.
- In our context, analysis and design of polyhedral estimates under uncertain-but-bounded perturbations in the sensing matrix appears to be the most difficult; our very limited results on this subject form the subject of Section 3,

[1] Ellitopes are defined in Section 2; an immediate example is a bounded intersection of centered at the origin elliptic cylinders

Notation and assumptions. We denote by $\|\cdot\|$ the norm on $\mathbf{R}^\nu$ used to measure the estimation error. In what follows, $\|\cdot\|$ is the maximum of Euclidean seminorms

$$\|u\| = \max_{\ell\le L}\sqrt{u^TR_\ell u}\quad R_\ell\succeq 0,\ \ell=1,...,L,\ \sum\nolimits_\ell R_\ell\succ 0.$$

We denote by $\phi_Y(z)=\sup_{y\in Y} z^Ty$ the support function of a set $Y\subset\mathbf{R}^k$. Throughout the paper, we assume that observation noise ξ (its distribution P_x may depend on x) is zero-mean sub-Gaussian, $\xi\sim\mathcal{SG}(0,\sigma^2 I)$, i.e., for all $t\in\mathbf{R}^m$m $x\in X$,

$$\mathbf{E}_{\xi\sim P_x}\left\{e^{t^T\xi}\right\}\le \exp\left(\tfrac{\sigma^2}{2}\|t\|_2^2\right). \tag{4}$$

We define the ϵ-risk of an estimate $\omega\mapsto\widehat{w}(\omega)$: we consider uniform over $x\in\mathcal{X}$ and $\eta\in\mathcal{U}$ ϵ-risk

$$\mathrm{Risk}_\epsilon[\widehat{w}|\mathcal{X}]=\sup_{x\in\mathcal{X},\eta\in\mathcal{U}}\inf\Big\{\rho:\mathrm{Prob}_{\xi\sim P_x}\{\|\widehat{w}(A[\eta]x+\xi)-Bx\|>\rho\}\le\epsilon\Big\}.$$

2 Design of presumably good linear estimate

Let us assume that the signal set $\mathcal{X}$ is *a basic ellitope*. By definition [13,14], a basic ellitope in $\mathbf{R}^n$ is a set of the form

$$\mathcal{X}=\{x\in\mathbf{R}^n:\exists t\in\mathcal{T}:\ z^TT_kz\le t_k,\ k\le K\} \tag{5}$$

where $T_k\in\mathbf{S}^n_+$, $T_k\succeq 0$, $\sum_k T_k\succ 0$, and $\mathcal{T}\subset\mathbf{R}^K_+$ is a convex compact set with a nonempty interior which is monotone: whenever $0\le t'\le t\in\mathcal{T}$ one has $t'\in\mathcal{T}$.

Observe that the error of the linear estimate $\widehat{w}^H(\omega)=H^T\omega$ satisfies

$$\begin{aligned}\|\widehat{w}(A[\eta]x+\xi)-Bx\|\le\|H^T\xi\|&+\max_{x\in\mathcal{X}}\|[B-H^TA]x\|\\&+\max_{x\in\mathcal{X},\eta\in\mathcal{U}}\left\|H^TD[\eta]x\right\|\end{aligned} \tag{6}$$

Design of a presumably good linear estimate $\widehat{x}_H(\omega)$ consists in minimizing over H the sum of tight efficiently computable upper bounds on the terms in the right-hand side of (6). The bounds on the first and the second term were already established in [1, Section 2.1]. Namely, we have

$$\mathrm{Prob}\left\{\|H^T\xi\|\ge[1+\sqrt{2\ln(L/\epsilon)}]\sigma\max_{\ell\le L}\sqrt{\mathrm{Tr}(HR_\ell H^T)}\right\}\le\epsilon\quad\forall\epsilon\in(0,1]$$

and $\max\limits_{x\in\mathcal{X}}\|[B-H^TA]x\|\le\mathfrak{r}_\ell(H)$ where

$$\mathfrak{r}_\ell(H)=\min_{\mu,\lambda}\left\{\lambda+\phi_{\mathcal{T}}(\mu):\ \mu\ge0,\ \left[\begin{array}{c|c}\lambda I_\nu & \frac12 R_\ell^{1/2}[B-H^TA]\\\hline \frac12[B-H^TA]^TR_\ell^{1/2} & \sum_k\mu_kT_k\end{array}\right]\succeq0\right\}$$

(cf. (35) and (36) in the proof of Proposition 2.1 of [1]). What is missing is a tight upper bound on

$$\mathfrak{s}(H) = \max_{x\in\mathcal{X},\eta\in\mathcal{U}} \left\|H^T D[\eta]x\right\|.$$

In the rest of this section we focus on building efficiently computable upper bound on $\mathfrak{s}(H)$ which is convex in H; the synthesis of the contrast H is then conducted by minimizing with respect to H the resulting upper bound on the estimation risk.

We assume from now on that $\mathcal{U}$ is a convex compact set in certain $\mathbf{R}^q$. In this case $\mathfrak{s}(H)$ is what in [12] was called the *robust norm*

$$\|\mathcal{Z}[H]\|_{\mathcal{X}} = \max_{Z\in\mathcal{Z}[H]} \|Z\|_{\mathcal{X}},\ \|Z\|_{\mathcal{X}} = \max_{x\in\mathcal{X}} \|Zx\|$$

of the *uncertain* $\nu \times n$ *matrix*

$$\mathcal{Z}[H] = \{Z = H^T D[\eta] : \eta \in \mathcal{U}\},$$

i.e., the maximum, over *instances* $Z \in \mathcal{Z}[H]$, of operator norms of the linear mappings $x \mapsto Zx$ induced by the norm with the unit ball $\mathcal{X}$ on the argument space and the norm $\|\cdot\|$ on the image space.

It is well known that aside of a very restricted family of special cases, robust norms do not allow for efficient computation. We are about to list known to us generic cases when these norms admit efficiently computable upper bounds which are tight within logarithmic factors.

2.1 Scenario uncertainty

This is the case where the nuisance set $\mathcal{U} = \mathrm{Conv}\{\eta^1, ..., \eta^S\}$ is given as a convex hull of moderate number of scenarios η^s. In this case, $\mathfrak{s}(H)$ the maximum of operator norms:

$$\begin{aligned}\mathfrak{s}(H) &= \max_{s\leq S}\max_{x\in\mathcal{X}} \|H^T D[\eta^s]x\| \\ &= \max_{s\leq S,\ell\leq L} \|\mathcal{M}_{s\ell}[H]\|_{\mathcal{X},2}, \quad \mathcal{M}_{s\ell}[H] = R_\ell^{1/2} H^T D[\eta^s],\end{aligned}$$

where, for $Q \in \mathbf{R}^{\nu\times n}$, $\|Q\|_{\mathcal{X},2} = \max_{x\in\mathcal{X}} \|Qx\|_2$ is the operator norm of the linear mapping $x \mapsto Qx : \mathbf{R}^n \to \mathbf{R}^\nu$ induced by the norm $\|\cdot\|_{\mathcal{X}}$ with the unit ball $\mathcal{X}$ on the argument space, and the Euclidean norm $\|\cdot\|_2$ on the image space. Note that this norm is efficiently computable in the *ellipsoid case* where $\mathcal{X} = \{x \in \mathbf{R}^n : x^T T x \leq 1\}$ with $T \succ 0$ (that is, for $K = 1$, $T_1 = T$, $\mathcal{T} = [0, 1]$ in (5))—one has $\|Q\|_{\mathcal{X},2} = \|QT^{-1/2}\|_{2,2}$. When $\mathcal{X}$ is a general ellitope, norm $\|\cdot\|_{\mathcal{X},2}$ is difficult to compute. However, it admits a tight efficiently computable convex in Q upper bound: it is shown in [12, Theorem 3.1] that function

$$\mathrm{Opt}[Q] = \min_{\lambda,\mu} \left\{ \lambda + \phi_{\mathcal{T}}(\mu) : \mu \geq 0, \left[\begin{array}{c|c} \lambda I_\nu & \frac{1}{2}Q \\ \hline \frac{1}{2}Q^T & \sum_k \mu_k T_k \end{array}\right] \succeq 0 \right\}$$

satisfies $\|Q\|_{\mathcal{X},2} \leq \mathrm{Opt}[Q] \leq 2.4\sqrt{\ln(4K)}\|Q\|_{\mathcal{X},2}$. As a result, under the circumstances,

$$\bar{\mathfrak{s}}(H) = \max_{s\leq S,\ell\leq L} \mathrm{Opt}_{s\ell}[H],$$
$$\mathrm{Opt}_{s\ell}[H] = \min_{\lambda_\ell,\mu^\ell} \left\{\lambda_\ell + \phi_{\mathcal{T}}(\mu^\ell) : \mu^\ell \geq 0, \left[\begin{array}{c|c} \lambda_\ell I_\nu & \frac{1}{2}R_\ell^{1/2}H^T D[\eta^s] \\ \hline \frac{1}{2}D^T[\eta^s]HR_\ell^{1/2} & \sum_k \mu^\ell{}_k T_k \end{array}\right] \succeq 0\right\},$$

is a tight within the factor $2.4\sqrt{\ln(4K)}$ efficiently computable convex in H upper bound on $\mathfrak{s}(H)$.

2.2 Box and structured norm-bounded uncertainty

In the case of *structured norm-bounded uncertainty* function $D[\eta]$ in the model (2) is of the form

$$\begin{array}{l} D[\eta] = \sum_{\alpha=1}^q P_\alpha^T \eta_\alpha Q_\alpha \quad [P_\alpha \in \mathbf{R}^{p_\alpha\times m}, Q_\alpha \in \mathbf{R}^{q_\alpha\times n}],\\ \mathcal{U} = \{\eta = (\eta_1, ..., \eta_q)\} = \mathcal{U}_1 \times ... \times \mathcal{U}_q,\\ \mathcal{U}_\alpha = \left\{\begin{array}{ll} \{\eta_\alpha = \delta I_{p_\alpha} : |\delta| \leq 1\} \subset \mathbf{R}^{p_\alpha\times p_\alpha}, & q_\alpha = p_\alpha,\ \alpha \leq q_s, \\ & [\text{"scalar perturbation blocks"}]\\ \{\eta_\alpha \in \mathbf{R}^{p_\alpha\times q_\alpha} : \|\eta_\alpha\|_{2,2} \leq 1\} & , q_s < \alpha \leq q. \\ & [\text{"general perturbation blocks"}] \end{array}\right. \end{array} \tag{7}$$

The special case of (7) where $q_s = q$, that is,

$$\mathcal{U} = \{\eta \in \mathbf{R}^q : \|\eta\|_\infty \leq 1\}\ \&\ A[\eta] = A + D[\eta] = A + \sum_{\alpha=1}^q \eta_\alpha A_\alpha$$

is referred to as box uncertainty. In this section we operate with structured norm-bounded uncertainty (7), assuming w.l.o.g. that all P_α and Q_α are nonzero. The main result here (for underlying rationale and proof, see Section A.2) is as follows:

Proposition 2.1 *Let $\mathcal{X} \subset \mathbf{R}^n$ be an ellitope: $\mathcal{X} = P\mathcal{Y}$, where*

$$\mathcal{Y} = \{y \in \mathbf{R}^N : \exists t \in \mathcal{T} : y^T T_k y \leq t_k, k \leq K\}$$

is a basic ellitope. Given the data of structured norm-bounded uncertainty (7), consider the efficiently computable convex function

$$\bar{\mathfrak{s}}(H) = \max_{\ell\leq L} \mathrm{Opt}_\ell(H),$$
$$\mathrm{Opt}_\ell(H) = \min_{\mu,\upsilon,\lambda,U_s,V_s,U^t,V^t} \left\{\begin{array}{l} \frac{1}{2}[\mu + \phi_{\mathcal{T}}(\upsilon)] :\ \mu \geq 0, \upsilon \geq 0, \lambda \geq 0 \\ \left[\begin{array}{c|c} U_s & -A_{s\ell}[H]P \\ \hline -P^T A_{s\ell}^T[H] & V_s \end{array}\right] \succeq 0,\ s \leq q_s \\ \left[\begin{array}{c|c} U^t & -L_{t\ell}^T[H] \\ \hline -L_{t\ell}[H] & \lambda_t I_{p_{q_s+t}} \end{array}\right] \succeq 0,\ t \leq q - q_s \\ V^t - \lambda_t P^T R_t^T R_t P \succeq 0,\ t \leq q - q_s \\ \mu I_\nu - \sum_s U_s - \sum_t U^t \succeq 0,\ \ \sum_k \upsilon_k T_k - \sum_s V_s - \sum_t V^t \succeq 0 \end{array}\right\}$$

where

$$A_{s\ell}[H] = R_\ell^{1/2} H^T P_s^T Q_s,\ 1 \le s \le q_{\rm s}$$
$$L_{t\ell}[H] = P_{q_{\rm s}+t} H R_\ell^{1/2},\ R_t = Q_{q_{\rm s}+t},\ 1 \le t \le q - q_{\rm s}.$$

Then

$$\mathfrak{s}(H) \le \bar{\mathfrak{s}}(H) \le \varkappa(K) \max[\vartheta(2\kappa), \pi/2] \mathfrak{s}(H),$$

where $\kappa = \max\limits_{\alpha \le q_{\rm s}} \min[p_\alpha, q_\alpha]$ *(*$\kappa = 0$ *when* $q_{\rm s} = 0$*),*

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

and $\vartheta(k)$ *is a universal function of integer* $k \ge 0$ *specified in (25) such that*

$$\vartheta(0) = 0,\ \vartheta(1) = 1,\ \vartheta(2) = \pi/2,\ \vartheta(3) = 1.7348...,\ \vartheta(k) \le \tfrac{1}{2}\pi\sqrt{k},\ \ k \ge 1.$$

Note that the "box uncertainty" version of Proposition 2.1 was derived in [12].

2.3 Robust estimation of linear forms

Until now, we imposed no restrictions on the matrix B. We are about to demonstrate that when we aim at recovering the value of a given linear form $b^T x$ of signal $x \in \mathcal{X}$, i.e., when B is a row vector:

$$Bx = b^T x \qquad [b \in \mathbf{R}^n], \tag{8}$$

we can handle much wider family of uncertainty sets $\mathcal{U}$ than those considered so far. Specifically, assume on the top of (8) that $\mathcal{U}$ is a *spectratope* (as is the case, e.g., with structured norm-bounded uncertainty) – a set of the form

$$\begin{array}{c}\mathcal{U} = \{\eta = Qv, v \in \mathcal{V}\},\ \mathcal{V} = \{v \in \mathbf{R}^M : \exists s \in \mathcal{S} : S_\ell^2[v] \preceq s_\ell I_{d_\ell},\ \ell \le L\}, \\ S_\ell[v] = \sum_{i=1}^M v_i S^{i\ell},\ S^{i\ell} \in \mathbf{S}^{d_\ell}\end{array} \tag{9}$$

where $\mathcal{S} \subset \mathbf{R}^L_+$ is a convex compact monotone (cf. definition of ellitope) set with a nonempty interior and $S_\ell[v] = 0,\ \ell \le L$, implies $v = 0$. Let $\mathcal{X}$ be a spectratope as well:

$$\begin{array}{c}\mathcal{X} = \{x = Py, y \in \mathcal{Y}\},\ \mathcal{Y} = \{y \in \mathbf{R}^N : \exists t \in \mathcal{T} : T_k^2[y] \preceq t_k I_{f_k},\ k \le K\}, \\ T_k[y] = \sum_{j=1}^N y_j T^{jk},\ T^{jk} \in \mathbf{S}^{f_k}.\end{array} \tag{10}$$

The contrast matrix H underlying a candidate linear estimate becomes a vector $h \in \mathbf{R}^m$, the associated linear estimate being $\widehat{w}_h(\omega) = h^T \omega$. In our present situation $\nu = 1$ we lose nothing when setting $\|\cdot\| = |\cdot|$. Representing $D[\eta]$ as $\sum_{\alpha=1}^q \eta_\alpha A_\alpha$, we get

$$\mathfrak{r}_b(h) = \max_{x \in \mathcal{X}, \eta \in \mathcal{U}} \left| h^T \sum\nolimits_\alpha \eta_\alpha A_\alpha x \right| = \max_{\eta \in \mathcal{U}, x \in \mathcal{X}} \eta^T A[h] x,\ A[h] = [h^T A_1; ...; h^T A_q].$$

In other words, $\mathfrak{r}_b(h)$ is the operator norm $\|A[h]\|_{\mathcal{X},\mathcal{U}_*}$ of the linear mapping $x \mapsto A[h]x$ induced by the norm $\|\cdot\|_{\mathcal{X}}$ with the unit ball $\mathcal{X}$ on the argument space and the norm with the unit ball $\mathcal{U}_*$—the polar of the spectratope $\mathcal{U}$—on the image space. Denote

$$\lambda[\Lambda] = [\mathrm{Tr}(\Lambda_1); ...; \mathrm{Tr}(\Lambda_K)], \quad \Lambda_k \in \mathbf{S}^{f_k},$$
$$\lambda[\Upsilon] = [\mathrm{Tr}(\Upsilon_1); ...; \mathrm{Tr}(\Upsilon_L)], \quad \Upsilon_\ell \in \mathbf{S}^{d_\ell},$$

and for $Y \in \mathbf{S}^{d_\ell}$ and $X \in \mathbf{S}^{f_k}$

$$R_\ell^{+,*}[Y] = \left[\mathrm{Tr}(YS^{i\ell}S^{j\ell})\right]_{i,j\le M}, \quad T_k^{+,*}[X] = \left[\mathrm{Tr}(XT^{ik}T^{jk})\right]_{i,j\le N}.$$

Invoking [12, Theorem 7], we arrive at

Proposition 2.2 *In the case of (9) and (10), efficiently computable convex function*

$$\bar{\mathfrak{r}}_b(h) = \min_{\Lambda,\Upsilon} \left\{ \tfrac{1}{2}[\phi_{\mathcal{T}}(\lambda[\Lambda]) + \phi_{\mathcal{S}}(\lambda[\Upsilon])] : \begin{array}{l} \Lambda = \{\Lambda_k \in \mathbf{S}_+^{f_k}, k \le K\}, \Upsilon = \{\Upsilon_\ell \in \mathbf{S}_+^{d_\ell}, \ell \le L\} \\ \left[\begin{array}{c|c} \sum_\ell R_\ell^{+,*}[\Upsilon_\ell] & Q^T A[h] P \\ \hline P^T A^T[h] Q & \sum_k T_k^{+,*}[\Lambda_k] \end{array}\right] \succeq 0 \end{array} \right\} \quad (11)$$

is a reasonably tight upper bound on $\mathfrak{r}_b(h)$*:*

$$\mathfrak{r}_b(h) \le \bar{\mathfrak{r}}_b(h) \le \bar{\varsigma}\left(\sum_{k=1}^K f_k\right) \; \bar{\varsigma}\left(\sum_{\ell=1}^L d_\ell\right) \mathfrak{r}_b(h)$$

where $\bar{\varsigma}(J) = \sqrt{2\ln(5J)}$.

3 Design of the robust polyhedral estimate

On a close inspection, the strategy for designing a presumably good polyhedral estimate developed in [1, Section 3] for the case of random uncertainty works in the case of uncertain-but-bounded perturbations $A[\eta] = A + \underbrace{\sum_\alpha \eta_\alpha A_\alpha}_{D[\eta]}, \eta \in \mathcal{U}$, provided that the constraints on the columns h of the contrast matrices are replaced with the constraint

$$\mathrm{Prob}_\xi\{|h^T\xi| > 1/2\} \le \delta/2, \quad (12a)$$
$$\left|\sum_{\alpha=1}^q [h^T A_\alpha x]\eta_\alpha\right| \le 1/2 \;\forall (x \in \mathcal{X}, \eta \in \mathcal{U}). \quad (12b)$$

Recalling that $h^T\xi$ is sub-Gaussian, $h^T\xi \sim \mathcal{SG}(0, \sigma^2\|h\|_2^2)$, and assuming that $\mathcal{U}$ and $\mathcal{X}$ are the spectratopes (9), (10), invoking Proposition 2.2, an efficiently verifiable sufficient condition for h to satisfy the constraints (12) is

$$\|h\|_2 \le 2\sigma\sqrt{2\ln(2/\delta)} \;\text{ and }\; \bar{\mathfrak{r}}_b(h) \le 1/2 \quad (13)$$

(see (11)). It follows that in order to build an efficiently computable upper bound for the ϵ-risk of a polyhedral estimate associated with a given $m \times ML$ contrast matrix $H = [H_1, .., H_L]$, $H_\ell \in \mathbf{R}^{m\times M}$, it suffices to check whether the columns of H satisfy constraints (13) with $\delta = \epsilon/ML$. If the answer is positive, one can upper-bound the risk utilizing the following spectratopic version of [1, Proposition 2.3]:

Proposition 3.1 *In the situation of this section, let $\epsilon \in (0,1)$, and let $H = [H_1, ..., H_L]$ be $m\times ML$ matrix with L blocks $H_\ell \in \mathbf{R}^{m\times M}$ such that all columns of H satisfy (13) with $\delta = \epsilon/ML$. Consider optimization problem*

$$\mathfrak{p}_+[H] = 2 \min_{\lambda_\ell, \Upsilon^\ell, \upsilon^\ell, \rho} \left\{ \rho : \begin{array}{l} \upsilon^\ell \geq 0,\ \Upsilon^\ell = \{\Upsilon^\ell_k \in \mathbf{S}^{f_k}_+, k \leq K\},\ \ell \leq L \\ \lambda_\ell + \phi_{\mathcal{T}}(\lambda[\Upsilon^\ell]) + \sum_{j=1}^M \upsilon^\ell_j \leq \rho,\ \ell \leq L \\ \left[\begin{array}{c|c} \lambda_\ell I_\nu & \frac{1}{2} R_\ell^{1/2} BP \\ \hline \frac{1}{2} P^T B^T R_\ell^{1/2} & P^T A^T H_\ell \mathrm{Diag}\{\upsilon^\ell\} H_\ell^T AP + \sum_k T_k^{+,*}[\Upsilon^\ell_k] \end{array}\right] \succeq 0,\ \ell \leq L \end{array} \right\} \tag{14}$$

where

$$\lambda[\Upsilon^\ell] = [\mathrm{Tr}(\Upsilon^\ell_1); ...; \mathrm{Tr}(\Upsilon^\ell_K)]$$

and

$$T_k^{+,*}[V] = \left[\mathrm{Tr}(V T^{ik} T^{jk})\right]_{1\leq i,j \leq N} \ \text{for}\ V \in \mathbf{S}^{f_k}.$$

Then

$$\mathrm{Risk}_\epsilon[\widehat{w}^H | \mathcal{X}] \leq \mathfrak{p}_+[H].$$

Remarks. When taken together, Propositions 2.2 and 3.1 allow to compute efficiently an upper bound on the ϵ-risk of the polyhedral estimate associated with a given $m \times ML$ contrast matrix H: when the columns of H satisfy (13) with $\delta = \epsilon/ML$, this bound is $\mathfrak{p}_+[H]$, otherwise it is, say, $+\infty$. The outlined methodology can be applied to *any* pair of spectratopes $\mathcal{X}$ and $\mathcal{Y}$. However, to design a presumably good polyhedral estimate, we need to optimize the risk bound obtained in H, and this seems to be difficult because the bound, same as its "random perturbation" counterpart derived in [1, Proposition 3.1], is nonconvex in H. At present, we know only one generic situation where the synthesis problem admits "presumably good" solution—the case where both $\mathcal{X}$ and $\mathcal{U}$ are ellipsoids. Applying appropriate one-to-one linear transformations to perturbation η and signal x, the latter situation can be reduced to that with

$$\mathcal{X} = \{x \in \mathbf{R}^n : \|x\|_2 \leq 1\},\ \ \mathcal{U} = \{\eta \in \mathbf{R}^q : \|\eta\|_2 \leq 1\}, \tag{15}$$

which we assume till the end of this section. In this case (13) reduces to

$$\|h\|_2 \leq [2\sigma\sqrt{2\ln(2/\delta)}]^{-1} \ \text{ and }\ \|\mathcal{A}[h]\|_{2,2} \leq 1/2 \tag{16}$$

where

$$\mathcal{A}[h] = [h^T A_1; h^T A_2; ...; h^T A_q].$$

As a result, (12) can be processed in the same fashion as corresponding constraints in [1, Sections 3.3 and 3.4] to yield a computationally efficient scheme for building a presumably good, in the case of (15), polyhedral estimate. This scheme is the same as that described at the end of [1, Section 3.6] with just one difference: the quantity $\chi(\delta)$ in the first semidefinite constraint of optimization problems (30) and (31) which define the upper bounds of the estimation risk should now be replaced with constant 2. For instance, when denoting by Opt the optimal value of the modified in the just explained way problem (30) in [1, Section 3.6], the ϵ-risk of the polyhedral estimate yielded by an optimal solution to the problem is upper-bounded by $2\sqrt{\varkappa}\mathrm{Opt}$, with $\varkappa = 4\ln(4m(m+n+q+1))$.

Acknowledgements This work was supported by Multidisciplinary Institute in Artificial intelligence MIAI @ Grenoble Alpes (ANR-19-P3IA-0003).

A Proofs

A.1 Proof of Proposition 3.1

The proof follows that of [1, Proposition 2.3]. All we need to prove is that if H satisfies the premise of the proposition and $\lambda_\ell, \Upsilon^\ell, \upsilon^\ell, \rho$ is a feasible solution to (14), then the inequality

$$\mathrm{Risk}_\epsilon[\widehat{w}^H_{\mathrm{poly}}|\mathcal{X}] \leq 2\rho \tag{17}$$

holds. Indeed, let us fix $x \in \mathcal{X}$ and $\eta \in \mathcal{U}$. Since the columns of H satisfy (13), the P_x-probability of the event

$$\mathcal{Z}_{x,\eta} = \{\xi : \|H^T[D[\eta]x + \xi\|_\infty \leq 1\}$$

is at least $1 - ML\delta = 1 - \epsilon$. Let us fix observation $\omega = Ax + D[\eta]x + \xi$ with $\xi \in \mathcal{Z}_{x,\eta}$. Then

$$\|H^T[\omega - Ax]\|_\infty = \|H^T[D[\eta]x + \xi]\|_\infty \leq 1, \tag{18}$$

implying that the optimal value in the optimization problem $\min_{u\in\mathcal{X}} \|H^T[Au - \omega\|_\infty$ is at most 1. Consequently, setting $\widehat{x} = \widehat{x}^H(\omega)$, we have $\widehat{x} \in \mathcal{X}$ and $\|H^T[A\widehat{x} - \omega]\|_\infty \leq 1$, see (3). These observations combine with (18) and the inclusion $x \in \mathcal{X}$ to imply that for $z = \frac{1}{2}[x - \widehat{x}]$ we have $z \in \mathcal{X}$ and $\|H^T z\|_\infty \leq 1$. Recalling what $\mathcal{X}$ is we conclude that $z = Py$ with $T_k^2[y] \preceq t_k I_{f_k}, k \leq K$ for some $t \in \mathcal{T}$ and

$$\|H_\ell^T APy\|_\infty = \|H_\ell^T Az\|_\infty \leq 1,\ \ell \leq L. \tag{19}$$

Now let $u \in \mathbf{R}^\nu$ with $\|u\|_2 \leq 1$. Semidefinite constraints in (14) imply that

$$\begin{aligned}
u^T R_\ell^{1/2} Bz &= u^T R_\ell^{1/2} BPy \leq u^T \lambda_\ell I_\nu u + y^T \left[PA^T H_\ell \mathrm{Diag}\{\upsilon^\ell\} H_\ell^T AP + \sum\nolimits_k T_k^{+,*}[\Upsilon_k^\ell]\right] y \\
&= \lambda_\ell u^T u + \sum\nolimits_j \upsilon_j^\ell \underbrace{[H_\ell^T APy]_j^2}_{\leq 1 \text{ by } (19)} + \sum\nolimits_k y^T T_k^{+,*}[\Upsilon_k^\ell] y \\
&\leq \lambda_\ell + \sum\nolimits_j \upsilon_j^\ell + \sum_k \sum_{i,j \leq N} y_i y_j \mathrm{Tr}(\Upsilon_k^\ell T^{ik} T^{jk}) \\
&= \lambda_\ell + \sum\nolimits_j \upsilon_j^\ell + \sum_k \mathrm{Tr}(\Upsilon_k^\ell T_k^2[y]) \\
&\leq \lambda_\ell + \sum\nolimits_j \upsilon_j^\ell + \sum_k t_k \mathrm{Tr}(\Upsilon_k^\ell) \text{ [due to } \Upsilon^\ell \succeq 0 \text{ and } T_k^2[y] \preceq t_k I_{f_k}] \\
&\leq \lambda_\ell + \sum\nolimits_j \upsilon_j^\ell + \phi_{\mathcal{T}}(\lambda[\Upsilon^\ell]) \leq \rho
\end{aligned} \tag{20}$$

where the concluding inequality follows from the constraints of (14). (20) holds true for all u with $\|u\|_2 \leq 1$, and we conclude that for $x \in \mathcal{X}$ and $\eta \in \mathcal{U}$ and $\xi \in \mathcal{Z}_{x,\eta}$ (recall that the latter inclusion takes place with P_x-probability $\geq 1-\epsilon$) we have

$$\|R_\ell^{1/2} B[\widehat{x}^H(Ax + D[\eta]x + \xi) - x]\|_2 \leq 2\rho, \; \ell \leq L.$$

Recalling what $\|\cdot\|$ is, we get

$$\forall (x \in \mathcal{X}, \eta \in \mathcal{U}) : \mathrm{Prob}_{\xi \sim P_x}\{\|B[x - \widehat{x}^H(Ax + D[\eta]x + \xi)]\| > 2\rho\| \leq \epsilon,$$

that is, $\mathrm{Risk}_\epsilon[\widehat{w}^H_{\mathrm{poly}}|\mathcal{X}] \leq 2\rho$. The latter relation holds true whenever ρ can be extended to a feasible solution to (14), and (17) follows. □

A.2 Robust norm of uncertain matrix with structured norm-bounded uncertainty

A.2.1 Situation and goal

Let matrices $A_s \in \mathbf{R}^{m \times n}$, $s \leq S$, and $L_t \in \mathbf{R}^{p_t \times m}$, $R_t \in \mathbf{R}^{q_t \times n}$, $t \leq T$, be given. These data specify *uncertain* $m \times n$ *matrix*

$$\mathcal{A} = \{A = \sum\nolimits_s \delta_s A_s + \sum\nolimits_t L_t^T \Delta_t R_t : |\delta_s| \leq 1 \, \forall s \leq S, \|\Delta_t\|_{2,2} \leq 1 \, \forall t \leq T\}. \tag{21}$$

Given ellitopes

$$\begin{aligned}
\mathcal{X} &= \{Py : y \in \mathcal{Y}\} \subset \mathbf{R}^n, \, \mathcal{Y} = \{y \in \mathbf{R}^N \; \& \; \exists t \in \mathcal{T} : y^T T_k y \leq t_k, k \leq K\}, \\
\mathcal{B}_* &= \{Qz : z \in \mathcal{Z}\} \subset \mathbf{R}^m, \, \mathcal{Z} = \{z \in \mathbf{R}^M : \exists s \in \mathcal{S} : z^T S_\ell z \leq s_\ell, \, \ell \leq L\},
\end{aligned} \tag{22}$$

we want to upper-bound the robust norm

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

of uncertain matrix $\mathcal{A}$ induced by the norm $\|\cdot\|_{\mathcal{X}}$ with the unit ball $\mathcal{X}$ in the argument space and the norm $\|\cdot\|_{\mathcal{B}}$ with the unit ball $\mathcal{B}$ which is the polar of $\mathcal{B}_*$ in the image space.

A.2.2 Main result

Proposition A.1 *Given uncertain matrix (21) and ellitopes (22), consider convex optimization problem*

$$\begin{array}{l}
\mathrm{Opt} = \min\limits_{\substack{\mu,\upsilon,\lambda,\\ U_s,V_s,U^t,V^t}} \frac{1}{2}[\phi_{\mathcal{S}}(\mu)+\phi_{\mathcal{T}}(\upsilon)]\\
\quad\text{subject to}\\
\quad\mu\geq 0,\ \upsilon\geq 0,\ \lambda\geq 0\\
\quad\left[\begin{array}{c|c} U_s & -Q^TA_sP\\ \hline -P^TA_s^TQ & V_s\end{array}\right]\succeq 0 \qquad (23a)\\
\quad\left[\begin{array}{c|c} U^t & -Q^TL_t^T\\ \hline -L_tQ & \lambda_tI_{p_t}\end{array}\right]\succeq 0,\ V^t-\lambda_tP^TR_t^TR_tP\succeq 0 \qquad (23b)\\
\quad\sum_\ell \mu_\ell S_\ell-\sum_s U_s-\sum_t U^t\succeq 0 \qquad (23c)\\
\quad\sum_k \upsilon_k T_k-\sum_s V_s-\sum_t V^t\succeq 0 \qquad (23d)
\end{array}$$

The problem is strictly feasible and solvable, and

$$\|\mathcal{A}\|_{\mathcal{X},\mathcal{B}}\leq \mathrm{Opt}\leq \varkappa(K)\varkappa(L)\max\left[\vartheta(2\kappa),\pi/2\right]\|\mathcal{A}\|_{\mathcal{X},\mathcal{B}} \qquad (24)$$

where

- *the function* $\vartheta(k)$ *of nonnegative integer* k *is given by* $\vartheta(0)=0$ *and* $k\geq 1$*,*

$$\vartheta(k)=\left[\min_\alpha\left\{(2\pi)^{-k/2}\int|\alpha_1u_1^2+...+\alpha_ku_k^2|\mathrm{e}^{-u^Tu/2}du,\ \alpha\in\mathbf{R}^k,\|\alpha\|_1=1\right\}\right]^{-1}. \qquad (25)$$

- $\kappa=\max\limits_{s\leq S}\mathrm{Rank}(A_s)$ *when* $S\geq 1$*, otherwise* $\kappa=0$*;*
- $\varkappa(\cdot)$ *is given by*

$$\varkappa(J)=\left\{\begin{array}{ll}1, & J=1,\\ \frac{5}{2}\sqrt{\ln(2J)}, & J>1.\end{array}\right. \qquad (26)$$

Remarks. The rationale behind (23) is as follows. Checking that the $\mathcal{X},\mathcal{B}$-norm of uncertain $m\times n$ matrix (21) is $\leq a\in\mathbf{R}$ is the same as to verify that for all $\delta_s\in[-1,1]$, Δ_t : $\|\Delta_t\|_{2,2}\leq 1$

$$\sum_s\delta_su^TA_sv+\sum_t u^TL_t^T\Delta_tR_tv\leq a\|u\|_{\mathcal{B}_*}\|v\|_{\mathcal{X}}\quad\forall(u\in\mathbf{R}^m,v\in\mathbf{R}^n),$$

or, which is the same due to what $\mathcal{B}_*$ and $\mathcal{X}$ are, that for all $\delta_s\in[-1,1],\Delta_t:\|\Delta_t\|_{2,2}\leq 1$

$$\sum_s\delta_sz^TQ^TA_sPy+\sum_t z^TQ^TL_t^T\Delta_tR_tPy\leq a\|z\|_{\mathcal{Z}}\|y\|_{\mathcal{Y}}\quad\forall(z\in\mathbf{R}^M,y\in\mathbf{R}^N). \qquad (27)$$

A simple certificate for (27) is a collection of positive semidefinite matrices U_s,V_s,U^t,V^t, U,V such that for all $z\in\mathbf{R}^M$, $y\in\mathbf{R}^N$ and all $s\leq S$, $t\leq T$ it holds

$$\begin{array}{ll}
2z^T[Q^TA_sP]y\leq z^TU_sz+y^TV_sy, & (28a)\\
2z^TQ^TL_t^T\Delta_tR_tPy\leq z^TU^tz+y^TV^ty\ \ \forall(\Delta_t:\|\Delta_t\|_{2,2}\leq 1), & (28b)\\
\sum_s U_s+\sum_t U^t\preceq U, & (28c)\\
\sum_s V_s+\sum_t V^t\preceq V, & (28d)\\
\max\limits_{z\in\mathcal{Z}}z^TUz+\max\limits_{y\in\mathcal{Y}}y^TVy\leq 2a. & (28e)
\end{array}$$

Now, (28a) clearly is the same as (23a). It is known (this fact originates from [7]) that (28b) is the same as existence of $\lambda_t \geq 0$ such that (23b) holds. Finally, existence of $\mu \geq 0$ such that $\sum_\ell \mu_\ell S_\ell \succeq U$ and $\upsilon \geq 0$ such that $\sum_k \upsilon_k T_k \succeq V$ (see (23c) and (23d)) implies due to the structure of $\mathcal{Z}$ and $\mathcal{Y}$ that $\max_{z\in\mathcal{Z}} z^T Uz \leq \phi_{\mathcal{S}}(\mu)$ and $\max_{y\in\mathcal{Y}} y^T Vy \leq \phi_{\mathcal{T}}(\upsilon)$. The bottom line is that a feasible solution to (23) implies the existence of a certificate

$$\left\{U_s, U^t, V_s, V^t, s \leq S, t \leq T, U = \sum\nolimits_\ell \mu_\ell S_\ell, V = \sum\nolimits_k \upsilon_k T_k\right\}$$

for relation (27) with $a = \frac{1}{2}[\phi_{\mathcal{S}}(\mu) + \phi_{\mathcal{T}}(\upsilon)]$.

Proof of Proposition A.1. 1^o. Strict feasibility and solvability of the problem are immediate consequences of $\sum_\ell S_\ell \succ 0$ and $\sum_k T_k \succ 0$.

Let us prove the first inequality in (24). All we need to show is that if

[a] $\mu, \upsilon, \lambda, U_s, V_s, U^t, V^t$ is feasible for (23),
[b] $x = Py$ with $y^T T_k y \leq \tau_k$, $k \leq K$, for some $\tau \in \mathcal{T}$ and $u = Qz$ for some z such that $z^T S_\ell z \leq \varsigma_\ell$, $\ell \leq L$, for some $\varsigma \in \mathcal{S}$, and
[c] δ_s, Δ_t satisfy $|\delta_s| \leq 1$, $\|\Delta_t\|_{2,2} \leq 1$,

then $\gamma := u^T[\sum_s \delta_s A_s + \sum_t L_t^T \Delta_t R_t]x \leq \frac{1}{2}[\phi_{\mathcal{S}}(\mu) + \phi_{\mathcal{T}}(\upsilon)]$. Assuming [a–c], we have

$$\begin{aligned}
\gamma &= \sum\nolimits_s \delta_s z^T Q^T A_s Py + \sum\nolimits_t z^T Q^T L_t^T \underbrace{\Delta_t R_t Py}_{\zeta_t} \\
&\leq \tfrac{1}{2} z^T \left[\sum\nolimits_s U_s\right] z + \tfrac{1}{2} y^T \left[\sum\nolimits_s V_s\right] y + \sum\nolimits_t \|L_t Qz\|_2 \|\zeta_t\|_2 \;\; \text{[by (23a) and due to } |\delta_s| \leq 1] \\
&\leq \tfrac{1}{2} z^T \left[\sum_s U_s\right] z + \tfrac{1}{2} y^T \left[\sum_s V_s\right] y + \sum_t \sqrt{(\lambda_t z^T U^t z)(y^T P^T R_t^T R_t Py)} \\
&\quad \text{[due to (23b) and } \|\Delta_t\|_{2,2} \leq 1] \\
&= \tfrac{1}{2} z^T \left[\sum\nolimits_s U_s\right] z + \tfrac{1}{2} y^T \left[\sum\nolimits_s V_s\right] y + \sum\nolimits_t \sqrt{(z^T U^t z)(\lambda_t y^T P^T R_t^T R_t Py)}.
\end{aligned}$$

Thus, by the second inequality of (23b),

$$\begin{aligned}
\gamma &\leq \tfrac{1}{2} z^T \left[\sum\nolimits_s U_s\right] z + \tfrac{1}{2} y^T \left[\sum\nolimits_s V_s\right] y + \sum\nolimits_t \sqrt{(z^T U^t z)(y^T V^t y)} \\
&\leq \tfrac{1}{2} z^T \left[\sum\nolimits_s U_s\right] z + \tfrac{1}{2} y^T \left[\sum\nolimits_s V_s\right] y + \tfrac{1}{2} \sum\nolimits_t [z^T U^t z + y^T V^t y] \\
&= \tfrac{1}{2} \left[z^T \left[\sum\nolimits_s U_s + \sum\nolimits_t U^t\right] z + y^T \left[\sum\nolimits_s V_s + \sum\nolimits_t V^t\right] y\right] \\
&\leq \tfrac{1}{2} \left[\sum\nolimits_\ell \mu_\ell z^T S_\ell z + \sum\nolimits_k y^T \upsilon_k T_k y\right] \quad \text{[by (23c) and (23d)]} \\
&\leq \tfrac{1}{2} \left[\sum\nolimits_\ell \mu_\ell \varsigma_\ell + \sum\nolimits_k \upsilon_k \tau_k\right] \quad \text{[due to } z^T S_\ell z \leq \varsigma_\ell, y^T T_k y \leq \tau_k] \\
&\leq \tfrac{1}{2} [\phi_{\mathcal{S}}(\mu) + \phi_{\mathcal{T}}(\upsilon)] \quad \text{[since } \varsigma \in \mathcal{S}, \tau \in \mathcal{T}]
\end{aligned} \tag{29}$$

as claimed.

2^o. Now, let us prove the second inequality in (24). Observe that

$$\mathbf{S} = \{0\} \cup \{[s;\sigma] : \sigma > 0, s/\sigma \in \mathcal{S}\}, \quad \mathbf{T} = \{0\} \cup \{[t;\tau] : \tau > 0, t/\tau \in \mathcal{T}\},$$

are regular cones with the duals

$$\mathbf{S}_* = \{[g;\sigma] : \sigma \geq \phi_{\mathcal{S}}(-g)\}, \quad \mathbf{T}_* = \{[h;\tau] : \tau \geq \phi_{\mathcal{T}}(-h)\},$$

and (23) can be rewritten as the conic problem

$$
\begin{array}{l}
2\mathrm{Opt} = \min\limits_{\substack{\alpha,\beta,\mu,\upsilon,\\ \lambda,U_s,V_s,U^t,V^t}} \alpha+\beta \qquad\qquad (P)\\
\text{subject to}\\
[-\mu;\alpha]^{[\overline{g},\overline{\alpha}]}\in\mathbf{S}_*,\ [-\upsilon;\beta]^{[\overline{h},\overline{\beta}]}\in\mathbf{T}_*,\ \mu^{\overline{\mu}}\geq 0,\ \upsilon^{\overline{\upsilon}}\geq 0,\ \lambda^{\overline{\lambda}}\geq 0\\
\left[\begin{array}{c|c} U_s & -Q^TA_sP\\ \hline -P^TA_s^TQ & V_s\end{array}\right]^{\left[\begin{array}{c|c}\overline{U}_s & \overline{A}_s\\ \hline \overline{A}_s^T & \overline{V}_s\end{array}\right]}\succeq 0,\ s\leq S\\
\left[\begin{array}{c|c} U^t & -Q^TL_t^T\\ \hline -L_tQ & \lambda_tI_{p_t}\end{array}\right]^{\left[\begin{array}{c|c}\overline{U}^t & \overline{L}_t^T\\ \hline \overline{L}_t & \overline{\Lambda}_t\end{array}\right]}\succeq 0,\ \ [V^t-\lambda_tP^TR_t^TR_tP]^{\overline{V}^t}\succeq 0,\ t\leq T\\
[\sum_\ell \mu_\ell S_\ell-\sum_s U_s-\sum_t U^t]^{\overline{S}}\succeq 0,\ \ [\sum_k \upsilon_kT_k-\sum_s V_s-\sum_t V^t]^{\overline{T}}\succeq 0 \qquad (30)
\end{array}
$$

(superscripts are the Lagrange multipliers for the corresponding constraints). (P) clearly is solvable and strictly feasible, so that 2Opt is the optimal value of the (solvable!) conic dual of (P):

$$
\begin{array}{l}
2\mathrm{Opt} = \max\limits_{\substack{\overline{\alpha},\overline{\beta},\overline{g},\overline{h},\overline{\mu},\overline{\upsilon},\overline{\lambda},\overline{S},\overline{T},\\ \overline{U}_s,\overline{V}_s,\overline{A}_s,\overline{U}^t,\overline{L}_t,\overline{\Lambda}_t,\overline{V}^t}} 2\sum_s \mathrm{Tr}(Q^TA_sP\overline{A}_s^T)+2\sum_t\mathrm{Tr}(Q^TL_t^T\overline{L}_t) \qquad (D)\\
\text{subject to}\\
[\overline{g};\overline{\alpha}]\in\mathbf{T},\ [\overline{h};\overline{\beta}]\in\mathbf{S},\ \overline{\mu}\geq 0,\overline{\upsilon}\geq 0,\ \overline{\lambda}\geq 0,\ \overline{V}^t\succeq 0,\overline{S}\succeq 0,\overline{T}\succeq 0\\
\left[\begin{array}{c|c}\overline{U}_s & \overline{A}_s\\ \hline \overline{A}_s^T & \overline{V}_s\end{array}\right]\succeq 0,\ \left[\begin{array}{c|c}\overline{U}^t & \overline{L}_t^T\\ \hline \overline{L}_t & \overline{\Lambda}_t\end{array}\right]\succeq 0\\
\overline{\alpha}=1,\ [\overline{g};\overline{\alpha}]\in\mathbf{S},\ \overline{\beta}=1,\ [\overline{h};\overline{\beta}]\in\mathbf{T},\\
-\overline{g}_\ell+\mathrm{Tr}(\overline{S}S_\ell)+\overline{\mu}_\ell=0,\ -\overline{h}_k+\mathrm{Tr}(\overline{T}T_k)+\overline{\upsilon}_k=0,\\
\mathrm{Tr}(\overline{\Lambda}_t)-\mathrm{Tr}(\overline{V}_tP^TR_t^TR_tP)+\overline{\lambda}_t=0,\\
\overline{U}_s=\overline{S},\ \overline{U}_t=\overline{S},\ \overline{V}_s=\overline{T},\ \overline{V}^t=\overline{T}
\end{array}
$$

(here and in what follows the constraints should be satisfied for all values of "free indexes" $s\leq S$, $t\leq T$, $\ell\leq L$, $k\leq K$). Taking into account that relation $\left[\begin{array}{c|c}X & Y\\ \hline Y^T & Z\end{array}\right]\succeq 0$ is equivalent to $X\succeq 0, Z\succeq 0$, and $Y=X^{1/2}\Delta Z^{1/2}$ with $\|\Delta\|_{2,2}\leq 1$, and that $[\overline{g};1]\in\mathbf{S}$, $[\overline{h};1]\in\mathbf{T}$ is the same as $\overline{g}\in\mathcal{S}$, $\overline{h}\in\mathcal{T}$, (D) boils down to

$$
\mathrm{Opt} = \max_{\substack{\overline{g},\overline{h},\overline{S},\overline{T},\\ \overline{\Delta}_s,\overline{\delta}_t,\overline{\Lambda}_t}} \left\{\sum_s\mathrm{Tr}(Q^TA_sP\overline{A}_s^T)+\sum_t\mathrm{Tr}(Q^TL_t^T\overline{L}_t):\ \begin{array}{l}\overline{g}\in\mathcal{T},\ \overline{h}\in\mathcal{S},\ \overline{S}\succeq 0,\ \overline{T}\succeq 0,\ \mathrm{Tr}(\overline{S}S_\ell)\leq\overline{g}_\ell,\ \mathrm{Tr}(\overline{T}T_k)\leq\overline{h}_k\\ \overline{A}_s=\overline{S}^{1/2}\overline{\Delta}_s\overline{T}^{1/2},\ \|\overline{\Delta}_s\|_{2,2}\leq 1,\overline{L}_t^T=\overline{S}^{1/2}\overline{\delta}_t\overline{\Lambda}_t^{1/2},\ \|\overline{\delta}_t\|_{2,2}\leq 1\\ \mathrm{Tr}(\overline{\Lambda}_t)\leq\mathrm{Tr}(\overline{T}^{1/2}P^TR_t{}^TR_tP\overline{T}^{1/2})\end{array}\right\}
$$

or, which is the same,

$$\mathrm{Opt} = \max_{\substack{\overline{g},\overline{h},\overline{S},\overline{T}\\ \overline{\Delta}_s,\overline{\delta}_t,\overline{\Lambda}_t,\overline{L}_t}} \left\{ \sum\nolimits_s \mathrm{Tr}(\overline{S}^{1/2}Q^T A_s P\overline{T}^{1/2}\overline{\Delta}_s^T) + 2\sum\nolimits_t \mathrm{Tr}(\overline{S}^{1/2}Q^T L_t^T \overline{\Lambda}_t^{1/2}\overline{\delta}_t^T) : \begin{array}{l} \overline{g}\in\mathcal{T},\, \overline{h}\in\mathcal{S},\, \overline{S}\succeq 0,\, \overline{T}\succeq 0,\, \mathrm{Tr}(\overline{S}S_\ell)\le \overline{g}_\ell,\, \mathrm{Tr}(\overline{T}T_k)\le\overline{h}_k \\ \|\overline{\Delta}_s\|_{2,2}\le 1,\, \|\overline{\delta}_t\|_{2,2}\le 1 \\ \mathrm{Tr}(\overline{\Lambda}_t)\le \mathrm{Tr}(\overline{T}^{1/2}P^T R_t^T R_t P\overline{T}^{1/2}),\, \overline{\Lambda}_t\succeq 0 \end{array}\right\} \tag{D'}$$

Note that for Δ and δ such that $\|\Delta\|_{2,2}\le 1$ and $\|\delta\|_{2,2}\le 1$ one has

$$\mathrm{Tr}(A\Delta)\le \|A\|_{\mathrm{nuc}} = \|\lambda(\mathcal{L}[A])\|_1,\ \mathcal{L}[A] = \left[\begin{array}{c|c} & \frac{1}{2}A \\ \hline \frac{1}{2}A^T & \end{array}\right]$$

and

$$\mathrm{Tr}(AB^T\delta) = \langle A,\delta^T B\rangle_{\mathrm{Fro}} \le \|A\|_{\mathrm{Fro}}\|\delta^T B\|_{\mathrm{Fro}} \le \|A\|_{\mathrm{Fro}}\|B\|_{\mathrm{Fro}}$$

(here $\|A\|_{\mathrm{nuc}}$ stands for the nuclear norm and $\lambda(A)$ for the vector of eigenvalues of a symmetric matrix A). Consequently, for a feasible solution to (D′) it holds

$$\mathrm{Tr}(\overline{S}^{1/2}Q^T A_s P\overline{T}^{1/2}\overline{\Delta}_s^T) \le \|\lambda(\mathcal{L}[\overline{S}^{1/2}Q^T A_s P\overline{T}^{1/2}])\|_1,$$

and

$$\mathrm{Tr}(\overline{S}^{1/2}Q^T L_t^T\overline{\Lambda}_t^{1/2}\overline{\delta}_t^T) \le \|\overline{S}^{1/2}Q^T L_t^T\|_{\mathrm{Fro}}\|\overline{\Lambda}_t^{1/2}\|_{\mathrm{Fro}}.$$

The latter bound combines with the last constraint in (D′) to imply that

$$\mathrm{Tr}(\overline{S}^{1/2}Q^T L_t^T\overline{\Lambda}_t^{1/2}\overline{\delta}_t^T) \le \|\overline{S}^{1/2}Q^T L_t^T\|_{\mathrm{Fro}}\|\overline{T}^{1/2}P^T R_t^T\|_{\mathrm{Fro}},$$

and we conclude that

$$\mathrm{Opt} \le \max_{\overline{S},\overline{g},\overline{T},\overline{h}} \left\{ \sum\nolimits_s \left\|\lambda(\mathcal{L}[\overline{S}^{1/2}Q^T A_s P\overline{T}^{1/2}])\right\|_1 + \sum\nolimits_t \left\|\overline{S}^{1/2}Q^T L_t^T\|_{\mathrm{Fro}}\|\overline{T}^{1/2}P^T R_t^T\right\|_{\mathrm{Fro}} : \begin{array}{l} \overline{S}\succeq 0, \overline{g}\in\mathcal{S},\, \mathrm{Tr}(\overline{S}S_\ell)\le\overline{g}_\ell,\, \ell\le L \\ \overline{T}\succeq 0, \overline{h}\in\mathcal{T},\, \mathrm{Tr}(\overline{T}T_k)\le\overline{h}_k,\, k\le K \end{array}\right\} \tag{31}$$

4^o. We need the following result:

Lemma A.1 *[4, Lemma 2.3] (cf. also [3, Lemma 3.4.3]) If the ranks of all matrices A_s (and thus—matrices $\overline{S}^{1/2}Q^T A_s P\overline{T}^{1/2}$) do not exceed a given $\kappa\ge 1$, then for $\omega\sim\mathcal{N}(0,I_{M+N})$ one has*

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

with $\vartheta(\cdot)$ as described in Proposition A.1.

Our next result is as follows (cf. [2, Proposition B.4.12])

Lemma A.2 *Let $\in\mathbf{R}^{p\times q}$, $B\in\mathbf{R}^{r\times q}$ and $\xi\sim\mathcal{N}(Q,I_q)$. Then*

$$\mathbf{E}_\xi\left\{\|A\xi\|_2\|B\xi\|_2\right\} \ge \frac{2}{\pi}\|A\|_{\mathrm{Fro}}\|\|B\|_{\mathrm{Fro}}.$$

Proof. Setting $A^TA = U\text{Diag}\{\lambda\}U^T$ with orthogonal U and $\zeta = U^T\xi$, we have

$$\mathbf{E}\left\{\|A\xi\|_2\|B\xi\|_2\right\} = \mathbf{E}\left\{\sqrt{\sum\nolimits_{i=1}^q \lambda_i [U^T\xi]_i^2}\|B\xi\|_2\right\}.$$

The right hand side is concave in λ, so that the infimum of this function in λ varying in the simplex $\sum_i \lambda_i = \text{Tr}(A^TA)$ is attained at an extreme point. In other words, there exists vector $a \in \mathbf{R}^q$ with $a^Ta = \|A\|_{\text{Fro}}^2$ such that

$$\mathbf{E}\left\{\|A\xi\|_2\|B\xi\|_2\right\} \geq \mathbf{E}_\xi\left\{|a^T\xi|\,\|B\xi\|_2\right\}.$$

Applying the same argument to $\|B\xi\|_2$-factor, we can now find a vector $b \in \mathbf{R}^q$, $b^Tb = \|B\|_{\text{Fro}}^2$, such that

$$\mathbf{E}_\xi\left\{|a^T\xi|\,\|B\xi\|_2\right\} \geq \mathbf{E}_\xi\left\{|a^T\xi|\,|b^T\xi|\right\}.$$

It suffices to prove that the concluding quantity is $\geq 2\|a\|_2\|b\|_2/\pi$. By homogeneity, this is the same as to prove that if $[s;t] \sim \mathcal{N}(0, I_2)$, then $\mathbf{E}\{|t|\,|\cos(\phi)t + \sin(\phi)s|\} \geq \frac{2}{\pi}$ for all $\phi \in [0, 2\pi)$, which is straightforward (for the justification, see the proof of Proposition 2.3 of [5]). □

The last building block is the following

Lemma A.3 *[12, Lemma 6] Let*

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

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

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

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

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

with $\varkappa(\cdot)$ given by (26).

4^o Now we can complete the proof of the second inequality in (24). Let $\kappa \geq 1$, and let $\overline{g}, \overline{S}, \overline{h}, \overline{T}$ be feasible for the optimization problem in (31). Denoting by $\|\cdot\|_{\mathcal{Q}}$ the norm with the unit ball $\mathcal{Q}$, for all $A \in \mathbf{R}^{m\times n}$, $u \in \mathbf{R}^m$, and $v \in \mathbf{R}^n$ we have

$$u^TAv \leq \|u\|_{\mathcal{B}_*}\|Av\|_{\mathcal{B}} \leq \|u\|_{\mathcal{B}_*}\|A\|_{\mathcal{X},\mathcal{B}}\|v\|_{\mathcal{X}},$$

so that for all $u \in \mathbf{R}^m$ and $v \in \mathbf{R}^n$

$$\begin{aligned}\|u\|_{\mathcal{B}_*}\|v\|_{\mathcal{X}}\|\mathcal{A}\|_{\mathcal{X},\mathcal{B}} &\geq \max_{\substack{\epsilon_s, |\epsilon_s|\leq 1,\\ \delta_t, \|\delta_t\|_{2,2}\leq 1}} \left[\sum\nolimits_s \epsilon_s u^TA_sv + \sum\nolimits_t u^TL_t^T\delta_tR_tv\right]\\ &= \sum\nolimits_s |u^TA_sv| + \sum\nolimits_t \|L_tu\|_2\|R_tv\|_2.\end{aligned}$$

Thus, for all $\overline{g}, \overline{S}, \overline{h}, \overline{T}$ which are feasible for (31) and $\xi \in \mathbf{R}^M$, $\eta \in \mathbf{R}^N$,

$$\begin{aligned}&\|\overline{S}^{1/2}\xi\|_{\mathcal{Z}}\|\overline{T}^{1/2}\eta\|_{\mathcal{Y}}\|\mathcal{A}\|_{\mathcal{X},\mathcal{B}} \geq \|Q\overline{S}^{1/2}\xi\|_{\mathcal{B}_*}\|P\overline{T}^{1/2}\eta\|_{\mathcal{X}}\|\mathcal{A}\|_{\mathcal{X},\mathcal{B}} \text{ [due to } \mathcal{B}_* = Q\mathcal{Z}, \mathcal{X} = P\mathcal{Y}]\\ &\geq \sum\nolimits_s |\xi^T\overline{S}^{1/2}Q^TA_sP\overline{T}^{1/2}\eta| + \sum\nolimits_t \|L_tQ\overline{S}^{1/2}\xi\|_2\|R_tP\overline{T}^{1/2}\eta\|_2\\ &= \sum\nolimits_s |[\xi;\eta]^T\mathcal{L}[\overline{S}^{1/2}Q^TA_sP\overline{T}^{1/2}][\xi;\eta]|\\ &+ \sum\nolimits_t \|[L_tQ\overline{S}^{1/2}, 0_{p_t\times N}][\xi;\eta]\|_2\|[0_{q_t\times M}, R_tP\overline{T}^{1/2}][\xi;\eta]\|_2.\end{aligned} \tag{32}$$

As a result, for $[\xi;\eta]\sim\mathcal{N}(0,I_{M+N})$, applying the bounds of Lemmas A.1 and A.2,

$$\begin{array}{l}
\mathbf{E}\left\{\left\|\overline{S}^{1/2}\xi\right\|_{\mathcal{Z}}\right\}\mathbf{E}\left\{\left\|\overline{T}^{1/2}\eta\right\|_{\mathcal{Y}}\right\}\|\mathcal{A}\|_{\mathcal{X},\mathcal{B}}=\mathbf{E}\left\{\left\|\overline{S}^{1/2}\xi\right\|_{\mathcal{Z}}\|\overline{T}^{1/2}\eta\|_{\mathcal{Y}}\|\mathcal{A}\|_{\mathcal{X},\mathcal{B}}\right\}\\
\geq\sum_s\mathbf{E}\left\{\left|[\xi;\eta]^T\mathcal{L}[\overline{S}^{1/2}Q^TA_sP\overline{T}^{1/2}][\xi;\eta]\right|\right\}\\
\quad+\sum_t\mathbf{E}\left\{\left\|[L_tQ\overline{S}^{1/2},0_{p_t\times N}][\xi;\eta]\right\|_2\left\|[0_{q_t\times M},R_tP\overline{T}^{1/2}][\xi;\eta]\right\|_2\right\}\\
\geq\vartheta(2\kappa)^{-1}\sum_s\left\|\lambda\left(\mathcal{L}[\overline{S}^{1/2}Q^TA_sP\overline{T}^{1/2}]\right)\right\|_1+\frac{2}{\pi}\sum_t\|L_tQ\overline{S}^{1/2}\|_{\rm Fro}\|R_tP\overline{T}^{1/2}\|_{\rm Fro}.
\end{array}$$

Besides this, by Lemma A.3 we have

$$\mathbf{E}\left\{\|\overline{S}^{1/2}\xi\|_{\mathcal{Z}}\right\}\leq\varkappa(L),\ \ \mathbf{E}\left\{\|\overline{T}^{1/2}\eta\|_{\mathcal{Y}}\right\}\leq\varkappa(K)$$

due to the fact that $\overline{g},\overline{S},\overline{h}$ and $\overline{T}$ are feasible for (31). This combines with (32) to imply that the value $\varkappa(L)\varkappa(K)\|\mathcal{A}\|_{\mathcal{X},\mathcal{B}}$ is lower bounded with the quantity

$$\begin{array}{l}
\max\left[\vartheta(2\kappa),\pi/2\right]^{-1}\left[\sum_s\left\|\lambda\left(\mathcal{L}[\overline{S}^{1/2}Q^TS_sP\overline{T}^{1/2}]\right)\right\|_1\right.\\
\qquad\qquad\left.+\sum_t\|\overline{S}^{1/2}Q^TL_t^T\|_{\rm Fro}\|\overline{T}^{1/2}P^TR_t^T\|_{\rm Fro}\right].
\end{array}$$

Invoking the inequality in (31), we arrive at the second inequality in (24). The above reasoning assumed that $\kappa\geq 1$, with evident simplifications, it is applicable to the case of $\kappa=0$ as well. □

A.2.3 Proof of Proposition 2.1

We put $S=q_{\rm S}$ and $T=q-q_{\rm S}$. In the situation of Proposition 2.1 we want to tightly upper-bound quantity

$$\begin{array}{rl}
\mathfrak{s}(H)=&\max\limits_{x\in\mathcal{X},\eta\in\mathcal{U}}\left\|H^TD[\eta]x\right\|\\
=&\max\limits_{\ell\leq L}\max\limits_{x\in\mathcal{X},\eta\in\mathcal{U}}\left\{\sqrt{[H^TD[\eta]x]^TR_\ell[H^TD[\eta]x]}\right\}\\
=&\max\limits_{\ell\leq L}\|\mathcal{A}_\ell[H]\|_{\mathcal{X},2},
\end{array}$$

where $\|\cdot\|_{\mathcal{X},2}$ is the operator norm induced by $\|\cdot\|_{\mathcal{X}}$ on the argument and $\|\cdot\|_2$ on the image space and the uncertain matrix $\mathcal{A}_\ell[H]$ is given by

$$\mathcal{A}_\ell=\left\{\sum_{s=1}^S\delta_s\underbrace{R_\ell^{1/2}H^TP_s^TQ_s}_{=:A_{s\ell}[H]}+\sum_{t=1}^T\underbrace{R_\ell^{1/2}H^TP_{S+t}^T}_{L_{t\ell}^T[H]}\Delta_s\underbrace{Q_{S+t}}_{=:R_t}:\begin{array}{ll}|\delta_s|\leq 1&,1\leq s\leq S\\ \|\Delta_s\|_{2,2}\leq 1&,1\leq t\leq T\end{array}\right\}$$

It follows that

$$\mathfrak{s}(H)=\max_{\ell\leq L}\|\mathcal{A}_\ell[H]\|_{\mathcal{X},2},$$

and Proposition A.1 provides us with the efficiently computable convex in H upper bound $\bar{\mathfrak{s}}(H)$ on $\mathfrak{s}(H)$:

$$\bar{\mathfrak{s}}(H) = \max_{\ell\leq L} \mathrm{Opt}_\ell(H),$$

$$\mathrm{Opt}_\ell(H) = \min_{\mu,\upsilon,\lambda,U_s,V_s,U^t,V^t} \left\{ \tfrac{1}{2}[\mu+\phi_{\mathcal{T}}(\upsilon)] : \mu\geq 0,\, \upsilon\geq 0,\, \lambda\geq 0 \right.$$
$$\left. \begin{array}{l} \left[\begin{array}{c|c} U_s & -A_{s\ell}[H]P \\ \hline -P^TA_{s\ell}^T[H] & V_s \end{array}\right] \succeq 0 \\ \left[\begin{array}{c|c} U^t & -L_{t\ell}^T[H] \\ \hline -L_{t\ell}[H] & \lambda_t I_{p_{q_{\mathrm{S}}+t}} \end{array}\right] \succeq 0 \\ V^t - \lambda_t P^T R_t^T R_t P \succeq 0 \\ \mu I_\nu - \sum_s U_s - \sum_t U^t \succeq 0 \\ \sum_k \upsilon_k T_k - \sum_s V_s - \sum_t V^t \succeq 0 \end{array} \right\}$$

and tightness factor of this bound does not exceed $\max[\vartheta(2\kappa), \pi/2]$ where $\kappa = \max_{\alpha\leq q_{\mathrm{S}}} \min[p_\alpha, q_\alpha]$.
□

A.3 Spectratopic version of Proposition A.1

Proposition A.1 admits a "spectratopic version," in which ellitopes $\mathcal{X}$ and $\mathcal{B}_*$ given by (22) are replaced by the pair of *spectratopes*

$$\begin{array}{l} \mathcal{X} = \{Py : y\in\mathcal{Y}\}\subset\mathbf{R}^n, \mathcal{Y} = \{y\in\mathbf{R}^N \ \&\ \exists t\in\mathcal{T} : T_k[y]^2 \preceq t_k I_{f_k}, k\leq K\}, \\ \quad T_k[y] = \sum_{j=1}^N y_j T^{jk},\ T^{jk}\in\mathbf{S}^{f_k}, \sum_k T_k^2[y] \succ 0\ \forall y\neq 0 \end{array} \tag{33a}$$

$$\begin{array}{l} \mathcal{B}_* = \{Qz : z\in\mathcal{Z}\}\subset\mathbf{R}^m,\ \mathcal{Z} = \{z\in\mathbf{R}^M : \exists s\in\mathcal{S} : S_\ell^2[z] \preceq s_\ell I_{d_\ell},\ \ell\leq L\}, \\ \quad S_\ell[z] = \sum_{j=1}^M z_j S^{jk\ell},\ S^{jk\ell}\in\mathbf{S}^{d_\ell}, \sum_\ell S_\ell^2[z]\succ 0\ \forall z\neq 0 \end{array} \tag{33b}$$

The spectratopic version of the statement reads as follows:

Proposition A.2 *Given uncertain matrix (21) and spectratopes (33a) and (33b), consider convex optimization problem*

$$\mathrm{Opt} = \min_{\mu,\upsilon,\lambda,U_s,V_s,U^t,V^t} \left\{ \tfrac{1}{2}[\phi_{\mathcal{S}}(\lambda[\mu]) + \phi_{\mathcal{T}}(\lambda[\upsilon])] : \right.$$

subject to

$$\mu = \{M_\ell\in\mathbf{S}_+^{d_\ell}, \ell\leq L\},\ \upsilon = \{\Upsilon_k\in\mathbf{S}_+^{f_k}, k\leq K\},\ \lambda\geq 0$$

$$\left[\begin{array}{c|c} U_s & -Q^TA_sP \\ \hline -P^TA_s^TQ & V_s \end{array}\right] \succeq 0 \tag{34a}$$

$$\left[\begin{array}{c|c} U^t & -Q^TL_t^T \\ \hline -L_tQ & \lambda_t I_{p_t} \end{array}\right] \succeq 0,\ V^t - \lambda_t P^T R_t^T R_t P \succeq 0 \tag{34b}$$

$$\sum_\ell S_\ell^{+,*}[M_\ell] - \sum_s U_s - \sum_t U^t \succeq 0 \tag{34c}$$

$$\sum_k T_k^{+,*}[\Upsilon_k] - \sum_s V_s - \sum_t V^t \succeq 0 \tag{34d}$$

where

$$\lambda[\zeta] = [\mathrm{Tr}(Z_1); ...; \mathrm{Tr}(Z_I)] \ \textit{for}\ \zeta = \{Z_i\in\mathbf{S}^{k_i}, i\leq I\}$$

and

$$S_\ell^{+,*}[V] = \left[\mathrm{Tr}(VS^{i\ell}S^{j\ell})\right]_{i,j\leq M}\ \textit{for}\ V\in\mathbf{S}^{d_\ell},\ T_k^{+,*}[U] = \left[\mathrm{Tr}(UT^{ik}T^{jk})\right]_{i,j\leq N}\ \textit{for}\ U\in\mathbf{S}^{f_k}.$$

Problem (34) is strictly feasible and solvable, and

$$\|\mathcal{A}\|_{\mathcal{X},\mathcal{B}} \le \mathrm{Opt} \le \varsigma\left(\sum\nolimits_k f_k\right)\varsigma\left(\sum\nolimits_\ell d_\ell\right)\max\left[\vartheta(2\kappa),\pi/2\right]\|\mathcal{A}\|_{\mathcal{X},\mathcal{B}}$$

where ϑ and κ are the same as in Proposition A.1 and

$$\varsigma(J) = 2\sqrt{2\ln(2J)}.$$

Proof. For $Y \in \mathbf{S}^M$ and $X \in \mathbf{S}^N$ let us set

$$S_\ell^+[Y] = \sum_{i,j=1}^M Y_{ij}S^{i\ell}S^{j\ell},\ T_k^+[X] = \sum_{i,j=1}^N X_{ij}T^{ik}T^{jk},$$

so that

$$S_\ell^+[zz^T] = S_\ell^2[z],\ T_k^+[yy^T] = T_k^2[y] \tag{35}$$

and

$$\begin{array}{l}\mathrm{Tr}(VS_\ell^+[Y]) = \mathrm{Tr}(S_\ell^{+,*}[V]Y) \text{ for } V \in \mathbf{S}^{d_\ell}, Y \in \mathbf{R}^M,\\ \mathrm{Tr}(UT_k^+[X]) = \mathrm{Tr}(T_k^{+,*}[U]X) \text{ for } U \in \mathbf{S}^{f_k}, X \in \mathbf{R}^N.\end{array} \tag{36}$$

The proof of Proposition A.2 is obtained from that (below referred to as "the proof") of Proposition A.1 by the following modifications:

1. All references to (23) should be replaced with references to (34). Item [b] in 1^o of the proof now reads
   [b′] $x = Py$ with $T_k^2[y] \preceq \tau_k I_{f_k}$, $k \le K$, for some $\tau \in \mathcal{T}$ and $u = Qz$ for some z such that $S_\ell^2[z] \preceq \varsigma_\ell I_{d_\ell}$, $\ell \le L$, for some $\varsigma \in \mathcal{S}$.
   The last three lines in the chain (29) are replaced with

$$\begin{aligned}\gamma &\le \frac{1}{2}\left[\sum_\ell \mathrm{Tr}([zz^T]S_\ell^{+,*}[M_\ell]) + \sum_k \mathrm{Tr}([yy^T]T_k^{+,*}[\Upsilon_k])\right] \text{ [by (34c) and (34d)]}\\ &= \frac{1}{2}\left[\sum_\ell \mathrm{Tr}(S_\ell^2[z]M_\ell) + \sum_k \mathrm{Tr}(T_k^2[y]\Upsilon_k)\right] \text{ [by (35) and (36)]}\\ &\le \frac{1}{2}\left[\sum_\ell \varsigma_\ell \mathrm{Tr}(M_\ell) + \sum_k \tau_k \mathrm{Tr}(\Upsilon_k)\right] \text{ [due to (b′) and } M_\ell \succeq 0, \Upsilon_k \succeq 0]\\ &\le \tfrac{1}{2}[\phi_{\mathcal{S}}(\lambda[\mu]) + \phi_{\mathcal{T}}(\lambda[\upsilon])] \text{ [since } \varsigma \in \mathcal{S}, \tau \in \mathcal{T}].\end{aligned}$$

2. Constraints (30) in (P) now read

$$\left[\sum_\ell S_\ell^{+,*}[M_\ell] - \sum_s U_s - \sum_t U^t\right]^{\overline{S}} \succeq 0,\quad \left[\sum_k T_k^{+,*}[\Upsilon_k] - \sum_s V_s - \sum_t V^t\right]^{\overline{T}} \succeq 0.$$

   As a result, (31) becomes

$$\mathrm{Opt} \le \max_{\overline{S},\overline{g},\overline{T},\overline{h}}\left\{\sum_s \left\|\lambda(\mathcal{L}[\overline{S}^{1/2}Q^TA_sP\overline{T}^{1/2}])\right\|_1 + \sum_t \|\overline{S}^{1/2}Q^TL_t^T\|_{\mathrm{Fro}}\|\overline{T}^{1/2}P^TR_t^T\|_{\mathrm{Fro}} : \begin{array}{l}\overline{S} \succeq 0, \overline{g} \in \mathcal{S}, S_\ell^+[\overline{S}] \preceq \overline{g}_\ell I_{d_\ell},\ \ell \le L\\ \overline{T} \succeq 0, \overline{h} \in \mathcal{T}, T_k^+[\overline{T}] \preceq \overline{h}_k I_{f_k},\ k \le K\end{array}\right\}$$

3. The role of Lemma A.3 in the proof is now played by the following fact.
   Lemma A.4 *[12, Lemma 8] Let*

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

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

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

   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

$$\mathbf{E}\{\rho(\omega)\} \le \varsigma\left(\sum\nolimits_j \nu_j\right),\ \varsigma(F) = 2\sqrt{2\ln(2F)}.$$

References

1. Y. Bekri, A. Juditsky, and A. Nemirovski. Estimation from indirect observations under stochastic uncertainty in observation matrix. `https://www.isye.gatech.edu/~nemirovs/jota_part1_submitted.pdf`, 2023. – to appear in *Journal of Optimization Theory and Applications*].
2. A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*, volume 28 of *Princeton Series in Applied Mathematics*. Princeton University Press, 2009.
3. A. Ben-Tal and A. Nemirovski. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, volume 2 of *SIAM Studies in Applied Mathematics*. SIAM, 2001.
4. A. Ben-Tal and A. Nemirovski. On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty. *SIAM Journal on Optimization*, 12(3):811–833, 2002.
5. A. Ben-Tal, A. Nemirovski, and C. Roos. Extended matrix cube theorems with applications to μ-theory in control. *Mathematics or Operations Research*, 28(3):497–523, 2003.
6. D. Bertsekas and I. Rhodes. Recursive state estimation for a set-membership description of uncertainty. *IEEE Transactions on Automatic Control*, 16(2):117–128, 1971.
7. S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *SIAM Studies in Applied Mathematics*. SIAM, 1994.
8. M. Casini, A. Garulli, and A. Vicino. Feasible parameter set approximation for linear models with bounded uncertain regressors. *IEEE Transactions on Automatic Control*, 59(11):2910–2920, 2014.
9. V. Cerone. Feasible parameter set for linear models with bounded errors in all variables. *Automatica*, 29(6):1551–1555, 1993.
10. J. Cope and B. Rust. Bounds on solutions of linear systems with inaccurate data. *SIAM Journal on Numerical Analysis*, 16(6):950–963, 1979.
11. N. J. Higham. *Accuracy and stability of numerical algorithms*. SIAM, 2002.
12. A. Juditsky, G. Kotsalis, and A. Nemirovski. Tight computationally efficient approximation of matrix norms with applications. *Open Journal of Mathematical Optimization*, 3(7):1–38, 2022.
13. A. Juditsky and A. Nemirovski. Near-optimality of linear recovery from indirect observations. *Mathematical Statistics and Learning*, 1(2):171–225, 2018.
14. A. Juditsky and A. Nemirovski. *Statistical Inference via Convex Optimization*. Princeton University Press, 2020.
15. V. Kreinovich, A. V. Lakeyev, and S. I. Noskov. Optimal solution of interval linear systems is intractable (np-hard). *Interval Computations*, 1:6–14, 1993.
16. A. Kurzhansky and I. Valyi. *Ellipsoidal calculus for estimation and control*. Birkhauser, 1997.
17. A. I. Matasov. *Estimators for uncertain dynamic systems*, volume 458. Springer Science & Business Media, 1998.
18. M. Milanese, J. Norton, H. Piet-Lahanier, and É. Walter. *Bounding approaches to system identification*. Springer Science & Business Media, 2013.
19. S. A. Nazin and B. T. Polyak. Interval parameter estimation under model uncertainty. *Mathematical and Computer Modelling of Dynamical Systems*, 11(2):225–237, 2005.
20. S. A. Nazin and B. T. Polyak. Ellipsoid-based parametric estimation in the linear multidimensional systems with uncertain model description. *Automation and Remote Control*, 68(6):993–1005, 2007.
21. A. Neumaier. *Interval methods for systems of equations*. Number 37. Cambridge University Press, 1990.
22. W. Oettli and W. Prager. Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides. *Numerische Mathematik*, 6(1):405–409, 1964.
23. B. 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.
24. R. Tempo and A. Vicino. Optimal algorithms for system identification: a review of some recent results. *Mathematics and computers in simulation*, 32(5-6):585–595, 1990.