Lecture 4

Globalized Robust Counterparts of Uncertain Linear and Conic Problems

In this lecture we extend the concept of Robust Counterpart in order to gain certain control on what happens when the actual data perturbations run out of the postulated perturbation set.

4.1 Globalized Robust Counterparts — Motivation and Definition

Let us come back to Assumptions A.1 – A.3 underlying the concept of Robust Counterpart and concentrate on **A.3**. This assumption is not a "universal truth" — in reality, there are indeed constraints that cannot be violated (e.g., you cannot order a negative supply), but also constraints whose violations, while undesirable, can be tolerated to some degree, (e.g., sometimes you can tolerate a shortage of a certain resource by implementing an "emergency measure" like purchasing it on the market, employing sub-contractors, taking out loans, etc.). Immunizing such "soft" constraints against data uncertainty should perhaps be done in a more flexible fashion than in the usual Robust Counterpart. In the latter, we ensure a constraint's validity for all realizations of the data from a given uncertainty set and do not care what happens when the data are outside of this set. For a soft constraint, we can take care of what happens in this latter case as well, namely, by ensuring *controlled deterioration* of the constraint when the data runs away from the uncertainty set. We are about to build a mathematically convenient model capturing the above requirements.

4.1.1 The Case of Uncertain Linear Optimization

Consider an uncertain linear constraint in variable x

$$[a^0 + \sum_{\ell=1}^{L} \zeta_\ell a^\ell]^T x \leq [b^0 + \sum_{\ell=1}^{L} \zeta_\ell b^\ell] \tag{4.1.1}$$

where ζ is the perturbation vector (cf. (1.3.4), (1.3.5)). Let $\mathcal{Z}_+$ be the set of all "physically possible" perturbations, and $\mathcal{Z} \subset \mathcal{Z}_+$ be the "normal range" of the perturbations — the one for which we insist on the constraint to be satisfied. With the usual RC approach, we treat $\mathcal{Z}$ as the only set of perturbations and require a candidate solution x to satisfy the constraint for all $\zeta \in \mathcal{Z}$. With our new approach, we add the requirement that *the violation of constraint in the case when* $\zeta \in \mathcal{Z}_+ \backslash \mathcal{Z}$ (that is a "physically possible" perturbation that is outside of the normal range) *should be bounded by a constant times the distance from* ζ *to* $\mathcal{Z}$. Both requirements — the validity of the constraint for $\zeta \in \mathcal{Z}$ and the bound on the constraint's violation when $\zeta \in \mathcal{Z}_+ \backslash \mathcal{Z}$ can be expressed by a single requirement

$$[a^0 + \sum_{\ell=1}^{L} \zeta_\ell a^\ell]^T x - [b^0 + \sum_{\ell=1}^{L} \zeta_\ell b^\ell] \leq \alpha \mathrm{dist}(\zeta, \mathcal{Z}) \quad \forall \zeta \in \mathcal{Z}_+,$$

where $\alpha \geq 0$ is a given "global sensitivity."

In order to make the latter requirement tractable, we add some structure to our setup. Specifically, let us assume that:

(G.a) The normal range $\mathcal{Z}$ of the perturbation vector ζ is a nonempty closed convex set;

(G.b) The set $\mathcal{Z}_+$ of all "physically possible" perturbations is the sum of $\mathcal{Z}$ and a closed convex cone $\mathcal{L}$:

$$\mathcal{Z}_+ = \mathcal{Z} + \mathcal{L} = \{\zeta = \zeta' + \zeta'' : \zeta' \in \mathcal{Z}, \zeta'' \in \mathcal{L}\}; \tag{4.1.2}$$

(G.c) We measure the distance from a point $\zeta \in \mathcal{Z}_+$ to the normal range $\mathcal{Z}$ of the perturbations in a way that is consistent with the structure (4.1.2) of $\mathcal{Z}_+$, specifically, by

$$\mathrm{dist}(\zeta, \mathcal{Z} \mid \mathcal{L}) = \inf_{\zeta'} \left\{ \|\zeta - \zeta'\| : \zeta' \in \mathcal{Z}, \zeta - \zeta' \in \mathcal{L} \right\}, \tag{4.1.3}$$

where $\|\cdot\|$ is a fixed norm on $\mathbb{R}^L$.

In what follows, we refer to a triple $(\mathcal{Z}, \mathcal{L}, \|\cdot\|)$ arising in (G.a–c) as a *perturbation structure* for the uncertain constraint (4.1.1).

Definition 4.1 *Given* $\alpha \geq 0$ *and a perturbation structure* $(\mathcal{Z}, \mathcal{L}, \|\cdot\|)$*, we say that a vector* x *is a globally robust feasible solution to uncertain linear constraint* (4.1.1) *with global sensitivity* α*, if* x *satisfies the semi-infinite constraint*

$$[a^0 + \sum_{\ell=1}^{L} \zeta_\ell a^\ell]^T x - [b^0 + \sum_{\ell=1}^{L} \zeta_\ell b^\ell] \leq \alpha \, \mathrm{dist}(\zeta, \mathcal{Z}|\mathcal{L}) \;\; \forall \zeta \in \mathcal{Z}_+ = \mathcal{Z} + \mathcal{L}. \tag{4.1.4}$$

We refer to the semi-infinite constraint (4.1.4) *as the Globalized Robust Counterpart (GRC) of the uncertain constraint* (4.1.1).

Note that global sensitivity $\alpha = 0$ corresponds to the most conservative attitude where the constraint must be satisfied for all physically possible perturbations; with $\alpha = 0$, the GRC becomes the usual RC of the uncertain constraint with $\mathcal{Z}_+$ in the role of the perturbation set. The larger α, the less conservative the GRC.

Now, given an uncertain Linear Optimization program with affinely perturbed data

$$\left\{ \min_x \left\{ c^T x : Ax \leq b \right\} : [A, b] = [A^0, b^0] + \sum_{\ell=1}^{L} \zeta_\ell [A^\ell, b^\ell] \right\} \tag{4.1.5}$$

(w.l.o.g., we assume that the objective is certain) and a perturbation structure $(\mathcal{Z}, \mathcal{L}, \|\cdot\|)$, we can replace every one of the constraints with its Globalized Robust Counterpart, thus ending up with the GRC of (4.1.5). In this construction, we can associate different sensitivity parameters α to different constraints. Moreover, we can treat these sensitivities as design variables rather than fixed parameters, add linear constraints on these variables, and optimize both in x *and* α an objective function that is a mixture of the original objective and a weighted sum of the sensitivities.

4.1.2 The Case of Uncertain Conic Optimization

Consider an uncertain conic problem (3.1.2), (3.1.3):

$$\min_x \left\{ c^T x + d : A_i x - b_i \in \mathbf{Q}_i,\ 1 \le i \le m \right\}, \tag{4.1.6}$$

where $\mathbf{Q}_i \subset \mathbb{R}^{k_i}$ are nonempty closed convex sets given by finite lists of conic inclusions:

$$\mathbf{Q}_i = \{u \in \mathbb{R}^{k_i} : Q_{i\ell} u - q_{i\ell} \in \mathbf{K}_{i\ell},\ \ell = 1, ..., L_i\}, \tag{4.1.7}$$

with closed convex pointed cones $\mathbf{K}_{i\ell}$, and let the data be affinely parameterized by the perturbation vector ζ:

$$(c, d, \{A_i, b_i\}_{i=1}^m) = (c^0, d^0, \{A_i^0, b_i^0\}_{i=1}^m) + \sum_{\ell=1}^L \zeta_\ell (c^\ell, d^\ell, \{A_i^\ell, b_i^\ell\}_{i=1}^m). \tag{4.1.8}$$

When extending the notion of Globalized Robust Counterpart from the case of Linear Optimization to the case of Conic one, we need a small modification. Assuming, same as in the former case, that the set $\mathcal{Z}_+$ of all "physically possible" realizations of the perturbation vector ζ is of the form $\mathcal{Z}_+ = \mathcal{Z} + \mathcal{L}$, where $\mathcal{Z}$ is the closed convex normal range of ζ and $\mathcal{L}$ is a closed convex cone, observe that in the conic case, as compared to the LO one, the left hand side of our uncertain constraint (4.1.6) is vector rather than scalar, so that a straightforward analogy of (4.1.4) does not make sense. Note, however, that when rewriting (4.1.1) in our present "inclusion form"

$$[a^0 + \sum_{\ell=1}^L \zeta_\ell a^\ell]^T x - [b^0 + \sum_{\ell=1}^L \zeta_\ell b^\ell] \in \mathbf{Q} \equiv \mathbb{R}_-, \tag{$*$}$$

relation (4.1.4) says exactly that the distance from the left hand side of $(*)$ to $\mathbf{Q}$ does not exceed $\alpha \mathrm{dist}(\zeta, \mathcal{Z}|\mathcal{L})$ for all $\zeta \in \mathcal{Z} + \mathcal{L}$. In this form, the notion of global sensitivity admits the following multi-dimensional extension:

Definition 4.2 *Consider an uncertain convex constraint*

$$[P_0 + \sum_{\ell=1}^L \zeta_\ell P_\ell] y - [p^0 + \sum_{\ell=1}^L \zeta_\ell p^\ell] \in \mathbf{Q}, \tag{4.1.9}$$

where $\mathbf{Q}$ is a nonempty closed convex subset in $\mathbb{R}^k$. Let $\|\cdot\|_Q$ be a norm on $\mathbb{R}^k$, $\|\cdot\|_Z$ be a norm on $\mathbb{R}^L$, $\mathcal{Z} \subset \mathbb{R}^L$ be a nonempty closed convex normal range of perturbation ζ, and $\mathcal{L} \subset\in \mathbb{R}^L$ be a *closed convex cone. We say that a candidate solution y is robust feasible, with global sensitivity α, for* (4.1.9)*, under the perturbation structure* $(\|\cdot\|_Q, \|\cdot\|_Z, \mathcal{Z}, \mathcal{L})$*, if*

$$\begin{array}{c}\mathrm{dist}([P_0+\sum_{\ell=1}^L \zeta_\ell P_\ell]y-[p^0+\sum_{\ell=1}^L \zeta_\ell p^\ell], \mathbf{Q}) \leq \alpha \mathrm{dist}(\zeta, \mathcal{Z}|\mathcal{L}) \\ \hfill \forall \zeta \in \mathcal{Z}_+ = \mathcal{Z}+\mathcal{L} \\ \left[\begin{array}{rcl}\mathrm{dist}(u,\mathbf{Q}) & = & \min\limits_v \{\|u-v\|_Q : v \in \mathbf{Q}\} \\ \mathrm{dist}(\zeta, \mathcal{Z}|\mathcal{L}) & = & \min\limits_v \{\|\zeta - v\|_Z : v \in \mathcal{Z}, \zeta - v \in \mathcal{L}\}\end{array}\right].\end{array} \tag{4.1.10}$$

Sometimes it is necessary to add some structure to the latter definition. Specifically, assume that the space $\mathbb{R}^L$ where ζ lives is given as a direct product:

$$\mathbb{R}^L = \mathbb{R}^{L_1} \times \ldots \times \mathbb{R}^{L_S}$$

and let $\mathcal{Z}^s \subset \mathbb{R}^{L_s}$, $\mathcal{L}^s \subset \mathbb{R}^{L_s}$, $\|\cdot\|_s$ be, respectively, closed nonempty convex set, closed convex cone and a norm on $\mathbb{R}^{L_s}$, $s = 1, ..., S$. For $\zeta \in \mathbb{R}^L$, let ζ^s, $s = 1, ..., S$, be the projections of ζ onto the direct factors $\mathbb{R}^{L_s}$ of $\mathbb{R}^L$. The "structured version" of Definition 4.2 is as follows:

Definition 4.3 *A candidate solution y to the uncertain constraint* (4.1.9) *is robust feasible with global sensitivities α_s, $1 \leq s \leq S$, under the perturbation structure* $(\|\cdot\|_Q, \{\mathcal{Z}^s, \mathcal{L}^s, \|\cdot\|_s\}_{s=1}^S)$*, if*

$$\begin{array}{c}\mathrm{dist}([P_0+\sum_{\ell=1}^L \zeta_\ell P_\ell]y-[p^0+\sum_{\ell=1}^L \zeta_\ell p^\ell], \mathbf{Q}) \leq \sum_{s=1}^S \alpha_s \mathrm{dist}(\zeta^s, \mathcal{Z}^s|\mathcal{L}^s) \\ \hfill \forall \zeta \in \mathcal{Z}_+ = \underbrace{(\mathcal{Z}^1 \times \ldots \times \mathcal{Z}^S)}_{\mathcal{Z}} + \underbrace{\mathcal{L}^1 \times \ldots \times \mathcal{L}^S}_{\mathcal{L}} \\ \left[\begin{array}{rcl}\mathrm{dist}(u,\mathbf{Q}) & = & \min\limits_v \{\|u-v\|_Q : v \in \mathbf{Q}\} \\ \mathrm{dist}(\zeta^s, \mathcal{Z}^s|\mathcal{L}^s) & = & \min\limits_{v^s} \{\|\zeta^s - v^s\|_s : v^s \in \mathcal{Z}^s, \zeta^s - v^s \in \mathcal{L}^s\}.\end{array}\right]\end{array} \tag{4.1.11}$$

Note that Definition 4.2 can be obtained from Definition 4.3 by setting $S = 1$. We refer to the semi-infinite constraints (4.1.10), (4.1.11) as to *Globalized Robust Counterparts* of the uncertain constraint (4.1.9) w.r.t. the perturbations structure in question. When building the GRC of uncertain problem (4.1.6), (4.1.8), we first rewrite it as an uncertain problem

$$\min_{y=(x,t)} \left\{ t : \begin{array}{l} \overbrace{c^Tx+d-t \equiv [c^0+\sum_{\ell=1}^L \zeta_\ell c^\ell]^T x + [d^0+\sum_{\ell=1}^L \zeta_\ell d^\ell] - t}^{[P_{00}+\sum_{\ell=1}^L \zeta_\ell P_{0\ell}]y-[p_0^0+\sum_{\ell=1}^L \zeta_\ell p_0^\ell]} \in \mathbf{Q}_0 \equiv \mathbb{R}_- \\ A_i x - b_i \equiv \underbrace{[A_i^0+\sum_{\ell=1}^L \zeta_\ell A_i^\ell]x - [b_i^0+\sum_{\ell=1}^L \zeta_\ell b_i^\ell]}_{[P_{i0}+\sum_{\ell=1}^L \zeta_\ell P_{i\ell}]y-[p_i^0+\sum_{\ell=1}^L \zeta_\ell p_i^\ell]} \in \mathbf{Q}_i,\ 1 \leq i \leq m \end{array} \right\}$$

with certain objective, and then replace the constraints with their Globalized RCs. The underlying perturbation structures and global sensitivities may vary from constraint to constraint.

4.1.3 Safe Tractable Approximations of GRCs

A Globalized RC, the same as the plain one, can be computationally intractable, in which case we can look for the second best thing — a safe tractable approximation of the GRC. This notion is defined as follows (cf. Definition 3.2):

Definition 4.4 *Consider the uncertain convex constraint* (4.1.9) *along with its GRC* (4.1.11). *We say that a system $\mathcal{S}$ of convex constraints in variables y, $\alpha = (\alpha_1, ..., \alpha_S) \geq 0$, and, perhaps, additional variables u, is a safe approximation of the GRC, if the projection of the feasible set of $\mathcal{S}$ on the space of (y, α) variables is contained in the feasible set of the GRC:*

$$\begin{array}{l}\forall(\alpha = (\alpha_1, ..., \alpha_S) \geq 0, y):\\ (\exists u: \ (y, \alpha, u) \textit{ satisfies } \mathcal{S}) \Rightarrow (y, \alpha) \textit{ satisfies } (4.1.11).\end{array}$$

This approximation is called tractable, provided that $\mathcal{S}$ is so, (e.g., $\mathcal{S}$ is an explicit system of CQIs/LMIs of, more general, the constraints in $\mathcal{S}$ are efficiently computable).

When quantifying the tightness of an approximation, we, as in the case of RC, assume that the normal range $\mathcal{Z} = \mathcal{Z}^1 \times ... \times \mathcal{Z}^S$ of the perturbations contains the origin and is included in the single-parametric family of normal ranges:

$$\mathcal{Z}_\rho = \rho\mathcal{Z}, \ \rho > 0.$$

As a result, the GRC (4.1.11) of (4.1.9) becomes a member, corresponding to $\rho = 1$, of the single-parametric family of constraints

$$\begin{array}{c}\mathrm{dist}([P_0 + \sum\limits_{\ell=1}^{L} \zeta_\ell P_\ell]y - [p^0 + \sum\limits_{\ell=1}^{L} \zeta_\ell p^\ell], \mathbf{Q}) \leq \sum\limits_{s=1}^{S} \alpha_s \mathrm{dist}(\zeta^s, \mathcal{Z}^s|\mathcal{L}^s)\\ \forall \zeta \in \mathcal{Z}_+^\rho = \underbrace{\rho(\mathcal{Z}^1 \times ... \times \mathcal{Z}^S)}_{\mathcal{Z}_\rho} + \underbrace{\mathcal{L}^1 \times ... \times \mathcal{L}^S}_{\mathcal{L}}\end{array} \qquad (\mathrm{GRC}_\rho)$$

in variables y, α. We define the tightness factor of a safe tractable approximation of the GRC as follows (cf. Definition 3.3):

Definition 4.5 *Assume that we are given an approximation scheme that associates with* (GRC_ρ) *a finite system $\mathcal{S}_\rho$ of efficiently computable convex constraints on variables y, α and, perhaps, additional variables u, depending on $\rho > 0$ as a parameter. We say that this approximation scheme is a safe tractable approximation of the GRC tight, within tightness factor $\vartheta \geq 1$, if*

(i) *For every $\rho > 0$, $\mathcal{S}_\rho$ is a safe tractable approximation of* (GRC_ρ)*: whenever $(y, \alpha \geq 0)$ can be extended to a feasible solution of $\mathcal{S}_\rho$, (y, α) satisfies* (GRC_ρ)*;*

(ii) *Whenever $\rho > 0$ and $(y, \alpha \geq 0)$ are such that (y, α) cannot be extended to a feasible solution of $\mathcal{S}_\rho$, the pair $(y, \vartheta^{-1}\alpha)$ is not feasible for* $(\mathrm{GRC}_{\vartheta\rho})$.

4.2 Tractability of GRC in the Case of Linear Optimization

As in the case of the usual Robust Counterpart, the central question of computational tractability of the Globalized RC of an uncertain LO reduces to a similar question for the GRC (4.1.4) of a single uncertain linear constraint (4.1.1). The latter question is resolved to a large extent by the following observation:

Proposition 4.1 *A vector x satisfies the semi-infinite constraint* (4.1.4) *if and only if x satisfies the following pair of semi-infinite constraints:*

$$\begin{array}{ll}(a) & \left[a^0 + \sum_{\ell=1}^{L} \zeta_\ell a^\ell\right]^T x \le \left[b^0 + \sum_{\ell=1}^{L} \zeta_\ell b^\ell\right] \ \forall \zeta \in \mathcal{Z} \\ (b) & \left[\sum_{\ell=1}^{L} \Delta_\ell a^\ell\right]^T x \le \left[\sum_{\ell=1}^{L} \Delta_\ell b^\ell\right] + \alpha \ \forall \Delta \in \widetilde{\mathcal{Z}} \equiv \{\Delta \in \mathcal{L} : \|\Delta\| \le 1\}.\end{array} \tag{4.2.1}$$

Remark 4.1 Proposition 4.1 implies that the GRC of an uncertain linear inequality is *equivalent* to a pair of semi-infinite linear inequalities of the type arising in the usual RC. Consequently, we can invoke the representation results of section 1.3 to show that *under mild assumptions on the perturbation structure, the GRC* (4.1.4) *can be represented by a "short" system of explicit convex constraints.*

Proof of Proposition 4.1. Let x satisfy (4.1.4). Then x satisfies (4.2.1.a) due to $\mathrm{dist}(\zeta, \mathcal{Z}|\mathcal{L}) = 0$ for $\zeta \in \mathcal{Z}$. In order to demonstrate that x satisfies (4.2.1.b) as well, let $\bar\zeta \in \mathcal{Z}$ and $\Delta \in \mathcal{L}$ with $\|\Delta\| \le 1$. By (4.1.4) and since $\mathcal{L}$ is a cone, for every $t > 0$ we have $\zeta_t := \bar\zeta + t\Delta \in \mathcal{Z} + \mathcal{L}$ and $\mathrm{dist}(\zeta_t, \mathcal{Z}|\mathcal{L}) \le \|t\Delta\| \le t$; applying (4.1.4) to $\zeta = \zeta_t$, we therefore get

$$\left[a^0 + \sum_{\ell=1}^{L} \bar\zeta_\ell a^\ell\right]^T x + t\left[\sum_{\ell=1}^{L} \Delta_\ell a^\ell\right]^T x \le \left[b^0 + \sum_{\ell=1}^{L} \bar\zeta_\ell b^\ell\right] + t\left[\sum_{\ell=1}^{L} \Delta_\ell b^\ell\right] + \alpha t.$$

Dividing both sides in this inequality by t and passing to limit as $t \to \infty$, we see that the inequality in (4.2.1.b) is valid at our Δ. Since $\Delta \in \widetilde{\mathcal{Z}}$ is arbitrary, x satisfies (4.2.1.b), as claimed.

It remains to prove that if x satisfies (4.2.1), then x satisfies (4.1.4). Indeed, let x satisfy (4.2.1). Given $\zeta \in \mathcal{Z} + \mathcal{L}$ and taking into account that $\mathcal{Z}$ and $\mathcal{L}$ are closed, we can find $\bar\zeta \in \mathcal{Z}$ and $\Delta \in \mathcal{L}$ such that $\bar\zeta + \Delta = \zeta$ and $t := \mathrm{dist}(\zeta, \mathcal{Z}|\mathcal{L}) = \|\Delta\|$. Representing $\Delta = te$ with $e \in \mathcal{L}$, $\|e\| \le 1$, we have

$$\begin{array}{l}\left[a^0 + \sum_{\ell=1}^{L} \zeta_\ell a^\ell\right]^T x - \left[b^0 + \sum_{\ell=1}^{L} \zeta_\ell b^\ell\right] \\ = \underbrace{\left[a^0 + \sum_{\ell=1}^{L} \bar\zeta_\ell a^\ell\right]^T x - \left[b^0 + \sum_{\ell=1}^{L} \bar\zeta_\ell b^\ell\right]}_{\le 0 \text{ by } (4.2.1.a)} + \underbrace{\left[\sum_{\ell=1}^{L} \Delta_\ell a^\ell\right]^T x - \left[\sum_{\ell=1}^{L} \Delta_\ell b^\ell\right]}_{\substack{= t\left[\left[\sum_{\ell=1}^{L} e_\ell a^\ell\right]^T x - \left[\sum_{\ell=1}^{L} e_\ell b^\ell\right]\right] \\ \le t\alpha \text{ by } (4.2.1.b)}} \\ \le t\alpha = \alpha \mathrm{dist}(\zeta, \mathcal{Z}|\mathcal{L}).\end{array}$$

Since $\zeta \in \mathcal{Z} + \mathcal{L}$ is arbitrary, x satisfies (4.1.4). □

Example 4.1 Consider the following 3 perturbation structures $(\mathcal{Z}, \mathcal{L}, \|\cdot\|)$:
(a) $\mathcal{Z}$ is a box $\{\zeta : |\zeta_\ell| \le \sigma_\ell, 1 \le \ell \le L\}$, $\mathcal{L} = \mathbb{R}^L$ and $\|\cdot\| = \|\cdot\|_1$;
(b) $\mathcal{Z}$ is an ellipsoid $\{\zeta : \sum_{\ell=1}^{L} \zeta_\ell^2/\sigma_\ell^2 \le \Omega^2\}$, $\mathcal{L} = \mathbb{R}^L_+$ and $\|\cdot\| = \|\cdot\|_2$;
(c) $\mathcal{Z}$ is the intersection of a box and an ellipsoid: $\mathcal{Z} = \{\zeta : |\zeta_\ell| \le \sigma_\ell, 1 \le \ell \le L, \sum_{\ell=1}^{L} \zeta_\ell^2/\sigma_\ell^2 \le \Omega^2\}$, $\mathcal{L} = \mathbb{R}^L$, $\|\cdot\| = \|\cdot\|_\infty$.

In these cases the GRC of (4.1.1) is equivalent to the finite systems of explicit convex inequalities as follows:
Case (a):

$$\begin{array}{ll} (a) & [a^0]^T x + \sum_{\ell=1}^{L} \sigma_\ell |[a^\ell]^T x - b^\ell| \le b^0 \\ (b) & |[a^\ell]^T x - b^\ell| \le \alpha,\ \ell = 1, ..., L \end{array}$$

Here (a) represents the constraint (4.2.1.a) (cf. Example 1.4), and (b) represents the constraint (4.2.1.b) (why?)
Case (b):

$$\begin{array}{ll} (a) & [a^0]^T x + \Omega \left(\sum_{\ell=1}^{L} \sigma_\ell^2 ([a^\ell]^T x - b^\ell)^2 \right)^{1/2} \le b^0 \\ (b) & \left(\sum_{\ell=1}^{L} \max^2[[a^\ell]^T x - b^\ell, 0] \right)^{1/2} \le \alpha. \end{array}$$

Here (a) represents the constraint (4.2.1.a) (cf. Example 1.5), and (b) represents the constraint (4.2.1.b).
Case (c):

$$\begin{array}{ll} (a.1) & [a^0]^T x + \sum_{\ell=1}^{L} \sigma_\ell |z_\ell| + \Omega \left(\sum_{\ell=1}^{L} \sigma_\ell^2 w_\ell^2 \right)^{1/2} \le b^0 \\ (a.2) & z_\ell + w_\ell = [a^\ell]^T x - b^\ell,\ \ell = 1, ..., L \\ (b) & \sum_{\ell=1}^{L} |[a^\ell]^T x - b^\ell| \le \alpha. \end{array}$$

Here $(a.1$–$2)$ represent the constraint (4.2.1.a) (cf. Example 1.6), and (b) represents the constraint (4.2.1.b).

4.2.1 Illustration: Antenna Design

We are about to illustrate the GRC approach by applying it to the Antenna Design problem (Example 1.1), where we are interested in the uncertain LP problem of the form

$$\left\{ \min_{x,\tau} \{\tau : \|d - D(I + \mathrm{Diag}\{\zeta\})x\|_\infty \le \tau\} : \|\zeta\|_\infty \le \rho \right\}; \tag{4.2.2}$$

here D is a given $m \times L$ matrix and $d \in \mathbb{R}^m$ is a given vector. The RC of the problem is equivalent to

$$\mathrm{Opt}(\rho) = \min_x \left\{ F_x(\rho) := \max_{\|\zeta\|_\infty \le \rho} \|d - D(I + \mathrm{Diag}\{\zeta\})x\|_\infty \right\} \tag{RC$_\rho$}$$

For a candidate design x, the function $F_x(\rho)$ of the uncertainty level ρ has a very transparent interpretation: it is the worst-case, over perturbations ζ with $\|\zeta\|_\infty \le \rho$, loss (deviation of the synthesized diagram from the target one) of this design. This clearly is a convex and nondecreasing function of ρ.

Let us fix a "reference uncertainty level" $\bar\rho \ge 0$ and equip our uncertain problem with the perturbation structure

$$\mathcal{Z} = \{\zeta : \|\zeta\|_\infty \le \bar\rho\},\ \mathcal{L} = \mathbb{R}^L,\ \|\cdot\| = \|\cdot\|_\infty. \tag{4.2.3}$$

With this perturbation structure, it can be immediately derived from Proposition 4.1 (do it!) that a pair (τ, x) is a robust feasible solution to the GRC with global sensitivity α if and only if

$$\tau \ge F_x(\bar\rho)\ \&\ \alpha \ge \alpha(x) = \lim_{\rho\to\infty} \frac{d}{d\rho} F_\rho(x) = \max_{i \le m} \sum_{j=1}^{L} |D_{ij}||x_j|.$$

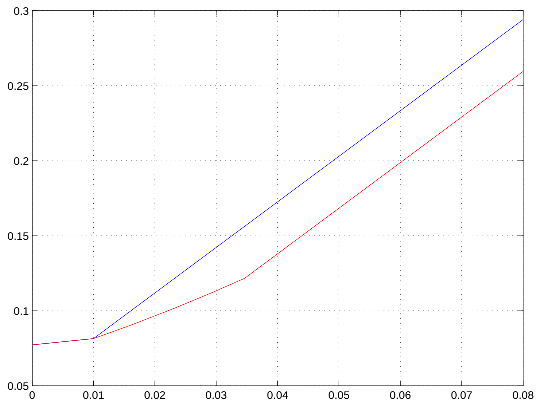

Figure 4.1: Bound (4.2.5) with $\rho_0 = 0.01$ (blue) on the loss $F_x(\rho)$ (red) associated with the optimal solution x to $(\mathrm{RC}_{0.01})$. The common values of the bound and the loss at $\rho_0 = 0.01$ is the optimal value of $(\mathrm{RC}_{0.01})$.

Note that the best (the smallest) value $\alpha(x)$ of global sensitivity which, for appropriately chosen τ, makes (τ, x) feasible for the GRC depends solely on x; we shall refer to this quantity as to the *global sensitivity of* x. Due to its origin, and since $F_x(\rho)$ is convex and nondecreasing, $\alpha(x)$ and a value of $F_x(\cdot)$ at a particular $\rho = \rho_0 \geq 0$ imply a piecewise linear upper bound on $F_x(\cdot)$:

$$\forall \rho \geq 0 : F_x(\rho) \leq \left\{ \begin{array}{ll} F_x(\rho_0), & \rho < \rho_0 \\ F_x(\rho_0) + \alpha(x)[\rho - \rho_0], & \rho \geq \rho_0 \end{array} \right. . \tag{4.2.4}$$

Taking into account also the value of F_x at 0, we can improve this bound to

$$\forall \rho \geq 0 : F_x(\rho) \leq \left\{ \begin{array}{ll} \frac{\rho_0 - \rho}{\rho_0} F_x(0) + \frac{\rho}{\rho_0} F_x(\rho_0), & \rho < \rho_0 \\ F_x(\rho_0) + \alpha(x)[\rho - \rho_0], & \rho \geq \rho_0 \end{array} \right. . \tag{4.2.5}$$

In figure 4.1, we plot the latter bound and the true $F_x(\cdot)$ for the robust design built in section 1.4.1 (that is, the optimal solution to $(\mathrm{RC}_{0.01})$), choosing $\rho_0 = 0.01$. When designing a robust antenna, our "ideal goal" would be to choose the design x which makes the loss $F_x(\rho)$ as small as possible *for all values of* ρ; of course, this goal usually cannot be achieved. With the usual RC approach, we fix the uncertainty level at $\bar{\rho}$ and minimize over x the loss at this particular value of ρ, with no care of how rapidly this loss grows when the true uncertainty level ρ exceeds our guess $\bar{\rho}$. This makes sense when we understand well what is the uncertainty level at which our system should work, which sometimes is not the case. With the GRC approach, we can, to some extent, take care of both the value of F_x at $\rho = \bar{\rho}$ and of the rate at which the loss grows with ρ, thus making our design better suited to the situations when it should be used in a wide range of uncertainty levels. For example, we can act as follows:

- We first solve $(\mathrm{RC}_{\bar{\rho}})$, thus getting the "reference" design $\bar{x}$ with the loss at the uncertainty level $\bar{\rho}$ as small as possible, so that $F_{\bar{x}}(\bar{\rho}) = \mathrm{Opt}(\mathrm{RC}_{\bar{\rho}})$;
- We then increase the resulting loss by certain percentage δ (say, $\delta = 0.1$) and choose, as

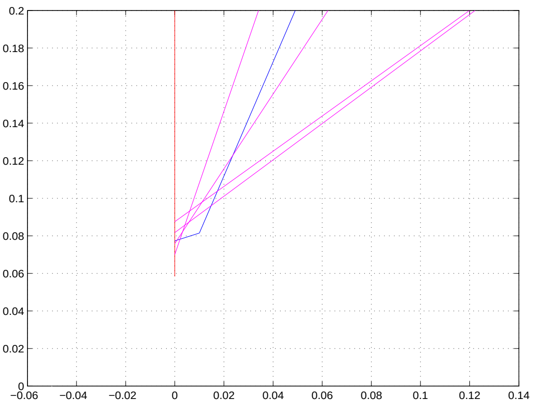

Figure 4.2: Red and magenta: bounds (4.2.4) on the losses for optimal solutions to (4.2.6) for the values of δ listed in table 4.1; the bounds correspond to $\rho_0 = 0$. Blue: bound (4.2.5) with $\rho_0 = 0.01$ on the loss $F_x(\rho)$ associated with the optimal solution x to ($\mathrm{RC}_{0.01}$).

δ	0	0.2	0.3	0.4	0.5
$\beta_*(\delta)$	2.6×10^5	3.8134	1.9916	0.9681	0.9379

Table 4.1: Tolerances δ and quantities $\beta_*(\delta)$ (the global sensitivities of the optimal solutions to (4.2.6)). Pay attention to how huge is the global sensitivity $\beta_*(0)$ of the nominal optimal design. For comparison: the global sensitivity of the robust design built in section 1.4.1 is just 3.0379.

the actual design, the solution to the minimization problem

$$\min_x \left\{\alpha(x) : F_x(\bar{\rho}) \leq (1+\delta)\mathrm{Opt}(\mathrm{RC}_{\bar{\rho}})\right\}.$$

In other words, we allow for a controlled sacrifice in the loss at the "nominal" uncertainty level $\bar{\rho}$ in order to get as good as possible upper bound (4.2.4) on the loss in the range $\rho \geq \bar{\rho}$.

In figure 4.2 and in table 4.1, we illustrate the latter approach in the special case when $\bar{\rho} = 0$. In this case, we want from our design to perform nearly as well as (namely, within percentage δ of) the nominally optimal design in the ideal case of no actuation errors, and optimize under this restriction the global sensitivity of the design w.r.t. the magnitude of actuation errors. Mathematically, this reduces to solving a simple LP problem

$$\beta_*(\delta) = \min_x \left\{ \max_{i \leq m} \sum_{j=1}^{L} |D_{ij}||x_j| : \|d - Dx\|_\infty \leq (1+\delta) \min_u \|d - Du\|_\infty \right\}. \tag{4.2.6}$$

4.3 Tractability of GRC in the Case of Conic Optimization

4.3.1 Decomposition

Preliminaries

Recall the notion of the *recessive cone* of a closed and nonempty convex set $\mathbf{Q}$:

Definition 4.6 *Let $\mathbf{Q} \subset \mathbb{R}^k$ be a nonempty closed convex set and $\bar{x} \in \mathbf{Q}$. The recessive cone* $\mathrm{Rec}(\mathbf{Q})$ *of Q is comprised of all rays emanating from $\bar{x}$ and contained in $\mathbf{Q}$:*

$$\mathrm{Rec}(\mathbf{Q}) = \{h \in \mathbb{R}^k : \bar{x} + th \in \mathbf{Q} \forall t \geq 0\}.$$

(Due to closedness and convexity of $\mathbf{Q}$, the right hand side set in this formula is independent of the choice of $\bar{x} \in \mathbf{Q}$ and is a nonempty closed convex cone in $\mathbb{R}^k$.)

Example 4.2
(i) The recessive cone of a nonempty bounded and closed convex set $\mathbf{Q}$ is trivial: $\mathrm{Rec}(\mathbf{Q}) = \{0\}$;
(ii) The recessive cone of a closed convex cone $\mathbf{Q}$ is $\mathbf{Q}$ itself;
(iii) The recessive cone of the set $\mathbf{Q} = \{x : Ax - b \in \mathbf{K}\}$, where $\mathbf{K}$ is a closed convex cone, is the set $\{h : Ah \in \mathbf{K}\}$;
(iv.a) Let $\mathbf{Q}$ be a closed convex set and $e_i \to e$, $i \to \infty$, $t_i \geq 0$, $t_i \to \infty$, $i \to \infty$, be sequences of vectors and reals such that $t_i e_i \in \mathbf{Q}$ for all i. Then $e \in \mathrm{Rec}(\mathbf{Q})$.
(iv.b) Vice versa: every $e \in \mathrm{Rec}(\mathbf{Q})$ can be represented in the form of $e = \lim_{i\to\infty} e_i$ with vectors e_i such that $ie_i \in \mathbf{Q}$.
Proof. (iv.a): Let $\bar{x} \in \mathbf{Q}$. With our e_i and t_i, for every $t > 0$ we have $\bar{x} + te_i - t/t_i\bar{x} = (t/t_i)(t_ie_i) + (1 - t/t_i)\bar{x}$. For all but finitely many values of i, the right hand side in this equality is a convex combination of two vectors from $\mathbf{Q}$ and therefore belongs to $\mathbf{Q}$; for $i \to \infty$, the left hand side converges to $\bar{x} + te$. Since $\mathbf{Q}$ is closed, we conclude that $\bar{x} + te \in \mathbf{Q}$; since $t > 0$ is arbitrary, we get $e \in \mathrm{Rec}(\mathbf{Q})$.
(iv.b): Let $e \in \mathrm{Rec}(\mathbf{Q})$ and $\bar{x} \in \mathbf{Q}$. Setting $e_i = i^{-1}(\bar{x} + ie)$, we have $ie_i \in \mathbf{Q}$ and $e_i \to e$ as $i \to \infty$. □

The Main Result

The following statement is the "multi-dimensional" extension of Proposition 4.1:

Proposition 4.2 *A candidate solution y is feasible for the GRC* (4.1.11) *of the uncertain constraint* (4.1.9) *if and only if x satisfies the following system of semi-infinite constraints:*

$$\begin{array}{ll} (a) & \overbrace{[P_0 + \sum_{\ell=1}^{L} \zeta_\ell P_\ell]^T y - [p^0 + \sum_{\ell=1}^{L} \zeta_\ell p^\ell]}^{P(y,\zeta)} \in \mathbf{Q} \\ & \qquad \forall \zeta \in \mathcal{Z} \equiv \mathcal{Z}^1 \times ... \times \mathcal{Z}^S \\ (b_s) & \mathrm{dist}(\overbrace{\sum_{\ell=1}^{L} [P_\ell y - p^\ell](E_s\zeta^s)_\ell}^{\Phi(y)E_s\zeta^s}, \mathrm{Rec}(\mathbf{Q})) \leq \alpha_s \\ & \qquad \forall \zeta^s \in \mathcal{L}^s_{\|\cdot\|_s} \equiv \{\zeta^s \in \mathcal{L}^s : \|\zeta^s\|_s \leq 1\}, \quad s = 1, ..., S, \end{array} \tag{4.3.1}$$

where E_s is the natural embedding of $\mathbb{R}^{L_s}$ into $\mathbb{R}^L = \mathbb{R}^{L_1} \times ... \times \mathbb{R}^{L_S}$ and $\mathrm{dist}(u, \mathrm{Rec}(\mathbf{Q})) = \min\limits_{v \in \mathrm{Rec}(\mathbf{Q})} \|u - v\|_Q$.

Proof. Assume that y satisfies (4.1.11), and let us verify that y satisfies (4.3.1). Relation (4.3.1.a) is evident. Let us fix $s \leq S$ and verify that y satisfies (4.3.1.b_s). Indeed, let $\bar{\zeta} \in \mathcal{Z}$ and $\zeta^s \in \mathcal{L}^s_{\|\cdot\|_s}$. For $i = 1, 2, ...,$ let ζ_i be given by $\zeta_i^r = \bar{\zeta}^r$, $r \neq s$, and $\zeta_i^s = \bar{\zeta}^s + i\zeta^s$, so that $\mathrm{dist}(\zeta_i^r, \mathcal{Z}^r|\mathcal{L}^r)$ is 0 for $r \neq s$ and is $\leq i$ for $r = s$. Since y is feasible for (4.1.11), we have

$$\mathrm{dist}(\underbrace{[P_0 + \sum_{\ell=1}^{L} (\zeta_i)_\ell P_\ell]y - [p^0 + \sum_{\ell=1}^{L} (\zeta_i)_\ell p^\ell]}_{P(y,\zeta_i)=P(y,\bar{\zeta})+i\Phi(y)E_s\zeta^s}, \mathbf{Q}) \leq \alpha_s i,$$

that is, there exists $q_i \in \mathbf{Q}$ such that

$$\|P(y, \bar{\zeta}) + i\Phi(y)E_s\zeta^s - q_i\|_Q \leq \alpha_s i.$$

From this inequality it follows that $\|q_i\|_Q/i$ remains bounded when $i \to \infty$; setting $q_i = ie_i$ and passing to a subsequence $\{i_\nu\}$ of indices i, we may assume that $e_{i_\nu} \to e$ as $\nu \to \infty$; by item (iv.a) of Example 4.2, we have $e \in \mathrm{Rec}(\mathbf{Q})$. We further have

$$\begin{array}{rcl} \|\Phi(y)E_s\zeta^s - e_{i_\nu}\|_Q & = & i_\nu^{-1}\|i_\nu\Phi(y)E_s\zeta - q_{i_\nu}\|_Q \\ & \leq & i_\nu^{-1}\left[\|P(y, \bar{\zeta}) + i_\nu\Phi(y)E_s\zeta^s - q_{i_\nu}\|_Q + i_\nu^{-1}\|P(y, \bar{\zeta})\|_Q\right] \\ & \leq & \alpha_s + i_\nu^{-1}\|P(y, \bar{\zeta})\|_Q, \end{array}$$

whence, passing to limit as $\nu \to \infty$, $\|\Phi(y)E_s\zeta^s - e\|_Q \leq \alpha_s$, whence, due to $e \in \mathrm{Rec}(\mathbf{Q})$, we have $\mathrm{dist}(\Phi(y)E_s\zeta^s, \mathrm{Rec}(\mathbf{Q})) \leq \alpha_s$. Since $\zeta^s \in \mathcal{L}^s_{\|\cdot\|_s}$ is arbitrary, (4.3.1.b_s) holds true.

Now assume that y satisfies (4.3.1), and let us prove that y satisfies (4.1.11). Indeed, given $\zeta \in \mathcal{Z} + \mathcal{L}$, we can find $\bar{\zeta}^s \in \mathcal{Z}^s$ and $\delta^s \in \mathcal{L}^s$ in such a way that $\zeta^s = \bar{\zeta}^s + \delta^s$ and $\|\delta^s\|_s = \mathrm{dist}(\zeta^s, \mathcal{Z}^s|\mathcal{L}^s)$. Setting $\bar{\zeta} = (\bar{\zeta}^1, ..., \bar{\zeta}^S)$ and invoking (4.3.1.a), the vector $\bar{u} = P(y, \bar{\zeta})$ belongs to $\mathbf{Q}$. Further, for every s, by (4.3.1.b_s), there exists $\delta u^s \in \mathrm{Rec}(\mathbf{Q})$ such that $\|\Phi(y)E_s\delta^s - \delta u^s\|_Q \leq \alpha_s\|\delta^s\|_s = \alpha_s \mathrm{dist}(\zeta^s, \mathcal{Z}^s|\mathcal{L}^s)$. Since $P(y, \zeta) = P(y, \bar{\zeta}) + \sum_s \Phi(y)E_s\delta^s$, we have

$$\|P(y, \zeta) - \underbrace{[\bar{u} + \sum_s \delta u^s]}_{v}\|_Q \leq \|\underbrace{P(y, \bar{\zeta}) - \bar{u}}_{=0}\|_Q + \sum_s \underbrace{\|\Phi(y)E_s\delta^s - \delta u^s\|_Q}_{\leq \alpha_s \mathrm{dist}(\zeta^s, \mathcal{Z}^s|\mathcal{L}^s)};$$

since $\bar{u} \in \mathbf{Q}$ and $\delta u^s \in \mathrm{Rec}(\mathbf{Q})$ for all s, we have $v \in \mathbf{Q}$, so that the inequality implies that

$$\mathrm{dist}(P(y, \zeta), \mathbf{Q}) \leq \sum_s \alpha_s \mathrm{dist}(\zeta^s, \mathcal{Z}^s|\mathcal{L}^s).$$

Since $\zeta \in \mathcal{Z} + \mathcal{L}$ is arbitrary, y satisfies (4.1.11). □

Consequences of Main Result

Proposition 4.2 demonstrates that the GRC of an uncertain constraint (4.1.9) is equivalent to the explicit system of semi-infinite constraints (4.3.1). We are well acquainted with the constraint (4.3.1.a) — it is nothing but the RC of the uncertain constraint (4.1.9) with the normal range $\mathcal{Z}$ of the perturbations in the role of the uncertainty set. As a result, we have certain knowledge of how to convert this semi-infinite constraint into a tractable form or how to build its tractable safe approximation. What is new is the constraint (4.3.1.b), which is of the following generic form:

We are given

• an Euclidean space E with inner product $\langle\cdot,\cdot\rangle_E$, a norm (not necessarily the Euclidean one) $\|\cdot\|_E$, and a closed convex cone K^E in E;

• an Euclidean space F with inner product $\langle\cdot,\cdot\rangle_F$, norm $\|\cdot\|_F$ and a closed convex cone K^F in F.

These data define a function on the space $\mathcal{L}(E,F)$ of linear mappings $\mathcal{M}$ from E to F, specifically, the function

$$\begin{array}{c}\Psi(\mathcal{M}) = \max\limits_{e}\left\{\mathrm{dist}_{\|\cdot\|_F}(\mathcal{M}e, K^F) : e\in K^E, \|e\|_E\leq 1\right\},\\ \mathrm{dist}_{\|\cdot\|_F}(f, K^F) = \min\limits_{g\in K^F}\|f-g\|_F.\end{array}\tag{4.3.2}$$

Note that $\Psi(\mathcal{M})$ is a kind of a norm: it is nonnegative, satisfies the requirement $\Psi(\lambda\mathcal{M}) = \lambda\Psi(\mathcal{M})$ when $\lambda\geq 0$, and satisfies the triangle inequality $\Psi(\mathcal{M}+\mathcal{N})\leq\Psi(\mathcal{M})+\Psi(\mathcal{N})$. The properties of a norm that are missing are symmetry (in general, $\Psi(-\mathcal{M})\neq\Psi(\mathcal{M})$) and strict positivity (it may happen that $\Psi(\mathcal{M}) = 0$ for $\mathcal{M}\neq 0$). Note also that in the case when $K^F=\{0\}$, $K^E=E$, $\Psi(\mathcal{M}) = \max\limits_{e:\|e\|_E\leq 1}\|\mathcal{M}e\|_F$ becomes the usual norm of a linear mapping induced by given norms in the origin and the destination spaces.

The above setting gives rise to a convex inequality

$$\Psi(\mathcal{M})\leq\alpha \tag{4.3.3}$$

in variables $\mathcal{M},\alpha$. Note that every one of the constraints (4.3.1.b) is obtained from a convex inequality of the form (4.3.3) by affine substitution

$$\mathcal{M}\leftarrow H_s(y),\ \alpha\leftarrow\alpha_s$$

where $H_s(y)\in\mathcal{L}(E_s,F_s)$ is affine in y. Indeed, (4.3.1.b_s) is obtained in this fashion when specifying

- $(E,\langle\cdot,\cdot\rangle_E)$ as the Euclidean space where $\mathcal{Z}^s,\mathcal{L}^s$ live, and $\|\cdot\|_E$ as $\|\cdot\|_s$;
- $(F,\langle\cdot,\cdot\rangle_F)$ as the Euclidean space where $\mathbf{Q}$ lives, and $\|\cdot\|_F$ as $\|\cdot\|_Q$;
- K^E as the cone $\mathcal{L}^s$, and K^F as the cone $\mathrm{Rec}(\mathbf{Q})$;
- $H(y)$ as the linear map $\zeta^s\mapsto\Phi(y)E_s\zeta^s$.

It follows that efficient processing of constraints (4.3.1.b) reduces to a similar task for the associated constraints

$$\Psi_s(\mathcal{M}_s)\leq\alpha_s \tag{$\mathcal{C}_s$}$$

of the form (4.3.3). Assume, e.g., that we are smart enough to build, for certain $\vartheta\geq 1$,

(i) a ϑ-tight safe tractable approximation of the semi-infinite constraint (4.3.1.a) with $\mathcal{Z}_\rho=\rho\mathcal{Z}_1$ in the role of the perturbation set. Let this approximation be a system $\mathcal{S}^a_\rho$ of explicit convex constraints in variables y and additional variables u;

(ii) for every $s=1,...,S$ a ϑ-tight efficiently computable upper bound on the function $\Psi_s(\mathcal{M}_s)$, that is, a system $\mathcal{S}^s$ of efficiently computable convex constraints on matrix variable $\mathcal{M}_s$, real variable τ_s and, perhaps, additional variables u^s such that

(a) whenever $(\mathcal{M}_s,\tau_s)$ can be extended to a feasible solution of $\mathcal{S}^s$, we have $\Psi_s(\mathcal{M}_s)\leq\tau_s$,

(b) whenever $(\mathcal{M}_s, \tau_s)$ cannot be extended to a feasible solution of $\mathcal{S}^s$, we have $\vartheta\Psi_s(\mathcal{M}_s) > \tau_s$.

In this situation, we can point out a safe tractable approximation, tight within the factor ϑ (see Definition 4.5), of the GRC in question. To this end, consider the system of constraints in variables $y, \alpha_1, ..., \alpha_S, u, u^1, ..., u^S$ as follows:

$$(y, u) \text{ satisfies } \mathcal{S}^a_\rho \text{ and } \{(H_s(y), \alpha_s, u^s) \text{ satisfies } \mathcal{S}^s,\ s = 1, ..., S\}, \qquad (\mathcal{S}_\rho)$$

and let us verify that this is a ϑ-tight safe computationally tractable approximation of the GRC. Indeed, $\mathcal{S}_\rho$ is an explicit system of efficiently computable convex constraints and as such is computationally tractable. Further, $\mathcal{S}_\rho$ is a safe approximation of the (GRC_ρ). Indeed, if (y, α) can be extended to a feasible solution of $\mathcal{S}_\rho$, then y satisfies $(4.3.1.a)$ with $\mathcal{Z}_\rho$ in the role of $\mathcal{Z}$ (since (y, u) satisfies $\mathcal{S}^a_\rho$) and (y, α_s) satisfies $(4.3.1.b_s)$ due to (ii.a) (recall that $(4.3.1.b_s)$ is equivalent to $\Psi_s(H_s(y)) \le \alpha_s$). Finally, assume that (y, α) cannot be extended to a feasible solution of $(\mathcal{S}_\rho)$, and let us prove that then $(y, \vartheta^{-1}\alpha)$ is not feasible for $(\mathrm{GRC}_{\vartheta\rho})$. Indeed, if (y, α) cannot be extended to a feasible solution to $\mathcal{S}_\rho$, then either y cannot be extended to a feasible solution of $\mathcal{S}^a_\rho$, or for certain s (y, α_s) cannot be extended to a feasible solution of $\mathcal{S}^s$. In the first case, y does not satisfy $(4.3.1.a)$ with $\mathcal{Z}_{\vartheta\rho}$ in the role of $\mathcal{Z}$ by (i); in the second case, $\vartheta^{-1}\alpha_s < \Psi_s(H_s(y))$ by (ii.b), so that in both cases the pair $(y, \vartheta^{-1}\alpha)$ is not feasible for $(\mathrm{GRC}_{\vartheta\rho})$.

We have reduced the tractability issues related to Globalized RCs to similar issues for RCs (which we have already investigated in the CO case) and to the issue of efficient bounding of $\Psi(\cdot)$. The rest of this section is devoted to investigating this latter issue.

4.3.2 Efficient Bounding of $\Psi(\cdot)$

Symmetry

We start with observing that the problem of efficient computation of (a tight upper bound on) $\Psi(\cdot)$ possesses a kind of symmetry. Indeed, consider a setup

$$\Xi = (E, \langle\cdot,\cdot\rangle_E, \|\cdot\|_E, K^E; F, \langle\cdot,\cdot\rangle_F, \|\cdot\|_F, K^F)$$

specifying Ψ, and let us associate with Ξ its *dual* setup

$$\Xi_* = (F, \langle\cdot,\cdot\rangle_F, \|\cdot\|^*_F, K^F_*; E, \langle\cdot,\cdot\rangle_E, \|\cdot\|^*_E, K^E_*),$$

where

- for a norm $\|\cdot\|$ on a Euclidean space $(G, \langle\cdot,\cdot\rangle_G)$, its *conjugate* norm $\|\cdot\|^*$ is defined as

$$\|u\|^* = \max_v \{\langle u, v\rangle_G : \|v\| \le 1\};$$

- For a closed convex cone K in a Euclidean space $(G, \langle\cdot,\cdot\rangle_G)$, its *dual* cone is defined as

$$K_* = \{y : \langle y, h\rangle_G \ge 0 \ \ \forall h \in K\}.$$

.

Recall that the *conjugate* to a linear map $\mathcal{M} \in \mathcal{L}(E, F)$ from Euclidean space E to Euclidean space F is the linear map $\mathcal{M}^* \in \mathcal{L}(F, E)$ uniquely defined by the identity

$$\langle \mathcal{M}e, f\rangle_F = \langle e, \mathcal{M}^* f\rangle_E \quad \forall (e \in E, f \in F);$$

representing linear maps by their matrices in a fixed pair of orthonormal bases in E, F, the matrix representing $\mathcal{M}^*$ is the transpose of the matrix representing $\mathcal{M}$. Note that twice taken dual/conjugate of an entity recovers the original entity: $(K_*)_* = K$, $(\|\cdot\|^*)^* = \|\cdot\|$, $(\mathcal{M}^*)^* = \mathcal{M}$, $(\Xi_*)_* = \Xi$.

Recall that the functions $\Psi(\cdot)$ are given by setups Ξ of the outlined type according to

$$\Psi(\mathcal{M}) \equiv \Psi_\Xi(\mathcal{M}) = \max_{e \in E} \left\{ \mathrm{dist}_{\|\cdot\|_F}(\mathcal{M}e, K^F) : e \in K^E, \|e\|_E \le 1 \right\}.$$

The aforementioned symmetry is nothing but the following simple statement:

Proposition 4.3 *For every setup $\Xi = (E, ..., K^F)$ and every $\mathcal{M} \in \mathcal{L}(E, F)$ one has*

$$\Psi_\Xi(\mathcal{M}) = \Psi_{\Xi_*}(\mathcal{M}^*).$$

Proof. Let $H, \langle\cdot,\cdot\rangle_H$ be a Euclidean space. Recall that the *polar* of a closed convex set $X \subset H$, $0 \in X$, is the set $X^o = \{y \in H : \langle y, x\rangle_H \le 1 \quad \forall x \in X\}$. We need the following facts:

(a) If $X \subset H$ is closed, convex and $0 \in X$, then so is X^o, and $(X^o)^o = X$ [87];

(b) If $X \subset H$ is convex compact, $0 \in X$, and $K^H \subset H$ is closed convex cone, then $X + K^H$ is closed and

$$(X + K^H)^o = X^o \cap (-K^H_*).$$

Indeed, the arithmetic sum of a compact and a closed set is closed, so that $X + K^H$ is closed, convex, and contains 0. We have

$$f \in (X + K^H)^o \Leftrightarrow 1 \ge \sup_{x \in X, h \in K^H} \langle f, x + h\rangle_H = \sup_{x \in X} \langle f, x\rangle_H + \sup_{h \in K^H} \langle f, h\rangle_H;$$

since K^H is a cone, the concluding inequality is possible iff $f \in X^o$ and $f \in -K^H_*$.

(c) Let $\|\cdot\|$ be a norm in H. Then for every $\alpha > 0$ one has $(\{x : \|x\| \le \alpha\})^o = \{x : \|x\|^* \le 1/\alpha\}$ (evident).

When $\alpha > 0$, we have

$$\begin{array}{ll}
 & \Psi_\Xi(\mathcal{M}) \le \alpha \\
\Leftrightarrow & \left\{ \begin{array}{l} \forall e \in K^E \cap \{e : \|e\|_E \le 1\} : \\ \mathcal{M}e \in \{f : \|f\|_F \le \alpha\} + K^F \end{array} \right. \qquad \text{[by definition]} \\
\Leftrightarrow & \left\{ \begin{array}{l} \forall e \in K^E \cap \{e : \|e\|_E \le 1\} : \\ \mathcal{M}e \in \left[[\{f : \|f\|_F \le \alpha\} + K^F]^o\right]^o \end{array} \right. \qquad \text{[by (a)]} \\
\Leftrightarrow & \left\{ \begin{array}{l} \forall e \in K^E \cap \{e : \|e\|_E \le 1\} : \\ \langle \mathcal{M}e, f\rangle_F \le 1 \; \forall f \in \underbrace{[\{f : \|f\|_F \le \alpha\} + K^F]^o}_{=\{f : \|f\|^*_F \le \alpha^{-1}\} \cap (-K^F_*)} \end{array} \right. \qquad \text{[by (b), (c)]}
\end{array}$$

$$
\begin{array}{lll}
\Leftrightarrow & \left\{\begin{array}{l}\forall e\in K^E\cap\{e:\|e\|_E\leq 1\}:\\ \langle e,\mathcal{M}^*f\rangle_E\leq 1\ \forall f\in\{f:\|f\|_F^*\leq\alpha^{-1}\}\cap(-K^F_*)\end{array}\right. & \\
\Leftrightarrow & \left\{\begin{array}{l}\forall e\in K^E\cap\{e:\|e\|_E\leq \alpha^{-1}\}:\\ \langle e,\mathcal{M}^*f\rangle_E\leq 1\ \forall f\in\{f:\|f\|_F^*\leq 1\}\cap(-K^F_*)\end{array}\right. & \text{[evident]}\\
\Leftrightarrow & \left\{\begin{array}{l}\forall e\in [-(-K^E_*)_*]\cap[\{e:\|e\|_E^*\leq \alpha\}^o]:\\ \langle e,\mathcal{M}^*f\rangle_E\leq 1\ \forall f\in\{f:\|f\|_F^*\leq 1\}\cap(-K^F_*)\end{array}\right. & \text{[by (c)]}\\
\Leftrightarrow & \left\{\begin{array}{l}\forall e\in \left[(-K^E_*)+\{e:\|e\|_E^*\leq \alpha\}\right]^o:\\ \langle \mathcal{M}^*f,e\rangle_E\leq 1\ \forall f\in\{f:\|f\|_F^*\leq 1\}\cap(-K^F_*)\end{array}\right. & \text{[by (b)]}\\
\Leftrightarrow & \left\{\begin{array}{l}\forall f\in\{f:\|f\|_F^*\leq 1\}\cap(-K^F_*):\\ \langle \mathcal{M}^*f,e\rangle_E\leq 1\,\forall e\in [(-K^E_*)+\{e:\|e\|_E^*\leq \alpha\}]^o\end{array}\right. & \\
\Leftrightarrow & \left\{\begin{array}{l}\forall f\in\{f:\|f\|_F^*\leq 1\}\cap(-K^F_*):\\ \mathcal{M}^*f\in(-K^E_*)+\{e:\|e\|_E^*\leq \alpha\}\end{array}\right. & \text{[by (a)]}\\
\Leftrightarrow & \left\{\begin{array}{l}\forall f\in K^*_F\cap\{f:\|f\|_F^*\leq 1\}:\\ \mathcal{M}^*f\in K^E_*+\{e:\|e\|_E^*\leq \alpha\}\end{array}\right. & \\
\Leftrightarrow & \Psi_{\Xi_*}(\mathcal{M}^*)\leq\alpha. & \square
\end{array}
$$

Good GRC setups

Proposition 4.3 says that "good" setups Ξ — those for which $\Psi_\Xi(\cdot)$ is efficiently computable or admits a tight, within certain factor ϑ, efficiently computable upper bound — always come in symmetric pairs: if Ξ is good, so is Ξ_*, and vice versa. In what follows, we refer to members of such a symmetric pair as to *counterparts* of each other. We are about to list a number of good pairs. From now on, we assume that all components of a setup in question are "computationally tractable," specifically, that the cones K^E, K^F and the epigraphs of the norms $\|\cdot\|_E$, $\|\cdot\|_F$ are given by LMI representations (or, more general, by systems of efficiently computable convex constraints). Below, we denote by B_E and B_F the unit balls of the norms $\|\cdot\|_E$, $\|\cdot\|_F$, respectively. Here are several good GRC setups:

A: $K^E=\{0\}$. The counterpart is

A*: $K^F=F$.

These cases are trivial: $\Psi_\Xi(\mathcal{M})\equiv 0$.

B: $K^E=E$, $B_E=\mathrm{Conv}\{e^1,...,e^N\}$, *the list* $\{e^i\}_{i=1}^N$ *is available.* The counterpart is the case

B*: $K^F=\{0\}$, $B_F=\{f:\langle f^i,f\rangle_F\leq 1, i=1,...,N\}$, *the list* $\{f^i\}_{i=1}^N$ *is available.*

Standard example for **B** is $E=\mathbb{R}^n$ with the standard inner product, $K^E=E$, $\|e\|=\|e\|_1\equiv\sum\limits_j|e_j|$. Standard example for **B*** is $F=\mathbb{R}^m$ with the standard inner product, $\|f\|_F=\|f\|_\infty=\max\limits_j|f_j|$.

The cases in question are easy. Indeed, in the case of **B** we clearly have

$$\Psi(\mathcal{M})=\max_{1\leq j\leq N}\mathrm{dist}_{\|\cdot\|_F}(\mathcal{M}e_i,K^F),$$

and thus $\Psi(\mathcal{M})$ is efficiently computable (as the maximum of a finite family of efficiently computable quantities $\mathrm{dist}_{\|\cdot\|_F}(\mathcal{M}e_i,K^F)$). Assuming, e.g., that E, F are, respectively, $\mathbb{R}^m$ and $\mathbb{R}^n$ with the standard inner products, and that K^F, $\|\cdot\|_F$ are given by strictly feasible conic representations:

$$
\begin{array}{l}
K^F=\{f:\exists u: Pf+Qu\in\mathbf{K}^1\},\\
\{t\geq\|f\|_F\}\Leftrightarrow\left\{\exists v: Rf+tr+Sv\in\mathbf{K}^2\right\}
\end{array}
$$

the relation

$$\Psi(\mathcal{M})\leq\alpha$$

can be represented equivalently by the following explicit system of conic constraints

$$\begin{array}{ll}(a) & Pf^i + Qu^i \in \mathbf{K}^1,\ i = 1, ..., N\\ (b) & R(\mathcal{M}e^i - f^i) + \alpha r + Sv^i \in \mathbf{K}^2,\ i = 1, ..., N\end{array}$$

in variables $\mathcal{M}, \alpha, u^i, f^i, v^i$. Indeed, relations (a) equivalently express the requirement $f^i \in K^F$, while relations (b) say that $\|\mathcal{M}e^i - f^i\|_F \leq \alpha$.

C: $K^E = E$, $K^F = \{0\}$. The counterpart case is exactly the same.

In the case of **C**, $\Psi(\cdot)$ is the norm of a linear map from E to F induced by given norms on the origin and the destination spaces:

$$\Psi(\mathcal{M}) = \max_e \{\|\mathcal{M}e\|_F : \|e\|_E \leq 1\}.$$

Aside of situations covered by **B**, **B***, there is only one generic situation where computing the norm of a linear map is easy — this is the situation where both $\|\cdot\|_E$ and $\|\cdot\|_F$ are Euclidean norms. In this case, we lose nothing by assuming that $E = \ell_2^n$ (that is, E is $\mathbb{R}^n$ with the standard inner product and the standard norm $\|e\|_2 = \sqrt{\sum_i e_i^2}$), $F = \ell_2^m$, and let M be the $m \times n$ matrix representing the map $\mathcal{M}$ in the standard bases of E and F. In this case, $\Psi(\mathcal{M}) = \|M\|_{2,2}$ is the maximal singular value of M and as such is efficiently computable. A semidefinite representation of the constraint $\|M\|_{2,2} \leq \alpha$ is

$$\left[\begin{array}{c|c}\alpha I_n & M^T\\ \hline M & \alpha I_m\end{array}\right] \succeq 0.$$

Now consider the case when $E = \ell_p^n$ (that is, E is $\mathbb{R}^n$ with the standard inner product and the norm

$$\|e\|_p = \left\{\begin{array}{ll}\left(\sum_j |e_j|^p\right)^{1/p} & , 1 \leq p < \infty\\ \max_j |e_j| & , p = \infty\end{array}\right.,$$

and $F = \ell_r^m$, $1 \leq r, p \leq \infty$. Here again we can naturally identify $\mathcal{L}(E, F)$ with the space $\mathbb{R}^{m\times n}$ of real $m \times n$ matrices, and the problem of interest is to compute

$$\|M\|_{p,r} = \max_e \{\|Me\|_r : \|e\|_p \leq 1\}.$$

The case of $p = r = 2$ is the just considered "purely Euclidean" situation; the cases of $p = 1$ and of $r = \infty$ are covered by **B**, **B***. These are the only 3 cases when computing $\|\cdot\|_{p,r}$ is known to be easy. It is also known that it is NP-hard to compute the matrix norm in question when $p > r$. However, in the case of $p \geq 2 \geq r$ there exists a tight efficiently computable upper bound on $\|M\|_{p,r}$ due to Nesterov [97, Theorem 13.2.4]. Specifically, Nesterov shows that when $\infty \geq p \geq 2 \geq r \geq 1$, the explicitly computable quantity

$$\Psi_{p,r}(M) = \frac{1}{2} \min_{\substack{\mu\in\mathbb{R}^n\\ \nu\in\mathbb{R}^m}} \left\{\|\mu\|_{\frac{p}{p-2}} + \|\nu\|_{\frac{r}{2-r}} : \left[\begin{array}{c|c}\mathrm{Diag}\{\mu\} & M^T\\ \hline M & \mathrm{Diag}\{\nu\}\end{array}\right] \succeq 0\right\}$$

is an upper bound on $\|M\|_{p,r}$, and this bound is tight within the factor $\vartheta = \left[\frac{2\sqrt{3}}{\pi} - \frac{2}{3}\right]^{-1} \approx 2.2936$:

$$\|M\|_{p,r} \leq \Psi_{p,r}(M) \leq \left[\frac{2\sqrt{3}}{\pi} - \frac{2}{3}\right]^{-1} \|M\|_{p,r}$$

(depending on values of p, r, the tightness factor can be improved; e.g., when $p = \infty, r = 2$, it is just $\sqrt{\pi/2} \approx 1.2533...$).

It follows that the explicit system of efficiently computable convex constraints

$$\left[\begin{array}{c|c} \mathrm{Diag}\{\mu\} & M^T \\ \hline M & \mathrm{Diag}\{\nu\} \end{array}\right] \succeq 0, \ \frac{1}{2}\left[\|\mu\|_{\frac{p}{p-2}} + \|\nu\|_{\frac{r}{2-r}}\right] \leq \alpha \tag{4.3.4}$$

in variables M, α, μ, ν is a safe tractable approximation of the constraint

$$\|M\|_{p,r} \leq \alpha,$$

which is tight within the factor ϑ. In some cases the value of the tightness factor can be improved; e.g., when $p = \infty$, $r = 2$ and when $p = 2$, $r = 1$, the tightness factor does not exceed $\sqrt{\pi/2}$.

Most of the tractable (or nearly so) cases considered so far deal with the case when $K^F = \{0\}$ (the only exception is the case $\mathbf{B}^*$ that, however, imposes severe restrictions on $\|\cdot\|_E$). In the GRC context, that means that we know nearly nothing about what to do when the recessive cone of the right hand side set $\mathbf{Q}$ in (4.1.6) is nontrivial, or, which is the same, $\mathbf{Q}$ is unbounded. This is not that disastrous — in many cases, boundedness of the right hand side set is not a severe restriction. However, it is highly desirable, at least from the academic viewpoint, to know something about the case when K^F is nontrivial, in particular, when K^F is a nonnegative orthant, or a Lorentz, or a Semidefinite cone (the two latter cases mean that (4.1.6) is an uncertain CQI, respectively, uncertain LMI). We are about to consider several such cases.

D: $F = \ell_\infty^m$, K^F *is a "sign" cone,* meaning that $K^F = \{u \in \ell_\infty^m : u_i \geq 0, i \in I_+, u_i \leq 0, i \in I_-, u_i = 0, i \in I_0\}$, *where I_+, I_-, I_0 are given non-intersecting subsets of the index set* $i = \{1, ..., m\}$.

The counterpart is

$\mathbf{D}^*$**:** $E = \ell_1^m$, $K^E = \{v \in \ell_1^m : v_j \geq 0, j \in J_+, v_j \leq 0, j \in J_-, v_j = 0, j \in J_0\}$, *where J_+, J_-, J_0 are given non-overlapping subsets of the index set* $\{1, ..., m\}$.

In the case of $\mathbf{D}^*$, assuming, for the sake of notational convenience, that $J_+ = \{1, ..., p\}$, $J_- = \{p+1, ..., q\}$, $J_0 = \{r+1, ..., m\}$ and denoting by e^j the standard basic orths in ℓ_1, we have

$$\begin{array}{rl} B \equiv & \{v \in K^E : \|v\|_E \leq 1\} = \mathrm{Conv}\{e^1, ..., e^p, -e^{p+1}, ..., -e^q, \pm e^{q+1}, ..., \pm e^r\} \\ \equiv & \mathrm{Conv}\{g^1, ..., g^s\}, s = 2r - q. \end{array}$$

Consequently,

$$\Psi(\mathcal{M}) = \max_{1 \leq j \leq s} \mathrm{dist}_{\|\cdot\|_F}(\mathcal{M}g^j, K^F)$$

is efficiently computable (cf. case $\mathbf{B}$).

E: $F = \ell_2^m$, $K^F = \mathbf{L}^m \equiv \{f \in \ell_2^m : f_m \geq \sqrt{\sum_{i=1}^{m-1} f_i^2}\}$, $E = \ell_2^n$, $K^E = E$.

The counterpart is

$\mathbf{E}^*$**:** $F = \ell_2^n$, $K^F = \{0\}$, $E = \ell_2^m$, $K^E = \mathbf{L}^m$.

In the case of $\mathbf{E}^*$, let $D = \{e \in K^E : \|e\|_2 \leq 1\}$, and let

$$B = \{e \in E : e_1^2 + ... + e_{m-1}^2 + 2e_m^2 \leq 1\}.$$

Let us represent a linear map $\mathcal{M} : \ell_2^m \to \ell_2^n$ by its matrix M in the standard bases of the origin and the destination spaces. Observe that

$$B \subset D_s \equiv \mathrm{Conv}\{D \cup (-D)\} \subset \sqrt{3/2}B \tag{4.3.5}$$

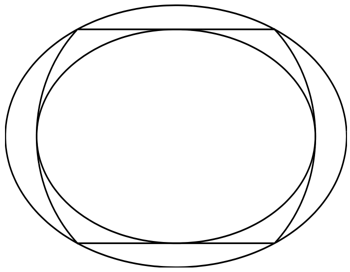

Figure 4.3: 2-D cross-sections of the solids B, $\sqrt{3/2}B$ (ellipses) and D_s by a 2-D plane passing through the common symmetry axis $e_1 = ... = e_{m-1} = 0$ of the solids.

(see figure 4.3). Now, let B_F be the unit Euclidean ball, centered at the origin, in $F = \ell_2^m$. By definition of $\Psi(\cdot)$ and due to $K^F = \{0\}$, we have

$$\Psi(\mathcal{M}) \leq \alpha \Leftrightarrow \mathcal{M}D \subset \alpha B_F \Leftrightarrow (\mathcal{M}D \cup (-\mathcal{M}D)) \subset \alpha B_F \Leftrightarrow \mathcal{M}D_s \subset \alpha B_F.$$

Since $D_s \subset \sqrt{3/2}B$, the inclusion $\mathcal{M}(\sqrt{3/2}B) \subset \alpha B_F$ is a sufficient condition for the validity of the inequality $\Psi(\mathcal{M}) \leq \alpha$, and since $B \subset D_s$, this condition is tight within the factor $\sqrt{3/2}$. (Indeed, if $\mathcal{M}(\sqrt{3/2}B) \not\subset \alpha B_F$, then $\mathcal{M}B \not\subset \sqrt{2/3}\alpha B_F$, meaning that $\Psi(\mathcal{M}) > \sqrt{2/3}\alpha$.) Noting that $\mathcal{M}(\sqrt{3/2}B) \leq \alpha$ if and only if $\|M\Delta\|_{2,2} \leq \alpha$, where $\Delta = \mathrm{Diag}\{\sqrt{3/2}, ..., \sqrt{3/2}, \sqrt{3/4}\}$, we conclude that the efficiently verifiable convex inequality

$$\|M\Delta\|_{2,2} \leq \alpha$$

is a safe tractable approximation, tight within the factor $\sqrt{3/2}$, of the constraint $\Psi(\mathcal{M}) \leq \alpha$.

F: $F = \mathbf{S}^m$, $\|\cdot\|_F = \|\cdot\|_{2,2}$, $K^F = \mathbf{S}^m_+$, $E = \ell_\infty^n$, $K^E = E$.
The counterpart is

F*: $F = \ell_1^n$, $K^F = \{0\}$, $E = \mathbf{S}^m$, $\|e\|_E = \sum\limits_{i=1}^m |\lambda_i(e)|$, *where* $\lambda_1(e) \geq \lambda_2(e) \geq ... \geq \lambda_m(e)$ *are the eigenvalues of* e, $K^E = \mathbf{S}^m_+$.
In the case of **F**, given $\mathcal{M} \in \mathcal{L}(\ell_\infty^n, \mathbf{S}^m)$, let $e^1, ..., e^n$ be the standard basic orths of ℓ_∞^n, and let $B_E = \{v \in \ell_\infty^n : \|u\|_\infty \leq 1\}$. We have

$$\begin{array}{l}\{\Psi(\mathcal{M}) \leq \alpha\} \Leftrightarrow \left\{\forall v \in B_E \, \exists V \succeq 0 : \max\limits_i |\lambda_i(\mathcal{M}v - V)| \leq \alpha\right\} \\ \Leftrightarrow \{\forall v \in B_E : \mathcal{M}v + \alpha I_m \succeq 0\}.\end{array}$$

Thus, the constraint

$$\Psi(\mathcal{M}) \leq \alpha \tag{$*$}$$

is equivalent to

$$\alpha I + \sum_{i=1}^n v_i(\mathcal{M}e^i) \succeq 0 \;\forall (v : \|v\|_\infty \leq 1).$$

It follows that the explicit system of LMIs

$$\begin{array}{l} Y_i \succeq \pm\mathcal{M}e^i, \; i = 1, ..., n \\ \alpha I_m \succeq \sum\limits_{i=1}^n Y_i \end{array} \tag{4.3.6}$$

in variables $\mathcal{M}$, α, $Y_1, ..., Y_n$ is a safe tractable approximation of the constraint $(*)$. Now let

$$\Theta(\mathcal{M}) = \vartheta(\mu(\mathcal{M})),\ \mu(\mathcal{M}) = \max_{1\leq i\leq n} \mathrm{Rank}(\mathcal{M}e^i),$$

where $\vartheta(\mu)$ is the function defined in the Real Case Matrix Cube Theorem, so that $\vartheta(1) = 1$, $\vartheta(2) = \pi/2$, $\vartheta(4) = 2$, and $\vartheta(\mu) \leq \pi\sqrt{\mu/2}$ for $\mu \geq 1$. Invoking this Theorem (see the proof of Theorem 3.4), we conclude that the *local* tightness factor of our approximation does not exceed $\Theta(\mathcal{M})$, meaning that if $(\mathcal{M}, \alpha)$ cannot be extended to a feasible solution of (4.3.6), then

$$\Theta(\mathcal{M})\Psi(\mathcal{M}) > \alpha.$$

4.3.3 Illustration: Robust Least Squares Antenna Design

We are about to illustrate our findings by applying a GRC-based approach to the Least Squares Antenna Design problem (see p. 104 and Example 1.1). Our motivation and course of actions here are completely similar to those we used in the case of $\|\cdot\|_\infty$ design considered in section 4.2.1. At present, we are interested in the uncertain conic problem

$$\left\{\min_{x,\tau} \left\{\tau : [h - H(I + \mathrm{Diag}\{\zeta\})x; \tau] \in \mathbf{L}^{m+1}\right\} : \|\zeta\|_\infty \leq \rho\right\}$$

where $H = WD \in \mathbb{R}^{m\times L}$ and $h = Wd \in \mathbb{R}^m$ are given matrix and vector, see p. 104. We equip the problem with the same perturbation structure (4.2.3) as in the case of $\|\cdot\|_\infty$-design:

$$\mathcal{Z} = \{\zeta \in \mathbb{R}^L : \|\zeta\|_\infty \leq \bar{\rho}\},\ \mathcal{L} = \mathbb{R}^L,\ \|\zeta\| \equiv \|\zeta\|_\infty$$

and augment this structure with the $\|\cdot\|_2$-norm on the embedding space $\mathbb{R}^{m+1}$ of the right hand side $\mathbf{Q} := \mathbf{L}^{m+1}$ of the conic constraint in question. Now the robust properties of a candidate design x are fully characterized by the worst-case loss

$$F_x(\rho) = \max_{\|\zeta\|_\infty\leq\rho} \|H(I + \mathrm{Diag}\{\zeta\})x - h\|_2,$$

which is a convex and nondecreasing function of $\rho \geq 0$. Invoking Proposition 4.2, it is easily seen that a par (τ, x) is feasible for the GRC of our uncertain problem with global sensitivity α if and only if

$$\begin{array}{c} F_{\bar{\rho}}(x) \leq \tau \ \&\ \ \alpha \geq \alpha(x) := \max_\zeta \left\{\mathrm{dist}_{\|\cdot\|_2}([D[x]\zeta; 0], \mathbf{L}^{m+1}) : \|\zeta\|_\infty \leq 1\right\} \\ D[x] = H\mathrm{Diag}\{x\} \in \mathbb{R}^{m\times L}. \end{array}$$

Similarly to the case of $\|\cdot\|_\infty$-synthesis, the minimal possible value $\alpha(x)$ of α depends solely on x; we call it *global sensitivity* of a candidate design x. Note that $F_x(\cdot)$ and $\alpha(x)$ are linked by the relation similar, although not identical to, the relation $\alpha(x) = \lim\limits_{\rho\to\infty} \frac{d}{d\rho}F_x(\rho)$ we had in the case of $\|\cdot\|_\infty$-design; now this relation modifies to

$$\alpha(x) = 2^{-1/2} \lim_{\rho\to+\infty} \frac{d}{d\rho}F_x(\rho).$$

Indeed, denoting $\beta(x) = \lim\limits_{\rho\to+\infty} \frac{d}{d\rho} F_x(\rho)$ and taking into account the fact that $F_x(\cdot)$ is a nondecreasing convex function, we have $\beta(x) = \lim\limits_{\rho\to\infty} F_x(\rho)/\rho$. Now, from the structure of F it immediately follows that $\lim\limits_{\rho\to\infty} F_x(\rho)/\rho = \max\limits_{\zeta}\{\|D[x]\zeta\|_2 : \|\zeta\|_\infty \leq 1\}$. Observing that the $\|\cdot\|_2$-distance from a vector $[u;0] \in \mathbb{R}^{m+1}$ to the Lorentz cone $\mathbf{L}^{m+1}$ clearly is $2^{-1/2}\|u\|_2$ and looking at the definition of $\alpha(x)$, we conclude that $\alpha(x) = 2^{-1/2}\beta(x)$, as claimed.

Now, in the case of $\|\cdot\|_\infty$-synthesis, the quantities $F_x(\rho)$ and $\alpha(x)$ were easy to compute, which is not the case now. However, we can built tight within the factor $\sqrt{\pi/2}$ tractable *upper bounds* on these quantities, namely, as follows.

We have

$$\begin{array}{l} F_x(\rho) = \max\limits_{\zeta:\|\zeta\|_\infty\leq\rho} \|h - H(1+\mathrm{Diag}\{\zeta\})x\|_2 = \max\limits_{\zeta:\|\zeta\|_\infty\leq\rho} \|h - Hx + D[x]\zeta\|_2 \\ \quad = \max\limits_{\|[t;\eta]\|_\infty\leq 1} \|t[h-Hx] + \rho D[x]\eta\|_2, \end{array}$$

and the concluding quantity is nothing but the norm of the linear map $[t:\zeta] \mapsto [h - Hx; \rho D[x]][t;\zeta]$ induced by the norm $\|\cdot\|_\infty$ in the origin and the norm $\|\cdot\|_2$ in the destination spaces. By Nesterov's theorem, see p. 194, the efficiently computable quantity

$$\widehat{F}_x(\rho) = \min_{\mu,\nu}\left\{\frac{\|\mu\|_1+\nu}{2} : \left[\begin{array}{c|c} \nu I_m & [h-Hx, \rho D[x]] \\ \hline [h-Hx,\rho D[x]]^T & \mathrm{Diag}\{\mu\}\end{array}\right] \succeq 0\right\} \tag{4.3.7}$$

is a tight within the factor $\sqrt{\pi/2}$ upper bound on $F_x(\rho)$; note that this bound, same as $F_x(\cdot)$ itself, is a convex and nondecreasing function of ρ.

Further, we have

$$\alpha(x) = \max_{\|\zeta\|_\infty\leq 1} \mathrm{dist}_{\|\cdot\|_2}([D[x]\zeta;0], \mathbf{L}^{m+1}) = 2^{-1/2}\max_{\|\zeta\|_\infty\leq 1}\|D[x]\zeta\|_2,$$

that is, $\alpha(x)$ is proportional to the norm of the linear mapping $\zeta \mapsto D[x]\zeta$ induced by the $\|\cdot\|_\infty$-norm in the origin and by the $\|\cdot\|_2$-norm in the destination spaces. Same as above, we conclude that the efficiently computable quantity

$$\widehat{\alpha}(x) = \min_{\mu,\nu}\left\{\frac{\|\mu\|_1+\nu}{2\sqrt{2}} : \left[\begin{array}{c|c} \nu I_m & D[x] \\ \hline D^T[x] & \mathrm{Diag}\{\mu\}\end{array}\right] \succeq 0\right\} \tag{4.3.8}$$

is a tight within the factor $\sqrt{\pi/2}$ upper bound on $\alpha(x)$. Note that $\widehat{F}_x(\cdot)$ and $\widehat{\alpha}(x)$ are linked by exactly the same relation as $F_x(\rho)$ and $\alpha(x)$, namely,

$$\widehat{\alpha}(x) = 2^{-1/2}\lim_{\rho\to\infty}\frac{d}{d\rho}\widehat{F}_x(\rho)$$

(why?).

Similarly to the situation of section 4.2.1, the function $\widehat{F}_x(\cdot)$ (and thus the true loss $F_x(\rho)$) admits the piecewise linear upper bound

$$\left\{\begin{array}{ll} \frac{\rho_0-\rho}{\rho_0}\widehat{F}_x(0) + \frac{\rho}{\rho_0}\widehat{F}_x(\rho_0), & 0 \leq \rho < \rho_0, \\ \widehat{F}_x(\rho_0) + 2^{1/2}\widehat{\alpha}(x)[\rho-\rho_0], & \rho \geq \rho_0, \end{array}\right. , \tag{4.3.9}$$

$\rho_0 \geq 0$ being the parameter of the bound.

We have carried out numerical experiments completely similar to those reported in section 4.2.1, that is, built solutions to the optimization problems

$$\widehat{\beta}_*(\delta) = \min_x\left\{\widehat{\alpha}(x) : F_x(0) \leq (1+\delta)\min_u\|Hu-h\|_2\right\}; \tag{4.3.10}$$

the results are presented in figure 4.4 and table 4.2.

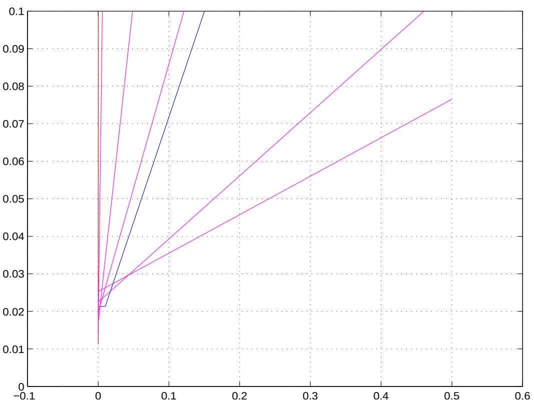

Figure 4.4: Red and magenta: bounds (4.3.9) on the losses for optimal solutions to (4.3.10) for the values of δ listed in table 4.2; the bounds correspond to $\rho_0 = 0$. Blue: bound (4.2.5) with $\rho_0 = 0.01$ on the loss $F_x(\rho)$ associated with the robust design built in section 3.3.1.

δ	0	0.25	0.50	0.75	1.00	1.25
$\widehat{\beta}_*(\delta)$	9441.4	14.883	1.7165	0.6626	0.1684	0.1025

Table 4.2: Tolerances δ and quantities $\widehat{\beta}_*(\delta)$ (tight within the factor $\pi/2$ upper bounds on global sensitivities of the optimal solutions to (4.3.10)). Pay attention to how huge is the global sensitivity ($\geq 2\widehat{\beta}_*(0)/\pi$) of the nominal Least Squares-optimal design. For comparison: the global sensitivity of the robust design $x_{\rm r}$ built in section 3.3.1 is $\leq \widehat{\alpha}(x_{\rm r}) = 0.3962$.

4.4 Illustration: Robust Analysis of Nonexpansive Dynamical Systems

We are about to illustrate the techniques we have developed by applying them to the problem of *robust nonexpansiveness analysis* coming from Robust Control; in many aspects, this problem resembles the Robust Lyapunov Stability Analysis problem we have considered in sections 3.4.2 and 3.5.1.

4.4.1 Preliminaries: Nonexpansive Linear Dynamical Systems

Consider an uncertain time-varying linear dynamical system (cf. (3.4.22)):

$$\begin{array}{rcl}\dot{x}(t) & = & A_t x(t) + B_t u(t)\\ y(t) & = & C_t x(t) + D_t u(t)\end{array} \tag{4.4.1}$$

where $x \in \mathbb{R}^n$ is the state, $y \in \mathbb{R}^p$ is the output and $u \in \mathbb{R}^q$ is the control. The system is assumed to be uncertain, meaning that all we know about the matrix $\Sigma_t = \left[\begin{array}{c|c} A_t & B_t \\ \hline C_t & D_t\end{array}\right]$ is that at every time instant t it belongs to a given *uncertainty set* $\mathcal{U}$.

System (4.4.1) is called *nonexpansive* (more precisely, *robustly nonexpansive w.r.t. uncertainty set* $\mathcal{U}$), if

$$\int\limits_0^t y^T(s)y(s)ds \leq \int\limits_0^t u^T(s)u(s)ds$$

for all $t \geq 0$ and for all trajectories of (all realizations of) the system such that $z(0) = 0$. In what follows, we focus on the simplest case of a system with $y(t) \equiv x(t)$, that is, on the case of $C_t \equiv I$, $D_t \equiv 0$. Thus, from now on the system of interest is

$$\begin{array}{rcl}\dot{x}(t) & = & A_t x(t) + B_t u(t)\\ & & [A_t, B_t] \in \mathcal{AB} \subset \mathbb{R}^{n\times m}\ \forall t,\\ & & m = n + q = \dim x + \dim u.\end{array} \tag{4.4.2}$$

Robust nonexpansiveness now reads

$$\int\limits_0^t x^T(s)x(s)ds \leq \int\limits_0^t u^T(s)u(s)ds \tag{4.4.3}$$

for all $t \geq 0$ and all trajectories $x(\cdot)$, $x(0) = 0$, of all realizations of (4.4.2).

Similarly to robust stability, robust nonexpansiveness admits a *certificate* that is a matrix $X \in \mathbf{S}^n_+$. Specifically, such a certificate is a solution of the following system of LMIs in matrix variable $X \in \mathbf{S}^m$:

$$\begin{array}{ll}(a) & X \succeq 0\\ (b) & \forall [A,B] \in \mathcal{AB}:\\ & \mathcal{A}(A,B;X) \equiv \left[\begin{array}{c|c} -I_n - A^TX - XA & -XB \\ \hline -B^TX & I_q\end{array}\right] \succeq 0.\end{array} \tag{4.4.4}$$

The fact that solvability of (4.4.4) is a *sufficient* condition for robust nonexpansiveness of (4.4.2) is immediate: if X solves (4.4.4), $x(\cdot)$, $u(\cdot)$ satisfy (4.4.2) and $x(0) = 0$, then

$$\begin{array}{l} u^T(s)u(s) - x^T(s)x(s) - \frac{d}{ds}\left[x^T(s)Xx(s)\right] = u^T(s)u(s) - x^T(s)x(s) \\ -[\dot{x}^T(s)Xx(s) + x^T(s)X\dot{x}(s)] = u^T(s)u(s) - x^T(s)x(s) \\ -[A_s x(s) + B_s u(s)]^T X x(s) - x^T(s)X[A_s x(s) + B_s u(s)] \\ = \left[x^T(s), u^T(s)\right] \mathcal{A}(A_s, B_s; X) \left[\begin{array}{c} x(s) \\ u(s) \end{array}\right] \geq 0, \end{array}$$

whence

$$\begin{array}{l} t > 0 \Rightarrow \int\limits_0^t [u^T(s)u(s) - x^T(s)x(s)]ds \geq x^T(t)Xx(t) - x^T(0)Xx(0) \\ = x^T(t)Xx(t) \geq 0. \end{array}$$

It should be added that when (4.4.2) is time-invariant, (i.e., $\mathcal{AB}$ is a singleton) and satisfies mild regularity conditions, the existence of the outlined certificate, (i.e., the solvability of (4.4.4)), is sufficient *and necessary* for nonexpansiveness.

Now, (4.4.4) is nothing but the RC of the system of LMIs in matrix variable $X \in \mathbf{S}^n$:

$$\begin{array}{ll} (a) & X \succeq 0 \\ (b) & \mathcal{A}(A, B; X) \in \mathbf{S}^m_+, \end{array} \tag{4.4.5}$$

the uncertain data being $[A, B]$ and the uncertainty set being $\mathcal{AB}$. From now on we focus on the *interval uncertainty*, where the uncertain data $[A, B]$ in (4.4.5) is parameterized by perturbation $\zeta \in \mathbb{R}^L$ according to

$$[A, B] = [A_\zeta, B_\zeta] \equiv [A^{\mathrm{n}}, B^{\mathrm{n}}] + \sum_{\ell=1}^{L} \zeta_\ell e_\ell f_\ell^T; \tag{4.4.6}$$

here $[A^{\mathrm{n}}, B^{\mathrm{n}}]$ is the nominal data and $e_\ell \in \mathbb{R}^n$, $f_\ell \in \mathbb{R}^m$ are given vectors.

Imagine, e.g., that the entries in the uncertain matrix $[A, B]$ drift, independently of each other, around their nominal values. This is a particular case of (4.4.6) where $L = nm$, $\ell = (i, j)$, $1 \leq i \leq n$, $1 \leq j \leq m$, and the vectors e_ℓ and f_ℓ associated with $\ell = (i, j)$ are, respectively, the i-th standard basic orth in $\mathbb{R}^n$ multiplied by a given deterministic real δ_ℓ ("typical variability" of the data entry in question) and the j-th standard basic orth in $\mathbb{R}^m$.

4.4.2 Robust Nonexpansiveness: Analysis via GRC

The GRC setup and its interpretation

We are about to consider the GRC of the uncertain system of LMIs (4.4.5) affected by interval uncertainty (4.4.6). Our "GRC setup" will be as follows:

1. We equip the space $\mathbb{R}^L$ where the perturbation ζ lives with the uniform norm $\|\zeta\|_\infty = \max_\ell |\zeta_\ell|$, and specify the normal range of ζ as the box

$$\mathcal{Z} = \{\zeta \in \mathbb{R}^L : \|\zeta\|_\infty \leq r\} \tag{4.4.7}$$

with a given $r > 0$.

2. We specify the cone $\mathcal{L}$ as the entire $E = \mathbb{R}^L$, so that all perturbations are "physically possible."

3. The only uncertainty-affected LMI in our situation is (4.4.5.b); the right hand side in this LMI is the positive semidefinite cone $\mathbf{S}^{n+m}_+$ that lives in the space $\mathbf{S}^m$ of symmetric $m \times m$ matrices equipped with the Frobenius Euclidean structure. We equip this space with the standard spectral norm $\|\cdot\| = \|\cdot\|_{2,2}$.

Note that our setup belongs to what was called "case **F**" on p. 196.

Before processing the GRC of (4.4.5), it makes sense to understand what does it actually mean that X is a feasible solution to the GRC with global sensitivity α. By definition, this means three things:

A. $X \succeq 0$;

B. X is a robust feasible solution to (4.4.5.b), the uncertainty set being

$$\mathcal{AB}_r \equiv \{[A_\zeta, b_\zeta] : \|\zeta\|_\infty \leq r\},$$

see (4.4.6); this combines with **A** to imply that if the perturbation $\zeta = \zeta^t$ underlying $[A_t, B_t]$ all the time remains in its normal range $\mathcal{Z} = \{\zeta : \|\zeta\|_\infty \leq r\}$, the uncertain dynamical system (4.4.2) is robustly nonexpansive.

C. When $\rho > r$, we have

$$\forall(\zeta, \|\zeta\|_\infty \leq \rho) : \mathrm{dist}(\mathcal{A}(A_\zeta, B_\zeta; X), \mathbf{S}^m_+) \leq \alpha \mathrm{dist}(\zeta, \mathcal{Z}|\mathcal{L}) = \alpha(\rho - r),$$

or, recalling what is the norm on $\mathbf{S}^m$,

$$\forall(\zeta, \|\zeta\|_\infty \leq \rho) : \mathcal{A}(A_\zeta, B_\zeta; X) \succeq -\alpha(\rho - r) I_m. \tag{4.4.8}$$

Now, repeating word for word the reasoning we used to demonstrate that (4.4.4) is sufficient for robust nonexpansiveness of (4.4.2), one can extract from (4.4.8) the following conclusion:

(!) *Whenever in uncertain dynamical system* (4.4.2) *one has* $[A_t, B_t] = [A_{\zeta^t}, B_{\zeta^t}]$ *and the perturbation* ζ^t *remains all the time in the range* $\|\zeta^t\|_\infty \leq \rho$, *one has*

$$(1 - \alpha(\rho - r)) \int_0^t x^T(s) x(s) ds \leq (1 + \alpha(\rho - r)) \int_0^t u^T(s) u(s) ds \tag{4.4.9}$$

for all $t \geq 0$ *and all trajectories of the dynamical system such that* $x(0) = 0$.

We see that global sensitivity α indeed controls "deterioration of nonexpansiveness" as the perturbations run out of their normal range $\mathcal{Z}$: when the $\|\cdot\|_\infty$ distance from ζ^t to $\mathcal{Z}$ all the time remains bounded by $\rho - r \in [0, \frac{1}{\alpha})$, relation (4.4.9) guarantees that the L_2 norm of the state trajectory on every time horizon can be bounded by constant times the L_2 norm of the control on the this time horizon. The corresponding constant $\left(\frac{1+\alpha(\rho-r)}{1-\alpha(\rho-r)}\right)^{1/2}$ is equal to 1 when $\rho = r$ and grows with ρ, blowing up to $+\infty$ as $\rho - r$ approaches the critical value α^{-1}, and the larger α, the smaller is this critical value.

Processing the GRC

Observe that (4.4.4) and (4.4.6) imply that

$$\begin{array}{c}\mathcal{A}(A_\zeta, B_\zeta; X) = \mathcal{A}(A^{\rm n}, B^{\rm n}; X) - \sum_{\ell=1}^{L} \zeta_\ell \left[L_\ell^T(X)R_\ell + R_\ell^T L_\ell(X)\right], \\ L_\ell^T(X) = [Xe_\ell; 0_{m-n,1}],\ R_\ell^T = f_\ell. \end{array} \tag{4.4.10}$$

Invoking Proposition 4.2, the GRC in question is equivalent to the following system of LMIs in variables X and α:

$$\begin{array}{ll} (a) & X \succeq 0 \\ (b) & \forall(\zeta, \|\zeta\|_\infty \leq r): \\ & \qquad \mathcal{A}(A^{\rm n}, B^{\rm n}; X) + \sum_{\ell=1}^{L} \zeta_\ell \left[L_\ell^T(X)R_\ell + R_\ell^T L_\ell(X)\right] \succeq 0 \\ (c) & \forall(\zeta, \|\zeta\|_\infty \leq 1): \sum_{\ell=1}^{L} \zeta_\ell \left[L_\ell^T(X)R_\ell + R_\ell^T L_\ell(X)\right] \succeq -\alpha I_m. \end{array} \tag{4.4.11}$$

Note that the semi-infinite LMIs $(4.4.11.b, c)$ are affected by structured norm-bounded uncertainty with 1×1 scalar perturbation blocks (see section 3.5.1). Invoking Theorem 3.13, the system of LMIs

$$\begin{array}{ll} (a) & X \succeq 0 \\ (b.1) & Y_\ell \succeq \pm\left[L_\ell^T(X)R_\ell + R_\ell^T L_\ell(X)\right],\ 1 \leq \ell \leq L \\ (b.2) & \mathcal{A}(A^{\rm n}, B^{\rm n}; X) - r\sum_{\ell=1}^{L} Y_\ell \succeq 0 \\ (c.1) & Z_\ell \succeq \pm\left[L_\ell^T(X)R_\ell + R_\ell^T L_\ell(X)\right],\ 1 \leq \ell \leq L \\ (c.2) & \alpha I_m - \sum_{\ell=1}^{L} Z_\ell \succeq 0 \end{array}$$

in matrix variables $X, \{Y_\ell, Z_\ell\}_{\ell=1}^{L}$ and in scalar variable α is a safe tractable approximation of the GRC, tight within the factor $\frac{\pi}{2}$. Invoking the result stated in (!) on p. 106, we can reduce the design dimension of this approximation; the equivalent reformulation of the approximation is the SDO program

$$\min \alpha$$

s.t.

$$\begin{array}{l} X \succeq 0 \\ \left[\begin{array}{c|ccc} \mathcal{A}(A^{\rm n}, B^{n}; X) - r\sum_{\ell=1}^{L} \lambda_\ell R_\ell^T R_\ell & L_1^T(X) & \cdots & L_L^T(X) \\ \hline L_1(X) & \lambda_1/r & & \\ \vdots & & \ddots & \\ L_L(X) & & & \lambda_L/r \end{array}\right] \succeq 0 \\ \left[\begin{array}{c|ccc} \alpha I_m - \sum_{\ell=1}^{L} \mu_\ell R_\ell^T R_\ell & L_1^T(X) & \cdots & L_L^T(X) \\ \hline L_1(X) & \mu_1 & & \\ \vdots & & \ddots & \\ L_L(X) & & & \mu_L \end{array}\right] \succeq 0 \end{array} \tag{4.4.12}$$

in variable $X \in \mathbf{S}^m$ and scalar variables α, $\{\lambda_\ell, \mu_\ell\}_{\ell=1}^{L}$. Note that we have equipped our (approximate) GRC with the objective to minimize the global sensitivity of X; of course, other choices of the objective are possible as well.

Numerical illustration

The data. In the illustration we are about to present, the state dimension is $n = 5$, and the control dimension is $q = 2$, so that $m = \dim x + \dim u = 7$. The nominal data (chosen at random) are as follows:

$$
\begin{array}{l}
[A^{\rm n}, B^{\rm n}] \\
= M := \left[\begin{array}{rrrrr|rr}
-1.089 & -0.079 & -0.031 & -0.575 & -0.387 & 0.145 & 0.241 \\
-0.124 & -2.362 & -2.637 & 0.428 & 1.454 & -0.311 & 0.150 \\
-0.627 & 1.157 & -1.910 & -0.425 & -0.967 & 0.022 & 0.183 \\
-0.325 & 0.206 & 0.500 & -1.475 & 0.192 & 0.209 & -0.282 \\
0.238 & -0.680 & -0.955 & -0.558 & -1.809 & 0.079 & 0.132
\end{array}\right].
\end{array}
$$

The interval uncertainty (4.4.6) is specified as

$$
[A_\zeta, b_\zeta] = M + \sum_{i=1}^{5}\sum_{j=1}^{7} \zeta_{ij} \underbrace{|M_{ij}| g_i}_{e_i} f_j^T,
$$

where g_i, f_j are the standard basic orths in $\mathbb{R}^5$ and $\mathbb{R}^7$, respectively; in other words, every entry in $[A, B]$ is affected by its own perturbation, and the variability of an entry is the magnitude of its nominal value.

Normal range of perturbations. Next we should decide how to specify the normal range $\mathcal{Z}$ of the perturbations, i.e., the quantity r in (4.4.7). "In reality" this choice could come from the nature of the dynamical system in question and the nature of its environment. In our illustration there is no "nature and environment," and we specify r as follows. Let r_* be the largest r for which the robust nonexpansiveness of the system at the perturbation level r, (i.e., the perturbation set being the box $B_r = \{\zeta : \|\zeta\|_\infty \le r\}$) admits a certificate. It would be quite reasonable to choose, as the normal range of perturbations $\mathcal{Z}$, the box B_{r_*}, so that the normal range of perturbations is the largest one where the robust nonexpansiveness still can be certified. Unfortunately, precise checking the existence of a certificate for a given box in the role of the perturbation set means to check the feasibility status of the system of LMIs

$$
\begin{array}{ll}
(a) & X \succeq 0 \\
(b) & \forall (\zeta, \|\zeta\|_\infty \le r) : \mathcal{A}(A_\zeta, B_\zeta; X) \succeq 0
\end{array}
$$

in matrix variable X, with $\mathcal{A}(\cdot, \cdot; \cdot)$ given in (4.4.4). This task seems to be intractable, so that we are forced to replace this system with its safe tractable approximation, tight within the factor $\pi/2$, specifically, with the system

$$
\begin{array}{l}
X \succeq 0 \\
\left[\begin{array}{c|ccc}
\mathcal{A}(A^{\rm n}, B^n; X) - r\sum_{\ell=1}^{L} \lambda_\ell R_\ell^T R_\ell & L_1^T(X) & \cdots & L_L^T(X) \\
\hline
L_1(X) & \lambda_1/r & & \\
\vdots & & \ddots & \\
L_L(X) & & & \lambda_L/r
\end{array}\right] \succeq 0
\end{array}
\qquad (4.4.13)
$$

in matrix variable X and scalar variables λ_ℓ (cf. (4.4.12)), with $R_\ell(X)$ and L_ℓ given by (4.4.10). The largest value r_1 of r for which the latter system is solvable (this quantity can be easily found by bisection) is a lower bound, tight within the factor $pi/2$, on r_*, and this is the quantity we use in the role of r when specifying the normal range of perturbations according to (4.4.7).

Applying this approach to the outlined data, we end up with

$$r = r_1 = 0.0346.$$

The results. With the outlined nominal and perturbation data and r, the optimal value in (4.4.12) turns out to be

$$\alpha_{\rm GRC} = 27.231.$$

It is instructive to compare this quantity with the global sensitivity of the RC-certificate $X_{\rm RC}$ of robust nonexpansiveness; by definition, $X_{\rm RC}$ is the X component of a feasible solution to (4.4.13) where r is set to r_1. This X clearly can be extended to a feasible solution to our safe tractable approximation (4.4.12) of the GRC; the smallest, over all these extensions, value of the global sensitivity α is

$$\alpha_{\rm RC} = 49.636,$$

which is by a factor 1.82 larger than $\alpha_{\rm GRC}$. It follows that the GRC-based analysis of the robust nonexpansiveness properties of the uncertain dynamical system in question provides us with essentially more optimistic results than the RC-based analysis. Indeed, a feasible solution $(\alpha, ...)$ to (4.4.12) provides us with the *upper bound*

$$C_*(\rho) \le C_\alpha(\rho) \equiv \left\{ \begin{array}{ll} 1, & 0 \le \rho \le r \\ \frac{1+\alpha(\rho-r)}{1-\alpha(\rho-r)}, & r \le \rho < r + \alpha^{-1} \end{array} \right. \tag{4.4.14}$$

(cf. (4.4.9)) on the "existing in the nature, but difficult to compute" quantity

$$\begin{array}{rl} C_*(\rho) = & \inf \Big\{ C : \int_0^t x^T(s)x(s)ds \le C \int_0^t u^T(s)u(s)ds \,\forall (t \ge 0, x(\cdot), u(\cdot)) : \\ & \qquad x(0) = 0,\, \dot{x}(s) = A_{\zeta^s}x(s) + B_{\zeta^s}u(s), \|\zeta^s\|_\infty \le \rho \,\forall s \Big\} \end{array}$$

responsible for the robust nonexpansiveness properties of the dynamical system. The upper bounds (4.4.14) corresponding to $\alpha_{\rm RC}$ and $\alpha_{\rm GRC}$ are depicted on the left plot in figure 4.5 where we see that the GRC-based bound is much better than the RC-based bound.

Of course, both the bounds in question are conservative, and their "level of conservatism" is difficult to access theoretically: while we do understand how conservative our tractable approximations to intractable RC/GRC are, we have no idea how conservative the sufficient condition (4.4.4) for robust nonexpansiveness is (in this respect, the situation is completely similar to the one in Lyapunov Stability Analysis, see section 3.5.1). We can, however, run a brute force simulation to bound $C_*(\rho)$ from below. Specifically, generating a sample of perturbations of a given magnitude and checking the associated matrices $[A_\zeta, B_\zeta]$ for nonexpansiveness, we can build an upper bound $\overline{\rho}_1$ on the largest ρ for which every matrix $[A_\zeta, B_\zeta]$ with $\|\zeta\|_\infty \le \rho$ generates a nonexpansive time-invariant dynamical system; $\overline{\rho}_1$ is, of course, greater than or equal to the largest $\rho = \rho_1$ for which $C_*(\rho) \le 1$. Similarly, testing matrices A_ζ for stability, we can build an upper bound $\overline{\rho}_\infty$ on the largest $\rho = \rho_\infty$ for which all matrices A_ζ, $\|\zeta\|_\infty \le \rho$, have all their eigenvalues in the closed left hand side plane; it is immediately seen that $C_*(\rho) = \infty$ when $\rho > \rho_\infty$. For our nominal and perturbation data, simulation yields

$$\overline{\rho}_1 = 0.310, \quad \overline{\rho}_\infty = 0.7854.$$

These quantities should be compared, respectively, to $r_1 = 0.0346$, (which clearly is a *lower* bound on the range ρ_1 of ρ's where $C_*(\rho) \le 1$) and $r_\infty = r_1 + \alpha_{\rm GRC}^{-1}$ (this is the range of values of ρ where the GRC-originating upper bound (4.4.14) on $C_*(\rho)$ is finite; as such, r_∞ is a *lower* bound on ρ_∞). We see that in our numerical example the conservatism of our approach is "within one order of magnitude": $\overline{\rho}_1/r_1 \approx 8.95$ and $\overline{\rho}_\infty/r_\infty \approx 11.01$.

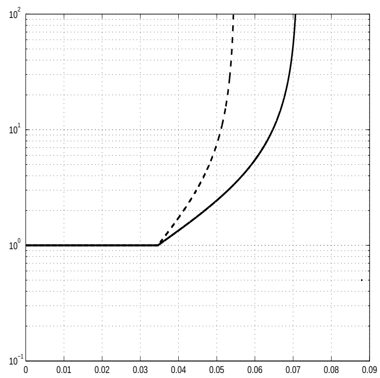

Figure 4.5: RC/GRC-based analysis: bounds (4.4.14) vs. ρ for $\alpha = \alpha_{\rm GRC}$ (solid) and $\alpha = \alpha_{\rm RC}$ (dashed).

4.5 Exercises

Exercise 4.1 Consider a situation as follows. A factory consumes n types of raw materials, coming from n different suppliers, to be decomposed into m pure components. The per unit content of component i in raw material j is $p_{ij} \geq 0$, and the necessary per month amount of component i is a given quantity $b_i \geq 0$. You need to make a long-term arrangement on the amounts of raw materials x_j coming every month from each of the suppliers, and these amounts should satisfy the system of linear constraints

$$Px \geq b, \; P = [p_{ij}].$$

The current per unit price of product j is c_j; this price, however, can vary in time, and from the history you know the volatilities $v_j \geq 0$ of the prices. How to choose x_j's in order to minimize the total cost of supply at the current prices, given an upper bound α on the sensitivity of the cost to possible future drifts in prices?

Test your model on the following data:

$$\begin{array}{c} n = 32, m = 8, p_{ij} \equiv 1/m, b_i \equiv 1.e3, \\ c_j = 0.8 + 0.2\sqrt{((j-1)/(n-1))}, v_j = 0.1(1.2 - c_j), \end{array}$$

and build the tradeoff curve "supply cost with current prices vs. sensitivity."

Exercise 4.2 Consider the Structural Design problem (p. 122 in section 3.4.2).

Recall that the nominal problem is

$$\min_{t,\tau} \left\{ \tau : t \in \mathcal{T}, \left[\begin{array}{c|c} 2\tau & f^T \\ \hline f & A(t) \end{array} \right] \succeq 0 \, \forall f \in \mathcal{F} \right\} \tag{SD}$$

where $\mathcal{T}$ is a given compact convex set of admissible designs, $\mathcal{F} \subset \mathbb{R}^m$ is a finite set of loads of interest and $A(t)$ is an $m \times m$ symmetric stiffness matrix affinely depending on t. Assuming that the load f "in reality" can run out of the set of scenarios $\mathcal{F}$ and thus can be treated as uncertain data element, we can look for a robust design. For the time being, we have already considered two robust versions of the problem. In the first, we extended the set $\mathcal{F}$ of actual loads of interest was extended to its union with an ellipsoid E centered at the origin (and thus we were interested to minimize the worst, with respect to loads of interest *and all "occasional" loads from E*, compliance of the construction. In the second (p. 154, section 3.6.2), we imposed an upper bound τ_* on the compliance w.r.t. the loads of interest and chance constrained the compliance w.r.t. Gaussian occasional loads. Along with these settings, it might make sense to control both the compliance w.r.t. loads of interest *and* its deterioration when these loads are perturbed. This is what we intend to consider now.

Consider the following "GRC-type" robust reformulation of the problem: we want to find a construction such that its compliance w.r.t. a load $f + \rho g$, where $f \in \mathcal{F}$, g is an occasional load from a given ellipsoid E centered at the origin, and $\rho \geq 0$ is a perturbation level, never exceeds a prescribed function $\phi(\rho)$. The corresponding robust problem is the semi-infinite program

(!) Given $\phi(\cdot)$, find $t \in \mathcal{T}$ such that

$$\left[\begin{array}{c|c} 2\phi(\rho) & [f+\rho g]^T \\ \hline f+\rho g & A(t) \end{array}\right] \succeq 0 \ \forall (f \in \mathcal{F}, g \in E).$$

1. Let us choose $\phi(\rho) = \tau + \alpha\rho$ – the choice straightforwardly inspired by the GRC methodology. Does the corresponding problem (!) make sense? If not, why the GRC methodology does not work in our case?

2. Let us set $\phi(\rho) = (\sqrt{\tau} + \rho\sqrt{\alpha})^2$. Does the corresponding problem (!) make sense? If yes, does (!) admit a computationally tractable reformulation?

Exercise 4.3 Consider the situation as follows:

Unknown signal z known to belong to a given ball $B = \{z \in \mathbb{R}^n : z^T z \leq 1\}$, $Q \succ 0$, is observed according to the relation

$$y = Az + \xi,$$

where y is the observation, A is a given $m \times n$ sensing matrix, and ξ is an observation error. Given y, we want to recover a linear form $f^T z$ of z; here f is a given vector. The normal range of the observation error is $\Xi = \{\xi \in \mathbb{R}^m : \|\xi\|_2 \leq 1\}$. We are seeking for a linear estimate $\widehat{f}(y) = g^T y$.

1. Formulate the problem of building the best, in the minimax sense (i.e., with the minimal worst case, w.r.t. $x \in B$ and $\xi \in \Xi$, recovering error), linear estimate as the RC of an uncertain LO problem and build tractable reformulation of this RC.

2. Formulate the problem of building a linear estimate with the worst case, over signals z with $\|z\|_2 \leq 1 + \rho_z$ and observation errors ξ with $\|\xi\|_2 \leq 1 + \rho_\xi$, risk for all $\rho_z, \rho_\xi \geq 0$, risk admitting the bound $\tau + \alpha_z\rho_z + \alpha_\xi\rho_\xi$ with given $\tau, \alpha_z, \alpha_x i$; thus, we want "desired performance" τ of the estimate in the normal range $B \times \Xi$ of $[z; \xi]$ and "controlled deterioration of this performance" when z and/or ξ run out of their normal ranges. Build a tractable reformulation of this problem.

Exercise 4.4 Consider situation completely similar to the one in Exercise 4.3, with the only difference that now we want to build a linear estimate Gy of the vector Cz, C being a given matrix, rather than to estimate a linear form of g.

1. Formulate the problem of building the best, in the minimax sense (i.e., with the minimal worst case, w.r.t. $z \in B$ and $\xi \in \Xi$, $\|\cdot\|_2$-recovering error), linear estimate of Cz as the RC of an uncertain Least Squares inequality and build a tight safe tractable approximation of this RC.

2. Formulate the problem of building a linear estimate of Cz with the worst case, over signals z with $\|z\|_2 \leq 1 + \rho_z$ and observation errors ξ with $\|\xi\|_2 \leq 1 + \rho_\xi$, risk admitting for all $\rho_z, \rho_\xi \geq 0$ the bound $\tau + \alpha_z \rho_z + \alpha_\xi \rho_\xi$ with given $\tau, \alpha_z, \alpha_x i$. Find a tight safe tractable approximation of this problem.