Linear Optimization Problem, its Data and Structure

♣ **Linear Optimization problem:**

$$\min_x \left\{ c^T x + d : Ax \le b \right\} \qquad \textbf{(LO)}$$

- $x \in \mathbb{R}^n$: vector of *decision variables*,
- $c \in \mathbb{R}^n$ and $d \in \mathbb{R}$ form the *objective*,
- A: an $m \times n$ *constraint matrix*,
- $b \in \mathbb{R}^m$: *right hand side*.

♠ **Problem's structure:** its sizes m, n.
♠ **Problem's data:** (c, d, A, b).

Data Uncertainty

♣ The data of typical real world LOs are partially *uncertain* — not known exactly when the problem is being solved.
♠ **Sources of data uncertainty:**

- **Prediction errors.** Some of data entries (future demands, returns, etc.) do not exist when the problem is solved and hence are replaced with their forecasts.

• **Measurement errors:** Some of the data (parameters of technological devices and processes, contents associated with raw materials, etc.) cannot be measured exactly, and their true values drift around the measured "nominal" values.

• **Implementation errors:** Some of the decision variables (planned intensities of technological processes, parameters of physical devices we are designing, etc.) cannot be implemented exactly as computed. *The implementation errors are equivalent to artificial data uncertainties.*

Indeed, the impact of implementation errors $x_j \mapsto (1+\epsilon_j)x_j + \delta_j$ on the validity of the constraint

$$a_{i1}x_1 + \ldots + a_{in}x_n \leq b_i$$

is *as if* there were no implementation errors, but the data of the constraint was subject to perturbations

$$a_{ij} \mapsto (1+\epsilon_j)a_{ij}, \;\; b_i \mapsto b_i - \sum_j a_{ij}\delta_j.$$

Data Uncertainty: Traditional Treatment and Dangers

♣ **Traditionally,**

♠ **"small" (fractions of percents) data uncertainty is just ignored, the problem is solved "as it is" – with the nominal data, and the resulting nominal optimal solution is forwarded to the end user;**

♠ **"large" data uncertainty is assigned with a probability distribution and is treated via Stochastic Programming techniques.**

♠ **Fact:** *in many cases, even small data uncertainty can make the nominal solution heavily infeasible and thus practically meaningless.*

♣ **Example: Antenna Design**

♠ [Physics:] *Directional density of energy transmitted by an monochromatic antenna placed at the origin is proportional to* $|D(\delta)|^2$, *where the antenna's diagram* $D(\delta)$ *is a complex-valued function of 3-D direction (unit 3-D vector)* δ.

♠ [Physics:] **For an** *antenna array* **— a complex antenna comprised of a number of antenna elements, the diagram is**

$$D(\delta) = \sum_j x_j D_j(\delta) \qquad (*)$$

- $D_j(\cdot)$: **diagrams of elements**
- x_j: **complex** *weights* **– design parameters responsible for how the elements in the array are invoked.**

♠ **Antenna Design problem:** *Given diagrams* $D_1(\cdot), ..., D_k(\cdot)$ *and a target diagram* $D_*(\cdot)$, *find the weights* $x_i \in \mathbb{C}$ *such that the synthesized diagram* $(*)$ *is as close as possible to the target diagram* $D_*(\cdot)$.

♡ **When** $D_j(\cdot)$, $D_*(\cdot)$, **same as the weights, are real and the "closeness' is quantified by the uniform norm on a finite grid** Γ **of directions, Antenna Design becomes the LO problem**

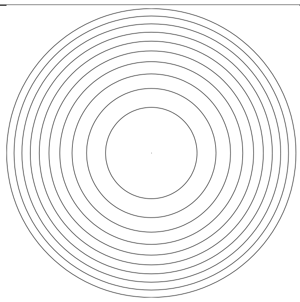
$$\min_{x \in \mathbb{R}^n, \tau} \left\{ \tau : -\tau \leq D_*(\delta) - \sum_j x_j D_j(\delta) \leq \tau \ \forall \delta \in \Gamma \right\}.$$

♠ **Example:** **Consider planar antenna array comprised of 10 elements (circle surrounded by 9 rings of equal areas) in the plane XY (Earth's surface"), and our goal is to send most of the energy "up," along the 12^o cone around the Z-axis:**

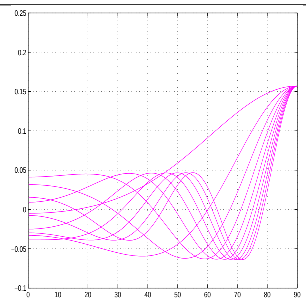
• **Diagram of a ring** $\{z=0, a \le \sqrt{x^2+y^2} \le b\}$**:**

$$D_{a,b}(\theta) = \frac{1}{2}\int\limits_a^b \left[\int\limits_0^{2\pi} r\cos\left(2\pi r\lambda^{-1}\cos(\theta)\cos(\phi)\right)d\phi\right]dr,$$

• θ: **altitude angle** • λ: **wavelength**



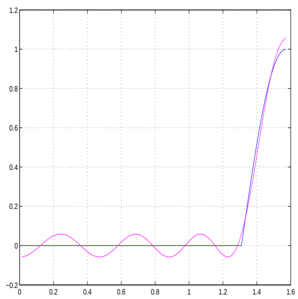
10 antenna elements,
equal areas, outer radius 1 m



Diagrams of the elements
vs the altitude angle θ, λ **=50 cm**

• **Nominal design problem:**

$$\tau_* = \min_{x\in\mathbb{R}^{10},\tau}\left\{\tau : -\tau \le D_*(\theta_i) - \sum_{j=1}^{10} x_j D_j(\theta_i) \le \tau,\ 1 \le i \le 240, \theta_i = \frac{i\pi}{480}\right\}$$

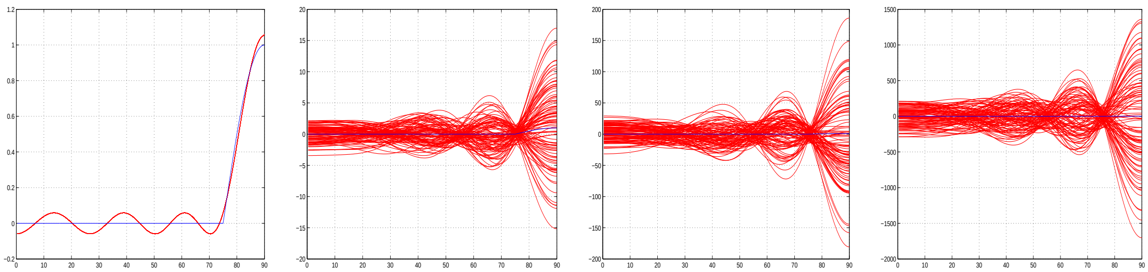


Target (blue) and nominal
optimal (magenta) diagrams,
$\tau_* = 0.0589$

But: *The design variables are characteristics of physical devices and as such they cannot be implemented exactly as computed. What happens when there are implementation errors:*

$$x_j^{\rm fact} = (1+\xi_j)x_j^{\rm comp},\ \ \xi_j \sim \textbf{Uniform}[-\rho,\rho]$$

with small ρ*?*



$\rho = 0$ $\rho = 0.0001$ $\rho = 0.001$ $\rho = 0.01$

"Dream and reality," nominal optimal design: samples of 100 actual diagrams (red) for different uncertainty levels. Blue: the target diagram

	Dream	Reality								
	$\rho = 0$	$\rho = 0.0001$			$\rho = 0.001$			$\rho = 0.01$		
	value	min	mean	max	min	mean	max	min	mean	max
$\|\cdot\|_\infty$-distance to target	0.059	1.280	5.671	14.04	11.42	56.84	176.6	39.25	506.5	1484
energy concentration	85.1%	0.5%	16.4%	51.0%	0.1%	16.5%	48.3%	0.5%	14.9%	47.1%

Quality of nominal antenna design: dream and reality. Data over 100 samples of actuation errors per each uncertainty level ρ.

♠ **Conclusion:** *Nominal optimal design is completely meaningless...*

♣ **Example: NETLIB Case Study.**

♠ **NETLIB: a collection of LO problems for testing LO algorithms.**

♠ **Constraint # 372 of the NETLIB problem** `PILOT4`**:**

$$
\begin{array}{rcl}
a^Tx & \equiv & -15.79081x_{826} - 8.598819x_{827} - 1.88789x_{828} - 1.362417x_{829} - 1.526049x_{830} \\
& & -0.031883x_{849} - 28.725555x_{850} - 10.792065x_{851} - 0.19004x_{852} - 2.757176x_{853} \\
& & -12.290832x_{854} + 717.562256x_{855} - 0.057865x_{856} - 3.785417x_{857} - 78.30661x_{858} \\
& & -122.163055x_{859} - 6.46609x_{860} - 0.48371x_{861} - 0.615264x_{862} - 1.353783x_{863} \\
& & -84.644257x_{864} - 122.459045x_{865} - 43.15593x_{866} - 1.712592x_{870} - 0.401597x_{871} \\
& & +x_{880} - 0.946049x_{898} - 0.946049x_{916} \\
& \geq & b \equiv 23.387405
\end{array}
\tag{C}
$$

The related *nonzero* coordinates in the optimal solution x^* of the problem as reported by `CPLEX` **are:**

$$
\begin{array}{lll}
x^*_{826} = 255.6112787181108 & x^*_{827} = 6240.488912232100 & x^*_{828} = 3624.613324098961 \\
x^*_{829} = 18.20205065283259 & x^*_{849} = 174397.0389573037 & x^*_{870} = 14250.00176680900 \\
x^*_{871} = 25910.00731692178 & x^*_{880} = 104958.3199274139 &
\end{array}
$$

This solution makes (C) an equality within machine precision.

♠ **Note: The coefficients in a, except for the coefficient 1 at x_{880}, are "ugly reals" like -15.79081 or -84.644257. Ugly coefficients characterize certain technological devices and processes; as such *they could hardly be known to high accuracy* and coincide with the "true" data within accuracy of 3-4 digits, not more.**

Question: *Assuming that the ugly entries in a are 0.1%-accurate approximations of the true data $\widetilde{a}$, what is the effect of this uncertainty on the validity of the "true" constraint $\widetilde{a}^Tx \geq b$ as evaluated at x^*?*

Answer:

• **The minimum, over all** 0.1% **perturbations** $a \mapsto \widetilde{a}$ **of ugly entries in** a**, value of** $\widetilde{a}^T x^* - b$**, is** < -104.9**, that is,** *with 0.1% perturbations of ugly coefficients, the violation of the constraint as evaluated at the nominal solution can be as large as* 450% *of the right hand side!*

• **With independent** *random* 0.1%**-perturbations of ugly coefficients,**

— **the violation of the constraint** *at average* **is as large as 125% of the right hand side;**

— **the probability of violating the constraint by** *at least* **150% of the right hand side is as large as 0.18.**

♣ **Among 90 NETLIB problems, perturbing ugly coefficients** *by just 0.01%* **results in violating some of the constraints, as evaluated at nominal optimal solutions,**

— **by more than 50% – in 13 problems,**

— **by more than 100% – in 6 problems.**

— **by 210,000% – in** `PILOT4`**.**

♣ **Conclusion:** *In applications of LO, there exists a real need of a technique capable of detecting cases when data uncertainty can heavily affect the quality of the nominal solution, and in these cases to generate a "reliable" solution, one that is immunized against uncertainty.*

Robust Optimization is aimed at satisfying the above need.

Uncertain Linear Optimization Problems

♣ **Definition:** *An uncertain LO problem is a collection*

$$\left\{\min_x \left\{c^T x + d : Ax \leq b\right\}\right\}_{(c,d,A,b)\in\mathcal{U}} \qquad (LO_{\mathcal{U}})$$

of LO problems (instances) $\min_x \left\{c^T x + d : Ax \leq b\right\}$ *of common structure* (i.e., with common numbers m of constraints and n of variables) *with the data varying in a given uncertainty set* $\mathcal{U} \subset \mathbb{R}^{(m+1)\times(n+1)}$.

♠ **Usually we assume that the uncertainty set is parameterized, in an affine fashion, by** *perturbation vector* ζ **varying in a given** *perturbation set* $\mathcal{Z}$:

$$\mathcal{U} = \left\{ \left[\begin{array}{c|c} c^T & d \\ \hline A & b \end{array}\right] = \underbrace{\left[\begin{array}{c|c} c_0^T & d_0 \\ \hline A_0 & b_0 \end{array}\right]}_{\substack{\text{nominal} \\ \text{data } D_0}} + \sum_{\ell=1}^{L} \zeta_\ell \underbrace{\left[\begin{array}{c|c} c_\ell^T & d_\ell \\ \hline A_\ell & b_\ell \end{array}\right]}_{\substack{\text{basic} \\ \text{shifts } D_\ell}} : \zeta \in \mathcal{Z} \subset \mathbb{R}^L \right\}.$$

Example: When speaking about PILOT4, we tacitly used the following model of uncertainty:
Uncertainty affects only the 'ugly" coefficients $\{a_{ij} : (i,j) \in \mathcal{J}\}$ in the constraint matrix, and every one of them is allowed to run, independently of all other coefficients, through the interval

$$[a_{ij}^{\rm n} - \rho_{ij}|a_{ij}^{\rm n}|, a_{ij}^{\rm n} + \rho_{ij}|a_{ij}^{\rm n}|]$$

- $a_{ij}^{\rm n}$**: nominal values of the data**
- ρ_{ij}**: perturbation levels (which in the experiment were set to $\rho = 0.001$).**

- **Perturbation set: The box**

$$\{\zeta = \{\zeta_{ij}\}_{(i,j)\in\mathcal{J}} : -\rho_{ij} \leq \zeta_{ij} \leq \rho_{ij}\}$$

- **Parameterization of the data by perturbation vector:**

$$\left[\begin{array}{c|c} c^T & d \\ \hline A & b \end{array}\right] = \left[\begin{array}{c|c} [c^{\rm n}]^T & d^{\rm n} \\ \hline A^{\rm n} & b^{\rm n} \end{array}\right] + \sum_{(i,j)\in\mathcal{J}} \zeta_{ij} \left[\begin{array}{c|c} & \\ \hline e_i e_j^T & \end{array}\right]$$

$$\left\{\min_x \left\{c^T x + d : Ax \leq b\right\}\right\}_{(c,d,A,b)\in\mathcal{U}}$$

$$\mathcal{U} = \left\{ \underbrace{\left[\begin{array}{c|c} c_0^T & d_0 \\ \hline A_0 & b_0 \end{array}\right]}_{\substack{\text{nominal} \\ \text{data } D_0}} + \sum_{\ell=1}^{L} \zeta_\ell \underbrace{\left[\begin{array}{c|c} c_\ell^T & d_\ell \\ \hline A_\ell & b_\ell \end{array}\right]}_{\substack{\text{basic} \\ \text{shifts } D_\ell}} : \zeta \in \mathcal{Z} \subset \mathbb{R}^L \right\}. \qquad (\mathbf{LO}_{\mathcal{U}})$$

♣ **There is no universally defined notion of a "solution to a *family* of optimization problems," like $(\mathbf{LO}_{\mathcal{U}})$.**

Consider "decision environment" as follows:

- **A.1. All decision variables in $(\mathbf{LO}_{\mathcal{U}})$ represent "here and now" decisions; they should be assigned specific numerical values as a result of solving the problem *before* the actual data "reveals itself."**
- **A.2. The decision maker is fully responsible for consequences of the decisions to be made when, and only when, the actual data is within the prespecified uncertainty set $\mathcal{U}$.**
- **A.3. The constraints in $(\mathbf{LO}_{\mathcal{U}})$ are *hard* — we cannot tolerate violations of constraints, even small ones, when the data is in $\mathcal{U}$.**

$$\left\{\min_x \left\{c^T x + d : Ax \le b\right\}\right\}_{(c,d,A,b)\in\mathcal{U}} \qquad (\mathbf{LO}_{\mathcal{U}})$$

♣ **In the above decision environment, the only meaningful candidate solutions to** $(\mathbf{LO}_{\mathcal{U}})$ **are the** *robust feasible* **ones.**

Definition: *$x \in \mathbb{R}^n$ is called a robust feasible solution to $(LO_{\mathcal{U}})$, if x is feasible for all instances:*

$$Ax \le b \;\forall (c,d,A,b) \in \mathcal{U}.$$

> **Indeed, by A.1 a meaningful candidate solution should be independent of the data, i.e., it should be just a fixed vector x. By A.2-3, it should satisfy the constraints, whatever be a realization of the data from $\mathcal{U}$.**

♠ **Acting in the same "worst-case-oriented" fashion, it makes sense to quantify the quality of a candidate solution x by the** *guaranteed* **(the worst, over the data from $\mathcal{U}$) value of the objective:**

$$\sup\{c^T x + d : (c,d,A,b) \in \mathcal{U}\}$$

$$\left\{\min_x \left\{c^T x + d : Ax \le b\right\}\right\}_{(c,d,A,b)\in\mathcal{U}} \qquad (\mathbf{LO}_{\mathcal{U}})$$

♠ **Now we can associate with $(\mathbf{LO}_{\mathcal{U}})$ the problem of finding the best,** *in terms of the guaranteed value of the objective*, **among the** *robust feasible* **solutions:**

$$\min_{t,x} \left\{t : c^T x + d \le t, Ax \le b \ \forall (c,d,A,b) \in \mathcal{U}\right\} \qquad \textbf{(RC)}$$

This is called the *Robust Counterpart* **of $(\mathbf{LO}_{\mathcal{U}})$.**

Note: Passing from LOs of the form

$$\min_x \left\{c^T x + d : Ax \le b\right\}$$

to their equivalents

$$\min_{t,x} \left\{t : c^T x + d \le t, Ax \le b\right\}$$

we always may assume that the objective is certain, and the RC respects this equivalence.

⇒ *We lose nothing by assuming the objective in $(LO_{\mathcal{U}})$ certain, in which case we can think of $\mathcal{U}$ as of the set in the space $\mathbb{R}^{m\times(n+1)}$ of the $[A,b]$-data, and the RC reads*

$$\min_x \left\{c^T x : Ax \le b \ \forall [A,b] \in \mathcal{U}\right\}. \qquad \textbf{(RC)}$$

$$\left\{\min_x \left\{c^T x : Ax \leq b\right\}\right\}_{(A,b)\in\mathcal{U}} \qquad (\mathbf{LO}_{\mathcal{U}})$$

$$\Downarrow$$

$$\min_x \left\{c^T x : Ax \leq b \;\forall [A,b] \in \mathcal{U}\right\} \quad \textbf{(RC)}$$

♣ **Fact I:** *The RC of uncertain LO with certain objective is a purely constraint-wise construction: when building the RC, we replace every constraint* $a_i^T x \leq b_i$ *of the instances with its RC*

$$a_i^T x \leq b_i \;\forall [a_i^T, b_i] \in \mathcal{U}_i \qquad (\mathrm{RC}_i)$$

where $\mathcal{U}_i$ *is the projection of the uncertainty set* $\mathcal{U}$ *on the space of data* $[a_i^T, b_i]$ *of* i*-th constraint.*

♣ **Fact II:** *The RC remains intact when extending the uncertainty set* $\mathcal{U}$ *to its closed convex hull.*

When $(\mathrm{LO}_{\mathcal{U}})$ *has certain objective, the RC remains intact when extending* $\mathcal{U}$ *to the direct product of closed convex hulls of* $\mathcal{U}_i$*. Thus, the transformation*

$$\mathcal{U} \mapsto \mathcal{U}^+ = [\mathrm{cl}\,\mathrm{Conv}(\mathcal{U}_1)] \times \ldots \times [\mathrm{cl}\,\mathrm{Conv}(\mathcal{U}_m)]$$

keeps the RC intact.

♠ *From now on, we always assume uncertainty set* $\mathcal{U}$ *convex, and perturbation set* $\mathcal{Z}$ *– convex and closed.*

$$\left\{\min_x \left\{c^T x : Ax \le b\right\}\right\}_{[A,b]\in\mathcal{U}} \qquad (\mathbf{LO}_{\mathcal{U}})$$

$$\Downarrow$$

$$\min_x \left\{c^T x : Ax \le b \; \forall [A,b] \in \mathcal{U}\right\} \quad \textbf{(RC)}$$

♣ **The central questions associated with the concept of RC are:**

A. *What is the "computational status" of the RC? When is it possible to process the RC efficiently?*

— to be addressed in-depth below.

B. *How to come-up with meaningful uncertainty sets?*

— modeling issue to be partly addressed in the sequel.

$$\min_x \left\{ c^T x : Ax \le b \ \forall [A,b] \in \mathcal{U} \right\} \qquad \textbf{(RC)}$$

♣ **Potentially bad news:** *The RC is a semi-infinite optimization problem (finitely many variables, infinitely many constraints) and as such can be computationally tractable.*

Example: Consider an "essentially linear" semi-infinite constraint

$$\|Px - p\|_1 \le 1, \ \forall [P,p] \in \mathcal{U}$$
$$\mathcal{U} = \{[P_*, p] : p = B\zeta, \ \|\zeta\|_2 \le 1\}$$

To check whether $x = 0$ is robust feasible is the same as to check whether

$$\max_{\zeta : \|\zeta\|_2 \le 1} \|B\zeta\|_1 \le 1. \qquad (!)$$

(!) is equivalent to

$$1 \ge \max_{\|\zeta\|_2 \le 1} \|B\zeta\|_1 = \max_{z:\|z\|_\infty \le 1, \zeta : \|\zeta\|_2 \le 1} z^T B\zeta$$
$$= \max_{z:\|z\|_\infty \le 1} \underbrace{\max_{\zeta:\|\zeta\|_2 \le 1} \zeta^T [B^T z]}_{\|B^T z\|_2} = \sqrt{\max_{z:\|z\|_\infty \le 1} z^T [BB^T] z}$$

Since BB^T can be an arbitrary symmetric positive semidefinite matrix, and finding the maximum of a nonnegative quadratic form over the box $\{\|z\|_\infty \le 1\}$ is NP-hard, even when relative accuracy like 4% is sought, *checking* **(!)** *is heavily computationally intractable.*

$$\min_x \left\{ c^T x : Ax \leq b \;\forall [A,b] \in \mathcal{U} \right\} \qquad \textbf{(RC)}$$

♣ **Good news:** *The RC of an uncertain LO problem is computationally tractable, provided the uncertainty set $\mathcal{U}$ is so.*

Explanation, I: The RC can be written down as the optimization problem

$$\min_x \left\{ c^T x : f_i(x) \leq 0,\; i = 1, ..., m \right\}$$
$$f_i(x) = \sup_{[A,b] \in \mathcal{U}} \left[a_i^T x - b_i \right]$$

• **The functions $f_i(x)$ are convex (due to their origin) and efficiently computable (as maxima of affine functions over computationally tractable convex sets).**

• **Thus, the RC is a Convex Programming program with efficiently computable objective and constraints, and problems of this type are efficiently solvable.**

♣ **The above "reasoning" refers to the notions of** *computationally tractable problem/convex set* **and on the fact that** *maximizing linear objective over a computationally tractable convex set*, **in particular, a convex set given by finitely many efficiently computable convex constraints,** *is a computationally tractable problem*. **While these notions and results can be rigorously defined and justified, it makes sense to present a somehow restricted "practical" version of them, highly instructive by its own rights and not requiring tedious an lengthy excursions to the complexity theory of continuous optimization.**

♣ **Recaling that the RC is a "constraint-wise" construction, all we need is to reformulate in a tractable form a** *single* **semi-infinite constraint**

$$\begin{array}{c}\forall \alpha = [a;b] \in \{\alpha_0 + \mathcal{A}\zeta : \zeta \in \mathcal{Z}\} \subset \mathbb{R}^{n+1}: \\ \alpha^T[x;1] \equiv a^T x + b \leq 0.\end{array} \qquad (*)$$

♠ **Consider several instructive cases when tractable reformulation of** $(*)$ **is easy – does not require any theory.**

1. Scenario uncertainty $\mathcal{Z} = \mathrm{Conv}\{\zeta^1, ..., \zeta^N\}$**. Setting** $\alpha^j = \alpha_0 + \mathcal{A}\zeta^j$, $1 \leq j \leq N$, **we get**

$$\mathcal{U} = \mathrm{Conv}\{\alpha^1, ..., \alpha^N\}$$

and therefore

$$(*) \Leftrightarrow \{\alpha^j[x;1] \leq 0,\ 1 \leq j \leq N\}$$

2. $\|\cdot\|_p$**-uncertainty** $\mathcal{Z} = \{\zeta \in \mathbb{R}^L : \|\zeta\|_p \leq 1\}$**. We have**

$$\begin{array}{rl} & \alpha^T[x;1] \leq 1\, \forall \alpha \in \mathcal{U} \\ \Leftrightarrow & [\alpha_0 + \mathcal{A}\zeta]^T[x;1] \leq 0\, \forall (\zeta : \|\zeta\|_p \leq 1) \\ \Leftrightarrow & \alpha_0^T[x;1] + \max\limits_{\|\zeta\|_p \leq 1} \zeta^T[\mathcal{A}^T[x;1]] \leq 0 \\ \Leftrightarrow & \alpha_0^T[x;1] + \|\mathcal{A}^T[x;1]\|_{p_*} \leq 0,\ \frac{1}{p} + \frac{1}{p_*} = 1 \end{array}$$

$$\begin{array}{c}\forall \alpha = [a;b] \in \{\alpha = \alpha_0 + \mathcal{A}\zeta : \zeta \in \mathcal{Z}\} \subset \mathbb{R}^{n+1} : \\ \alpha^T[x;1] := a^T x + b \le 0.\end{array} \qquad (*)$$

3. Intersection of simple perturbation sets: $\mathcal{Z} = \bigcap_{i=1}^k \mathcal{Z}_i$. **Let** $\mathcal{Z}_i, 0 \in \mathcal{Z}_i$, $1 \le i \le k$ **be convex compact sets such that** $\bigcap_{i=1}^k \mathrm{int}\mathcal{Z}_i \ne \emptyset$.

Fact from Convex Analysis: *For* $\mathcal{Z}, \mathcal{Z}_i$ *as above,*

$$\max_{\zeta\in\mathcal{Z}} \beta^T\zeta = \min_{\substack{\beta_1,\dots,\beta_k,\\ \beta_1+\dots+\beta_k=\beta}} \sum_{i=1}^k \max_{\zeta\in\mathcal{Z}_i} \beta_i^T\zeta.$$

Therefore,

$$\begin{array}{rl} & \alpha^T[x;1] \le 0 \ \forall \alpha \in \mathcal{U} \\ \Leftrightarrow & \alpha_0^T[x;1] + [\mathcal{A}\zeta]^T[x;1] \le 0 \,\forall \zeta \in \mathcal{Z} \\ \Leftrightarrow & \alpha_0^T[x;1] + \max_{\zeta\in\mathcal{Z}} \zeta^T[\mathcal{A}^T[x;1]] \le 0 \\ \Leftrightarrow & \exists \beta_1, \dots, \beta_k : \left\{\begin{array}{ll} \beta_1 + \dots + \beta_k = \mathcal{A}^T[x;1] & (a) \\ \alpha_0^T[x;1] + \sum_{i=1}^k \max_{\zeta\in\mathcal{Z}_i} \beta_i^T\zeta \le 0 & (b) \end{array}\right. \end{array}$$

Thus, $(*)$ *is represented by the system*

$$\left\{\begin{array}{ll} \beta_1 + \dots + \beta_k = \mathcal{A}^T[x;1] & (a) \\ \alpha_0^T[x;1] + \sum_{i=1}^k \max_{\zeta\in\mathcal{Z}_i} \beta_i^T\zeta \le 0 & (b) \end{array}\right. \qquad (S)$$

of constraints in variables x, $\beta_1, \dots, \beta_k$, **meaning that** x *can be extended to a feasible solution of* (S) *if and only if* x *is feasible for* $(*)$.

When $\mathcal{Z} = \bigcap\limits_{i=1}^{k} \mathcal{Z}_i$, $0 \in \mathcal{Z}$, $\bigcap_i \mathrm{int}\mathcal{Z}_i \neq \emptyset$, the system

$$\left\{ \begin{array}{ll} \beta_1 + ... + \beta_k = \mathcal{A}^T[x;1] & (a) \\ \alpha_0^T[x;1] + \sum_{i=1}^k \max\limits_{\zeta \in \mathcal{Z}_i} \beta_i^T \zeta \leq 0 & (b) \end{array} \right. \qquad (S)$$

of convex constraints in variables $x, \beta_1, ..., \beta_k$ represents the semi-infinite constraint

$$\alpha^T[x;1] \leq 0 \ \forall \alpha \in \{\alpha_0 + \mathcal{A}\zeta : \zeta \in \mathcal{Z}\}$$

Note: *When $\mathcal{Z}_i$ are simple, so that the convex functions $f_i(\beta_i) = \max\limits_{\zeta \in \mathcal{Z}_i} \beta_i^T \zeta$ are available in closed analytic form, (S) is a system of explicitly given convex constraints.*

Example: Ball-Box-Budgeted uncertainty

$\mathcal{Z} = \{\zeta : \|\zeta\|_\infty \leq \Omega_\infty\} \bigcap \{\zeta : \|\zeta\|_2 \leq \Omega_2\} \bigcap \{\zeta : \|\zeta\|_1 \leq \Omega_1\}$.

Here

$$\begin{array}{l} f_1(\beta) = \max\limits_{\zeta:\|\zeta\|_\infty \leq \Omega_\infty} \beta^T\zeta = \Omega_\infty \|\beta\|_1, \\ f_2(\beta) = \max\limits_{\zeta:\|\zeta\|_2 \leq \Omega_2} \beta^T\zeta = \Omega_2 \|\beta\|_2, \\ f_3(\beta) = \max\limits_{\zeta:\|\zeta\|_1 \leq \Omega_1} \beta^T\zeta = \Omega_1 \|\beta\|_\infty, \end{array}$$

and thus (S) is equivalent to the system of convex constraints

$$\left\{ \begin{array}{l} \beta_1 + \beta_2 + \beta_3 = \mathcal{A}^T[x;1] \\ \alpha_0^T[x;1] + \Omega_\infty \|\beta_1\|_1 + \Omega_2 \|\beta_2\|_2 + \Omega_1 \|\beta_3\|_\infty \leq 0 \end{array} \right.$$

General Well-Structured Case

Definition. *Let us say that a set* $\mathcal{X} \subset \mathbb{R}^N$ *is well-structured, if it admits a well-structured representation — a representation of the form*

$$\mathcal{X} = \left\{ x \in \mathbb{R}^N : \exists u \in \mathbb{R}^M : \left\{ \begin{array}{rcl} A_0x + B_0u + c_0 &=& 0 \\ A_1x + B_1u + c_1 &\in& \mathbf{K}_1 \\ &\cdots& \\ A_Kx + B_ku + c_k &\in& \mathbf{K}_K \end{array} \right\} \right\},$$

where $\mathbf{K}_k$*, for every* $k \leq K$*, is a simple cone, specifically,*
— either nonnegative orthant $\mathbb{R}^m_+ = \{x \in \mathbb{R}^m : x \geq 0\}$, $m = m_k$,
— or a Lorentz cone $\mathbf{L}^m = \{x \in \mathbb{R}^m : x_m \geq \sqrt{x_1^2 + ... + x_{m-1}^2}\}$, $m = m_k$,
— or a Semidefinite cone $\mathbf{S}^m_+$ *— the cone of positive semidefinite matrices in the space* $\mathbf{S}^m$ *of real symmetric* $m \times m$ *matrices,* $m = m_k$.

Example 1: *The set* $\mathcal{X} = \{x \in \mathbb{R}^N : \|x\|_1 \leq 1\}$ *admits polyhedral representation*

$$\mathcal{X} = \{x \in \mathbb{R}^N : \exists u \in \mathbb{R}^N : -u_i \leq x_i \leq u_i, \sum_i u_i \leq 1\}$$

$$= \left\{ x \in \mathbb{R}^n : \exists u \in \mathbb{R}^N : A_1 x + B_1 u + c_1 \equiv \begin{bmatrix} u_1 - x_1 \\ u_1 + x_1 \\ \vdots \\ u_N - x_N \\ u_N + x_N \\ 1 - \sum_i u_i \end{bmatrix} \in \mathbb{R}^{2N+1}_+ \right\}$$

Example 2: *The set* $\mathcal{X} = \{x \in \mathbb{R}^4_+ : x_1 x_2 x_3 x_4 \geq 1\}$ *admits conic quadratic representation*

$$\mathcal{X} = \left\{ x \in \mathbb{R}^4_+ : \exists u \in \mathbb{R}^3 : \left\{ \begin{array}{l} 0 \leq u_1 \leq \sqrt{x_1 x_2} \\ 0 \leq u_2 \leq \sqrt{x_3 x_4} \\ 1 \leq u_3 \leq \sqrt{u_1 u_2} \end{array} \right. \right\}$$

$$= \left\{ x \in \mathbb{R}^n : \exists u \in R^3 : \left\{ \begin{array}{rcl} [x_1; x_2; x_3; x_4; u_1; u_2; u_3 - 1] & \in & \mathbb{R}^7_+ \\ [2u_1; x_1 - x_2; x_1 + x_2] & \in & \mathbf{L}^3 \\ [2u_2; x_3 - x_4; x_3 + x_4] & \in & \mathbf{L}^3 \\ [2u_3; u_1 - u_2; u_1 + u_2] & \in & \mathbf{L}^3 \end{array} \right. \right\}$$

Example 3: *The set* $\mathcal{X}$ *of* $m \times n$ *matrices* X *with nuclear norm (sum of singular values)* ≤ 1 *admits semidefinite representation*

$$\mathcal{X} = \left\{ X \in \mathbb{R}^{m \times n} : \exists u = (U \in \mathbf{S}^m, V \in \mathbf{S}^n) : \begin{array}{rcl} \mathrm{Tr}(U) + \mathrm{Tr}(V) & \leq & 2 \\ \left[\begin{array}{c|c} U & X \\ \hline X^T & V \end{array}\right] & \succeq & 0 \end{array} \right\}.$$

$$\mathcal{X} = \left\{ x \in \mathbb{R}^N : \exists u \in \mathbb{R}^M : \left\{ \begin{array}{rcl} A_0x + B_0u + c_0 & = & 0 \\ A_1x + B_1u + c_1 & \in & \mathbf{K}_1 \\ & \cdots & \\ A_Kx + B_ku + c_k & \in & \mathbf{K}_K \end{array} \right\} \right\}, \qquad (*)$$

♣ **Good news** on **w**ell-**s**tructured **r**epresentations:

• **Computational tractability:** *Minimizing a linear objective over a set given by* $(*)$ *reduces to solving a well-structured conic program*

$$\min_{x,u} \left\{ c^T x : \left\{ \begin{array}{rcl} A_0x + B_0u + c_0 & = & 0 \\ A_1x + B_1u + c_1 & \in & \mathbf{K}_1 \\ & \cdots & \\ A_Kx + B_ku + c_k & \in & \mathbf{K}_K \end{array} \right\} \right\},$$

and thus can be done in a theoretically (and to some extent — also practically) *efficient manner by polynomial time interior point algorithms.*

• **Extremely powerful expressive abilities:** *w.-s.r.'s admit a simple fully algorithmic calculus which makes it easy to build a w.-s.r. for the result of a convexity-preserving operation with convex sets* (like taking intersections, direct sums, affine images, inverse affine images, polars, etc.) *via w.-s.r.'s of the operands.*

As a result, *for all practical purposes, all computationally tractable convex sets arising in Optimization admit explicit w.-s.r.'s.*

♣ **The RC Tractability Theorem:** *Let the perturbation set $\mathcal{Z}$ of a semi-infinite linear inequality*

$$\alpha^T[x;1] \leq 0 \ \forall \alpha \in \{\alpha_0 + \mathcal{A}\zeta : \zeta \in \mathcal{Z}\} \qquad (*)$$

be nonempty and be given by w.-s.r.

$$\mathcal{Z} = \left\{ \zeta \in \mathbb{R}^L : \exists u \in \mathbb{R}^M : \left\{ \begin{array}{rcl} A_0 x + B_0 u + c_0 & = & 0 \\ A_1 x + B_1 u + c_1 & \in & \mathbf{K}_1 \\ & \cdots & \\ A_K x + B_k u + c_k & \in & \mathbf{K}_K \end{array} \right. \right\} \qquad (!)$$

When not all the cones $\mathbf{K}_k$ are nonnegative orthants, assume that (!) is strictly feasible, that is, there exist $\bar{x}$ and $\bar{u}$ such that

$$A_0\bar{x} + B_0\bar{u} + c_0 = 0 \ \& \ A_k\bar{x} + B_k\bar{u} + c_k \in \mathrm{int}\mathbf{K}_k,\ 1 \leq k \leq K.$$

Then the feasible set $\mathcal{X}$ of $(*)$ admits an explicit w.-s.r., specifically,

$$\mathcal{X} = \left\{ x : \exists z = [z^0; ...; z^K] : \left\{ \begin{array}{l} \sum_{k=0}^K A_k^* z^k + \mathcal{A}^T[x;1] = 0 \\ \sum_{k=0}^K B_k^* z^k = 0 \\ \alpha_0^T[x;1] + \sum_{k=0}^K \langle z^k, c_k \rangle \leq 0 \\ z^k \in \mathbf{K}_k,\ 1 \leq k \leq K \end{array} \right. \right\}$$

Here for a linear map $e \mapsto Be$ from a Euclidean space $(E, \langle\cdot,\cdot\rangle_E)$ to a Euclidean space $(F, \langle\cdot,\cdot\rangle_F)$ the adjoint map $f \mapsto B^*f : F \to E$ is given by

$$\langle f, Be \rangle_F \equiv \langle B^* f, e \rangle_E$$

Proof heavily utilizes the *Conic Duality Theorem* which answers the following question:

♣ Consider a *conic program*

$$\mathrm{Opt}(P) = \min_y \left\{ \langle c, y\rangle : \left\{ \begin{array}{l} A_0 y - b_0 = 0 \\ A_k y - b_k \in \mathbf{K}_k, \\ 1 \le k \le K \end{array} \right\} \right\}, \qquad (P)$$

where $\mathbf{K}_k$ are cones (closed, convex, pointed and with a nonempty interior) in Euclidean spaces E_k, $1 \le k \le K$.

How to bound from below, in a systematic way, the optimal value of the program?

♠ Consider an approach as follows. Let

$$\mathbf{K}_k^* = \{u \in E_k : \langle u, v\rangle \ge 0 \,\forall v \in \mathbf{K}_k\}$$

be the cones *dual* to $\mathbf{K}_k$. Let us choose $z^0 \in \mathbb{R}^{\dim b_0}$ and $z^k \in \mathbf{K}_k^*$, $1 \le k \le K$, and let y be feasible for (P). By feasibility, we have

$$\langle z^k, A_k y - b_k\rangle \ge 0,\ 0 \le k \le K,$$

or, which is the same,

$$\langle A_k^* z^k, y\rangle \ge \langle z^k, b_k\rangle,\ 0 \le k \le K.$$

Summing up, we get

$$\langle \sum_{k=0}^K A_k^* z^k, y\rangle \ge \sum_{k=0}^K \langle z^k, b_k\rangle.$$

$$\mathrm{Opt}(P) = \min_y \left\{ \langle c, y\rangle : \left\{ \begin{array}{l} A_0 y - b_0 = 0 \\ A_k y - b_k \in \mathbf{K}_k, \\ 1 \leq k \leq K \end{array} \right\} \right\}, \qquad (P)$$

Intermediate summary: *Whenever $z^0 \in \mathbb{R}^{\dim b_0}$ and $z^k \in \mathbf{K}_k^*$, $1 \leq k \leq K$, every feasible solution y of (P) satisfies the inequality*

$$\langle \sum_{k=0}^{K} A_k^* z^k, y\rangle \geq \sum_{k=0}^{K} \langle z^k, b_k\rangle. \qquad (*)$$

Conclusion: *When the left hand side in $(*)$ is identically in $y \in \mathbb{R}^N$ equal to $\langle c, y\rangle$, the right hand side in $(*)$ is a lower bound on $\mathrm{Opt}(P)$.* **In other words,** *The optimal value $\mathrm{Opt}(D)$ in the conic dual of (P), that is, in the problem*

$$\mathrm{Opt}(D) = \max_{\{z^k\}} \left\{ \sum_{k=0}^{K} \langle z^k, b_k\rangle : \begin{array}{c} z^k \in \mathbf{K}_k^*, \\ 1 \leq k \leq K \\ \sum_{k=0}^{K} A_k^* z^k = c \end{array} \right\} \qquad (D)$$

is a lower bound on $\mathrm{Opt}(P)$. **[“Weak Duality”]**

♣ **Conic Duality Theorem:** *If (P) is strictly feasible and below bounded, then (D) is solvable, and $\mathrm{Opt}(P) = \mathrm{Opt}(D)$.*

Note: *When $\mathbf{K}_k = \mathbb{R}_+^{m_k}$ for all k, “strict feasibility” can be weakened to “feasibility.”*

$$\alpha^T[x;1] \leq 0 \;\forall \alpha \in \{\alpha_0 + \mathcal{A}\zeta : \zeta \in \mathcal{Z}\} \qquad (*)$$

$$\mathcal{Z} = \left\{ \zeta \in \mathbb{R}^L : \exists u \in \mathbb{R}^M : \left\{ \begin{array}{rcl} A_0x + B_0u + c_0 & = & 0 \\ A_1x + B_1u + c_1 & \in & \mathbf{K}_1 \\ & \cdots & \\ A_Kx + B_ku + c_k & \in & \mathbf{K}_K \end{array} \right. \right\} \qquad (!)$$

Observe that x is feasible for $(*)$ iff

$$\mathrm{Opt}(P) := \min_{\zeta \in \mathcal{Z}} \left\{ [-\mathcal{A}\zeta]^T[x;1] \right\} \geq \alpha_0^T[x;1],$$

or, which is the same, iff

$$\begin{array}{rcl} \mathrm{Opt}(P) & := & \min\limits_{\zeta,u} \left\{ [-\mathcal{A}^T[x;1]]^T\zeta : \left\{ \begin{array}{rcl} A_0x + B_0u + c_0 & = & 0 \\ A_1x + B_1u + c_1 & \in & \mathbf{K}_1 \\ & \cdots & \\ A_Kx + B_ku + c_k & \in & \mathbf{K}_K \end{array} \right. \right\} \\ & \geq & \alpha_0^T[x;1] \end{array}$$

By CDT, and noting that $\mathbf{K}_k^* = \mathbf{K}_k$ for our cones, this is the case iff the problem

$$\max_{[z^0;\ldots;z^K]} \left\{ -\sum_{k=0}^{L} \langle z^k, c_k \rangle : \left\{ \begin{array}{l} \sum_{k=0}^K A_k^* z^k = -\mathcal{A}^T[x;1] \\ \sum_{k=0}^K B_k^* z^k = 0 \\ z^k \in \mathbf{K}_k,\ 1 \leq k \leq K \end{array} \right. \right\}$$

has a solution with the value of the objective $\geq$ $\alpha_0^T[x;1]$.

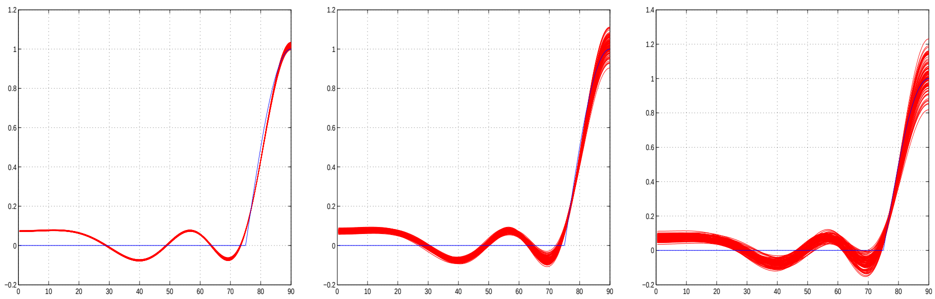
• **Thus, x is feasible for $(*)$ iff there exists $z = [z^0; ...; z^K]$ such that**

$$\begin{array}{l}\alpha_0^T[x;1] + \sum_{k=0}^K \langle z^k, c_k\rangle \leq 0\\ \sum_{k=0}^K A_k^* z^k + \mathcal{A}^T[x;1] = 0\\ \sum_{k=0}^K B_k^* z^k = 0\\ z^k \in \mathbf{K}_k,\ 1 \leq k \leq K\end{array} \qquad \Box$$

How it Works: Antenna Design

$$\min_{\tau,x}\left\{\tau : -\tau \le D_*(\theta_i) - \sum_{j=1}^{10} x_j D_j(\theta_i) \le \tau,\ 1 \le i \le I\right\}$$
$$x_j \mapsto (1+\zeta_j)x_j,\ -\rho \le \zeta_j \le \rho$$
$$\Rightarrow \min_{\tau,x}\left\{\tau : \begin{array}{c} D_*(\theta_i) - \sum_j x_j D_j(\theta_i) - \rho \sum_j |x_j||D_j(\theta_i)| \ge -\tau \\ D_*(\theta_i) - \sum_j x_j D_j(\theta_i) + \rho \sum_j |x_j||D_j(\theta_i)| \le \tau \end{array},\ 1 \le i \le I\right\} \quad \text{(RC)}$$

♠ **Solving (RC) at uncertainty level $\rho = 0.01$, we arrive at** *robust design*. **The robust optimal value is** 0.0815 **(39% more than the nominal optimal value** 0.0589**).**



$\rho = 0.01$ $\qquad$ $\rho = 0.05$ $\qquad$ $\rho = 0.1$

"Dream and reality," robust optimal design: samples of 100 actual diagrams (red) for different uncertainty levels. Blue: the target diagram.

	Reality								
	$\rho = 0.01$			$\rho = 0.05$			$\rho = 0.1$		
	min	mean	max	min	mean	max	min	mean	max
$\|\cdot\|_\infty$-distance to target	0.075	0.078	0.081	0.077	0.088	0.114	0.082	0.113	0.216
energy concentration	70.3%	72.3%	73.8%	63.6%	71.6%6	79.3%	52.2%	70.8%	87.5%

Robust optimal design, data over 100 samples of actuation errors per each uncertainty level ρ. For *nominal* design with $\rho = 0.001$, the average $\|\cdot\|_\infty$-distance to target is 56.8, and energy concentration is 16.5%.

How it Works: NETLIB Case Study

♣ **At uncertainty level** $\rho = 0.001$**,** *the RCs of all 90 NETLIB problems are feasible***, and** *the robust optimal values of all problems are within 1% of their nominal optimal values.*