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**

$$a_{i1}x_1 + ... + a_{in}x_n \le b_i$$

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* $D_1(\cdot), ..., D_n(\cdot)$ *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**

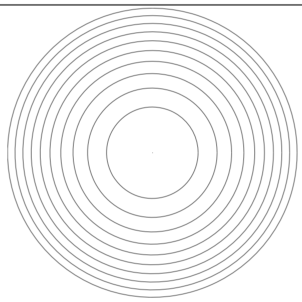
$$\min_{x \in \mathbb{R}^n, \tau} \left\{ \tau : -\tau \le D_*(\delta) - \sum_j x_j D_j(\delta) \le \tau \ \forall \delta \in \Gamma \right\}.$$

♠ **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:**

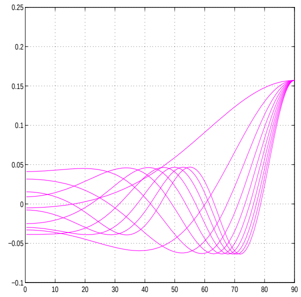
- **Diagram of a ring** $\{z = 0, a \leq \sqrt{x^2 + y^2} \leq b\}$:

$$D_{a,b}(\theta) = \frac{1}{2}\int_a^b \left[\int_0^{2\pi} r\cos\left(2\pi r\lambda^{-1}\cos(\theta)\cos(\phi)\right)d\phi\right]dr,$$

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



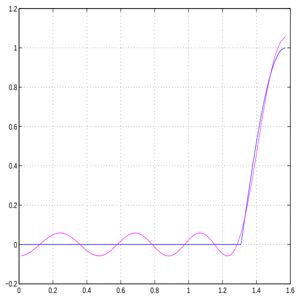
10 antenna elements, equal areas, outer radius 1 m



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

- **Nominal design problem:**

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



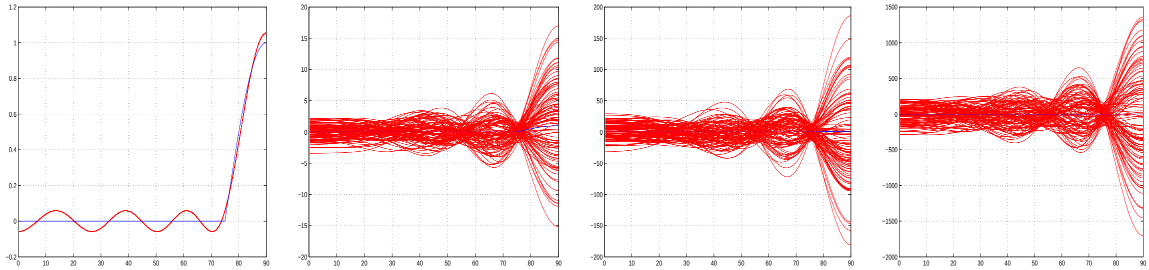
Target (blue) and nominal optimal (magenta) diagrams,

$\tau_* = 0.0589$

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:*

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

with small ρ?



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

"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`**:**

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

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 $\widetilde{a}^T x \geq b$ 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

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

- $a^{\rm n}_{ij}$: 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:**

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

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

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

♣ **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 $\mathcal{U}$.**

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

♣ **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:*

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

> **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.

$\Rightarrow$ *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*

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

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

$$\Downarrow$$

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

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

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

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.*

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

♣ **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

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

- **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**

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

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

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

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,

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

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

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

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

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

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*

$$\mathcal{X} = \left\{ x \in \mathbb{R}^N : \exists u \in \mathbb{R}^M : \left\{ \begin{array}{rcl} A_0 x + B_0 u + c_0 & = & 0 \\ A_1 x + B_1 u + c_1 & \in & \mathbf{K}_1 \\ & \dots & \\ A_K x + B_k u + c_k & \in & \mathbf{K}_K \end{array} \right. \right\},$$

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*

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

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

♣ **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*

$$\min_{x,u} \left\{ c^T x : \left\{ \begin{array}{rcl} A_0 x + B_0 u + c_0 & = & 0 \\ A_1 x + B_1 u + c_1 & \in & \mathbf{K}_1 \\ & \cdots & \\ A_K x + B_k u + c_k & \in & \mathbf{K}_K \end{array} \right\} \right\},$$

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.

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

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

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

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

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

$$\mathrm{Opt}(P)=\min_y\left\{\langle c,y\rangle:\left\{\begin{array}{l}A_0y-b_0=0\\ A_ky-b_k\in\mathbf{K}_k,\\ 1\leq k\leq K\end{array}\right\}\right\}, \qquad (P)$$

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*

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

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."*

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

$$\mathcal{Z} = \left\{ \zeta \in \mathbb{R}^L : \exists u \in \mathbb{R}^M : \left\{ \begin{array}{rcl} A_0\zeta + B_0u + c_0 & = & 0 \\ A_1\zeta + B_1u + c_1 & \in & \mathbf{K}_1 \\ & \cdots & \\ A_K\zeta + B_ku + c_k & \in & \mathbf{K}_K \end{array} \right. \right\} \qquad (!)$$

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

$$\begin{array}{rcl} \mathrm{Opt}(P) & := & \min\limits_{\zeta,u} \left\{ [-\mathcal{A}^T[x;1]]^T\zeta : \left\{ \begin{array}{rcl} A_0\zeta + B_0u + c_0 & = & 0 \\ A_1\zeta + B_1u + c_1 & \in & \mathbf{K}_1 \\ & \cdots & \\ A_K\zeta + B_ku + c_k & \in & \mathbf{K}_K \end{array} \right. \right\} \\ & \geq & \alpha_0^T[x;1] \end{array}$$

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**

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

$\square$

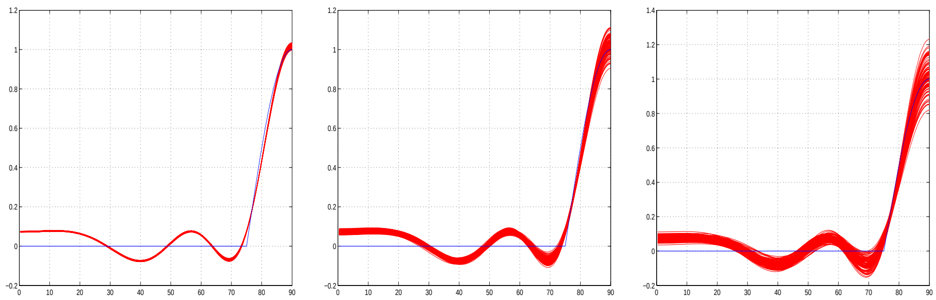
How it Works: Antenna Design

$$\min_{\tau,x}\left\{\tau : -\tau \leq D_*(\theta_i) - \sum_{j=1}^{10} x_j D_j(\theta_i) \leq \tau,\ 1 \leq i \leq I\right\}$$

$$x_j \mapsto (1+\zeta_j)x_j,\ -\rho \leq \zeta_j \leq \rho$$

$$\Rightarrow \min_{\tau,x}\left\{\tau : \begin{array}{l} D_*(\theta_i) - \sum_j x_j D_j(\theta_i) - \rho\sum\limits_j |x_j||D_j(\theta_i)| \geq -\tau \\ D_*(\theta_i) - \sum_j x_j D_j(\theta_i) + \rho\sum_j |x_j||D_j(\theta_i)| \leq \tau \end{array},\ 1 \leq i \leq I\right\} \quad \text{(RC)}$$

♠ **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).**



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

"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.*

Robust Linear Optimization
and
Chance Constraints

$$
\begin{array}{ll}
\left\{\alpha^T[x;1] \equiv a^T x + b \leq 0\right\},\ \alpha\text{: uncertain} & \textbf{(ULC)} \\
\Rightarrow \alpha^T[x;1] \leq 0\ \forall \alpha \in \mathcal{U} & \textbf{(RC)}
\end{array}
$$

♣ **Question:** *How to specify an uncertainty set?*

♠ **Answer:** *This is a modeling, heavily application-dependent, issue and as such it is beyond the scope of the RO theory.*

♠ **However:** *Sometimes we already have an uncertainty model, but a stochastic one rather than a model given in terms of an uncertainty/perturbation set.*

♠ **Claim:** *Given a stochastic uncertainty model, we can gain a lot by "translating" it into the RO paradigm.*

$$\left\{\alpha^T[x;1] \equiv a^T x + b \leq 0\right\}, \ \alpha\text{: uncertain} \qquad \textbf{(ULC)}$$

♣ **With the RO approach, we**

• *assume $\alpha = [a;b]$ to be affinely parameterized by a perturbation vector ζ:*

$$\alpha = \alpha_0 + \sum_{\ell=1}^{L} \zeta_\ell \alpha_\ell \qquad (*)$$

• *assume that ζ runs through a given perturbation set $\mathcal{Z} \subset \mathbb{R}^L$, and*

• *require from* (ULC) *to be valid for all realizations of α associated with $\zeta \in \mathcal{Z}$.*

The x's satisfying the latter requirement are treated as "uncertainty-immunized."

♠ **With the Chance Constrained Stochastic Optimization approach, we also assume** $(*)$**, but**

• *instead of specifying the range $\mathcal{Z}$ of ζ, treat ζ as a random variable with (partially) known distribution, and*

• *associate with* (ULC) *the chance constraint*

$$\mathrm{Prob}\left\{\zeta : \left[\alpha_0 + \sum_{\ell=1}^{L} \zeta_\ell \alpha_\ell\right]^T [x;1] > 0\right\} \leq \epsilon \qquad \text{(ChC)}$$

where $\epsilon \ll 1$ is a given tolerance.

The x's satisfying the latter requirement are treated as "uncertainty-immunized."

$$\left\{\alpha^T[x;1] \equiv a^T x + b \leq 0\right\},\ \alpha\text{: uncertain} \quad \textbf{(ULC)}$$

$$\Rightarrow \text{Prob}\left\{\zeta : \left[\alpha_0 + \sum_{\ell=1}^{L} \zeta_\ell \alpha_\ell\right]^T [x;1] > 0\right\} \leq \epsilon \qquad \text{(ChC)}$$

♣ Chance Constraints: pro & con

Good news on chance constraints: *ignoring the consequences of "rare events," our decision-making becomes less conservative than with the worst-case oriented RO approach.*

Bad news on chance constraints: *passing from an uncertain constraint* (ULC) *to its chance constrained version makes sense under four if's as follows:*

• **If** *there are reasons to believe that uncertain data indeed are of stochastic nature*, **which not always is the case E.g., when uncertainty comes from** *measurement errors*, **even those involving randomness, it perhaps makes sense to speak about** *distribution of nominal* **(measured)** *data, given the true data*, **but not about** *distribution of the true data, given the measurements.*

$$\left\{\alpha^T[x;1] \equiv a^T x + b \leq 0\right\},\ \alpha\text{: uncertain} \quad \textbf{(ULC)}$$

$$\Rightarrow \mathrm{Prob}\left\{\zeta : \left[\alpha_0 + \sum_{\ell=1}^{L} \zeta_\ell \alpha_\ell\right]^T [x;1] > 0\right\} \leq \epsilon \qquad \text{(ChC)}$$

• **If** *we are smart enough to identify the underlying data distribution.*

While the latter indeed is the case in some Engineering applications (Communications, Signal Processing, Control,...), it typically is *not* **the case in "decision-making proper." Given the "curse of dimensionality" when identifying multivariate distributions from historical data,** *assigning the uncertain data a particular distribution more often than not is an act of faith rather than a solid inference from the experimental data.*

• **If** *we are satisfied with probabilistic guarantees like "with such and such x, the probability of a disaster is $\leq$ 1.e-4 (or $\leq$ 1.e-8)"*

Probabilistic guarantees usually make sense if the situation repeats itself many times. Their attractiveness in the *single-outcome* **situation is much more problematic.**

• **If** *we are smart enough to process* (ChC) *in a computationally efficient manner.*

$$\left\{\alpha^T[x;1] \equiv a^T x + b \leq 0\right\},\ \alpha\text{: uncertain} \qquad \textbf{(ULC)}$$

$$\Rightarrow \left[\alpha_0 + \sum_{\ell=1}^{L} \zeta_\ell \alpha_\ell\right]^T [x;1] \leq 0 \qquad \text{(RC)}$$

$$\Rightarrow \text{Prob}\left\{\zeta : \left[\alpha_0 + \sum_{\ell=1}^{L} \zeta_\ell \alpha_\ell\right]^T [x;1] > 0\right\} \leq \epsilon \quad \text{(ChC)}$$

♠ **We have seen that (RC) is computationally tractable whenever the (convex) perturbation set $\mathcal{Z}$ is so.**

Unfortunately, there are no similar "general tractability results" for (ChC); **as a matter of fact,** *more often than not, chance constraints are computationally intractable.*

Reasons for intractability:

A. The Analysis problem associated with (ChC): *"given x, check whether x is feasible for* (ChC)*"* **usually is difficult: as a rule, the required probability is not available in a closed analytical form, while** *accurate* **numerical multi-dimensional integration is prohibitively time-consuming.**

Theorem [L. Khachiyan] *Consider the function*

$$vol(a) = \text{mes}_L\{\zeta \in \mathbb{R}^L : 0 \leq \zeta_\ell \leq 1\,\forall \ell, a^T\zeta \geq 1\}$$

where a is an integral vector. Unless P=NP, *no algorithm, given on input a and $\delta > 0$, is capable to compute* vol(a) *within accuracy δ in time polynomial in the bit length of a and in* $\ln(1/\delta)$.

$$\text{Prob}\left\{\zeta : \left[\alpha_0 + \sum_{\ell=1}^{L} \zeta_\ell \alpha_\ell\right]^T [x;1] > 0\right\} \le \epsilon \qquad \text{(ChC)}$$

♠ **Note: One can evaluate the probability in (ChC) by Monte Carlo simulation. However, the required sample size should be of order of $1/\epsilon$ and thus is prohibitively large for small ϵ, like $\epsilon = 1$.e-6 or $\epsilon = 1$.e-8.**

Question: Should we bother about small ϵ?

Answer: Sometimes this is a must. Think about

- **reliability of the steering mechanism in your car**
- **an LO problem with 10,000 randomly perturbed hard constraints**

B. The feasible set of (ChC) typically is non-convex, which makes problematic efficient minimization of linear objectives under (systems of) chance constraints.

Note: Essentially, the only known generic case where neither one of the above difficulties A, B occurs is the case of $\zeta \sim \mathcal{N}(\mu, \Sigma)$ and $\epsilon \le 1/2$. Here (ChC) is equivalent to

$$[\alpha_0 + \sum_{\ell=1}^{L} \mu_\ell \alpha_\ell]^T [x;1] + \text{ErfInv}(\epsilon)\sqrt{[x;1]^T A \Sigma A^T [x;1]} \le 0,$$

$$\bullet\ A = [\alpha_1, ..., \alpha_L] \qquad \bullet \int\limits_{\text{ErfInf}(r)}^{\infty} \frac{\mathrm{e}^{-s^2/2}}{\sqrt{2\pi}} ds \equiv r$$

$$\mathrm{Prob}\left\{\zeta : \left[\alpha_0 + \sum_{\ell=1}^{L}\zeta_\ell\alpha_\ell\right]^T [x;1] > 0\right\} \le \epsilon \qquad \text{(ChC)}$$

Note: **The body** $\left[\alpha_0 + \sum_{\ell=1}^{L}\zeta_\ell\alpha_\ell\right]^T [x;1]$ **of (ChC) can be rewritten as**

$$w_0[x] + \sum_{\ell=1}^{L}\zeta_\ell w_\ell[x]$$

where $w_0[x], ..., w_L[x]$ **are** *affine* **functions of** x**.**

♣ **When (ChC) "as it is" is computationally intractable, we can replace it with its** *safe tractable convex approximation* **defined as follows:**

Definition. *A safe convex approximation of the chance constraint*

$$\mathrm{Prob}\{w_0 + \sum_{\ell=1}^{L}\zeta_\ell w_\ell > 0\} \le \epsilon \qquad (\#)$$

in variables w is a convex subset W of the set of feasible solutions to (#).

Such an approximation is called tractable, if W is computationally tractable, i.e., it is given by an explicit finite system $\mathcal{S}$ of efficiently computable convex constraints $f_i(w,v) \le 0, i \le I$ in variables w and additional variables v.

$$\mathrm{Prob}\Big\{\zeta : \overbrace{\left[\alpha_0 + \sum\nolimits_{\ell=1}^{L}\zeta_\ell\alpha_\ell\right]^T [x;1]}^{\equiv w_0[x]+\sum_{\ell=1}^{m}\zeta_\ell w_\ell[x]} > 0\Big\} \leq \epsilon \quad \text{(ChC)}$$
$$\Rightarrow \mathrm{Prob}\{w_0 + \sum\nolimits_{\ell=1}^{L}\zeta_\ell w_\ell > 0\} \leq \epsilon \qquad (\#)$$

Note: *Given a safe tractable convex approximation* $f_i(w,v) \leq 0, i \leq I$ *of* $(\#)$*, the system of explicit efficiently computable convex constraints*

$$g_i(x,v) \equiv f_i(w[x],v) \leq 0,\ i \leq I$$

in variables x, v *possesses the following properties:*

- [tractability] *it is computationally tractable*
- [safety] *whenever* x *can be extended to a feasible solution to the system,* x *is feasible for the chance constraint* (ChC).

Conclusion: Given a Chance Constrained LO problem with certain objective and replacing every chance constraint with its safe convex tractable approximation, we end up with an efficiently solvable convex optimization problem which is a safe approximation of the problem of interest: *every feasible solution to the approximation is feasible for the chance constrained problem.*

Safe Tractable Approximations of Scalar Chance Constraints and Robust Optimization

$$\mathrm{Prob}\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell > 0\} \le \epsilon \qquad (\#)$$

♣ *From now on, let ζ_i be with finite means.*

Observation: The feasible set W_+ of (#) is

- *conic*: $w \in W_+, t \ge 0 \Rightarrow tw \in W_+$
- *closed*
- *possesses a nonempty interior*, **specifically,**

$$e := [-1; 0; ...; 0] \in \mathrm{int} W_+.$$

Definition: *A safe convex approximation of* (#) **(i.e. a convex set** $W \subset W_+$**)** *is called normal, if it inherits the above properties of W_+, that is, it is conic, closed and* $e \in \mathrm{int} W$.

Conclusion: *A normal safe convex approximation of* (#) *is a closed convex cone $W \subset \mathbb{R}^{L+1}$ which is contained in the feasible set of* (#) *and is such that* $e \in \mathrm{int} W$.

$$\text{Prob}\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell > 0\} \leq \epsilon \qquad (\#)$$

Conclusion: *A safe convex approximation of* $(\#)$ *is a closed convex cone* $W \subset \mathbb{R}^{L+1}$ *which is contained in the feasible set of* $(\#)$ *and is such that* $e = [-1; 0; ...; 0] \in \text{int} W$.

Facts: Let $\mathbf{K} \subset \mathbb{R}^{L+1}$ be a cone.

• **K is the dual of another cone, specifically, of its dual cone $\mathbf{K}_* = \{v \in \mathbb{R}^{L+1} : v^T w \geq 0 \,\forall w \in \mathbf{K}\}$:**

$$\mathbf{K} = \{w \in \mathbb{R}^{L+1} : w^T v \geq 0 \,\forall v \in \mathbf{K}_*\}$$

• **If $f \in \text{int}\mathbf{K}$, then $V = \{v \in \mathbf{K}_* : f^T v = 1\}$ is a convex compact set and**

$$\mathbf{K} = \{w \in \mathbb{R}^{L+1} : v^T w \geq 0 \;\forall v \in V\}$$

♠ **Let W be a normal safe convex approximation of $(\#)$. Applying Facts to W in the role of $\mathbf{K}$ and e in the role of f, we get the following:**

There exists a convex compact set $V \subset \mathbb{R}^{L+1}$, *specifically, the set*

$$\begin{aligned} V &= \{v = [v_0; v_1; ...; v_L] \in W_* : [-1; 0; ...; 0]^T v = 1\} \\ &= \{v = [-1; v_1; ...; v_L] \in W_*\} = \{[-1; -z] : z \in \mathcal{Z} \subset \mathbb{R}^L\} \end{aligned}$$

such that

$$\begin{aligned} W &= \{w \in \mathbb{R}^{L+1} : v^T w \geq 0 \,\forall v \in V\} \\ &= \{[w_0; ...; w_L] \in \mathbb{R}^{L+1} : w_0 + \sum_{\ell=1}^{L} z_\ell w_\ell \leq 0 \,\forall z \in \mathcal{Z}\} \end{aligned}$$

$$\mathrm{Prob}\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell > 0\} \leq \epsilon \qquad (\#)$$

♠ **We have arrived at the following**

Proposition: *Every normal safe convex approximation W of (#) is of the Robust Counterpart form: there exists a convex compact set $\mathcal{Z} \subset \mathbb{R}^L$ such that*

$$W = \{[w_0; w_1; ...; w_L] : w_0 + \sum_{\ell=1}^{L} z_L w_L \leq 0 \ \forall z \in \mathcal{Z}\}$$

that is, W is the set of robust feasible solutions to the uncertainty-affected inequality

$$w_0 + \sum_{\ell=1}^{L} z_\ell w_\ell \leq 0,$$

in variables w, the uncertain coefficients being $z_1, ..., z_L$, and the perturbation set being $\mathcal{Z}$.

On a closest inspection,

The convex compact set $\mathcal{Z}$ associated with a normal safe convex approximation of (#) is computationally tractable whenever the approximation itself is so.

♣ **For the time being, we were speaking on the chance constraints of the form**

$$\mathrm{Prob}\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell > 0\} \leq \epsilon$$

tacitly assuming that *the probability distribution of ζ is fixed.* **In reality, more often than not the probability distribution P of ζ is only** *partially known* **— all we know is that** *P belongs to a given family $\mathcal{P}$ of probability distributions on $\mathbb{R}^L$*. **Whenever this is the case, it is natural to associate with randomly perturbed constraint its** *ambiguously chance constrained version*

$$\forall P \in \mathcal{P} : \mathrm{Prob}_{\zeta\sim P}\left\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell > 0\right\} \leq \epsilon \qquad (!)$$

Assuming from now on that ζ is "$\mathcal{P}$ uniformly summable:"

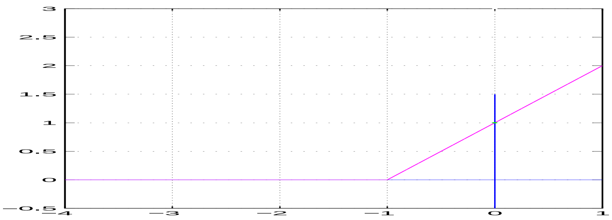
$$\sup_{P\in\mathcal{P}} \mathbf{E}_{\zeta\sim P}\{\|\zeta\|\} < \infty,$$

the notions of normal/safe/convex/tractable approximation of a "usual" chance constraint word by word extend to the case of ambiguous chance constraint, and Proposition extends on this case as well.

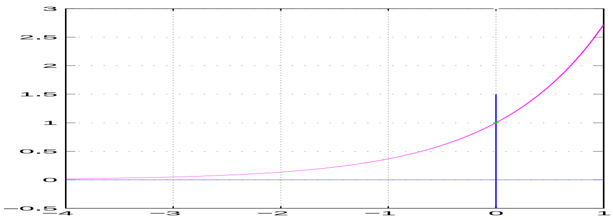
$$p(w) := \sup_{P\in\mathcal{P}} \mathrm{Prob}_{\zeta\sim P}\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell > 0\} \le \epsilon \qquad (!)$$

Generating-Function-Based Safe Convex Approximation of Chance Constraint

♣ **Definition:** *A generator is a convex function $\gamma(\cdot)$ on the axis such that $\gamma(s) \to 0$ as $s \to -\infty$ and $\gamma(0) \ge 1$.*



$\gamma(s) = \max[1+s, 0]$



$\gamma(s) = \exp\{s\}$

♠ **Observation: Let $\gamma(\cdot)$ be a generator. Then the function**

$$\Psi_P(w) = \mathbf{E}_{\zeta\sim P}\left\{\gamma(w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell)\right\} : \mathbb{R}^{L+1} \to \mathbb{R}\cup\{+\infty\}$$

is convex and lower semi-continuous (l.s.-c.), and this function is an upper bound on $p(w)$. Consequently, *the function*

$$\Psi(w) = \sup_{P\in\mathcal{P}} \Psi_P(w)$$

is a convex l.s.-c. upper bound on $p(w)$.

♠ **Further, we clearly have $\alpha > 0 \Rightarrow p(w) = p(w/\alpha)$, whence $\alpha > 0 \Rightarrow \Psi(w/\alpha) \ge p(w)\ \forall w$.**

$$p(w) := \sup_{P\in\mathcal{P}} \mathrm{Prob}_{\zeta\sim P}\{w_0 + \sum_{\ell=1}^L \zeta_\ell w_\ell > 0\} \leq \epsilon \qquad (!)$$

$$\gamma(\cdot) : \mathbb{R} \to \mathbb{R}_+ : \ \text{convex}, \gamma(-\infty) = 0, \gamma(0) \geq 1$$

$$\Rightarrow \Psi(w) := \sup_{P\in\mathcal{P}} \mathbf{E}_{\zeta\sim P}\left\{\gamma(w_0 + \sum_{\ell=1}^L \zeta_\ell w_\ell)\right\} \geq p(w) \ (+)$$

Corollary. *The system*

$$\alpha\Psi(w/\alpha) - \alpha\epsilon \leq 0, \ \alpha > 0 \qquad (S)$$

in variables w, α is a system of convex constraints which is a safe convex approximation of $(!)$: whenever w can be extended by properly chosen α to a feasible solution of (S), w is feasible for $(!)$.

Proof. $\Psi(w)$ is convex by its origin, which, by a well-known fact of Convex Analysis, implies that $\alpha\Psi(w/\alpha)$ is convex in w, α in the domain $\alpha > 0$. Thus, (S) is indeed a system of convex constraints in variables w, α.

Now let (w, α) be feasible for (S). Then $\Psi(w/\alpha) \leq \epsilon$, whence $p(w/\alpha) \leq \epsilon$ as well. Since $p(w) = p(w/\alpha) \leq \epsilon$, w is feasible for $(!)$, Q.E.D.

♠ A simple technical exercise allows to strengthen Corollary to the following

Proposition. *The convex constraint*

$$G(w) := \inf_{\alpha>0} \{\alpha\Psi(w/\alpha) - \alpha\epsilon\} \leq 0$$

is a safe convex approximation of $(!)$.

$$p(w) := \sup_{P\in\mathcal{P}} \mathrm{Prob}_{\zeta\sim P}\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell > 0\} \leq \epsilon \qquad (!)$$

$$\gamma(\cdot) : \mathbb{R} \to \mathbb{R}_+ : \text{ convex, } \gamma(0) \geq 1,\ \gamma(-\infty) = 0$$

$$\Rightarrow \Psi(w) := \sup_{P\in\mathcal{P}} \mathbf{E}_{\zeta\sim P}\left\{\gamma(w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell)\right\} \geq p(w) \quad (+)$$

Corollary of Proposition: *Let $\Psi^+(\cdot)$ be a convex upper bound on $\Psi(\cdot)$. Then the system of convex constraints in variables w, α*

$$\alpha\Psi^+(w/\alpha) - \alpha\epsilon \leq 0,\ \alpha > 0 \qquad (S)$$

and the convex constraint

$$G^+(w) := \inf_{\alpha>0} \left\{\alpha\Psi^+(w/\alpha) - \alpha\epsilon\right\} \leq 0$$

are safe convex approximations of $(!)$.

♣ **Example:** Assume that ζ_ℓ are known to be independent zero mean and taking values in $[-1,1]$, **or, equivalently,** *$\mathcal{P}$ is comprised of product-type distributions P with zero mean marginals P_ℓ supported on $[-1,1]$.*

♠ **Let us apply the above approximation scheme with the generator** $\gamma(s) = \exp\{s\}$**. When** $P = P_1 \times \ldots \times P_L \in \mathcal{P}$**, we have**

$$\begin{array}{l}\Psi_P(w) = \mathbf{E}_{\zeta\sim P_1\times\ldots\times P_L}\{\exp\{w_0 + \sum_{\ell=1}^L \zeta_\ell w_\ell\}\}\\ = \exp\{w_0\}\prod\limits_{\ell=1}^L \mathbf{E}_{\zeta_\ell\sim P_\ell}\{\exp\{\zeta_\ell w_\ell\}\}.\end{array}$$

Lemma: *Let Q be a zero mean distribution supported on $[-1,1]$. Then*

$$\int \exp\{ts\}dQ(s) \leq \cosh(t) \leq \exp\{t^2/2\}. \qquad (*)$$

Proof. The inequality $\cosh(t) \leq \exp\{t^2/2\}$ **is evident. To prove that** $\int \exp\{ts\}dQ(s) \leq \cosh(t)$**, let** $f(s) = \exp\{ts\} - s\sinh(t)$**. Since** Q **is with zero mean and is supported on** $[-1,1]$**, we have**

$$\begin{array}{l}\int \exp\{ts\}dQ(s) = \int f(s)dQ(s)\\ \leq \max\limits_{-1\leq s\leq 1} f(s) = \max\limits_{s=\pm 1} f(s) = \cosh(t).\end{array}$$

Note: *When Q is uniform on $\{-1;1\}$, the first inequality in $(*)$ becomes equality.*

♠ **We see that we are in the situation**

$$\begin{array}{l}\Psi(w) = \exp\{w_0\} \prod\limits_{\ell=1}^{L} \cosh(w_\ell) \\ \leq \Psi^+(w) := \exp\{w_0 + \frac{1}{2}\sum_{\ell=1}^{L} w_\ell^2\}.\end{array}$$

$\Rightarrow$ **For our $\mathcal{P}$, the safe convex approximation of the chance constraint**

$$\forall (P \in \mathcal{P}) : \mathrm{Prob}_{\zeta\sim P}\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell > 0\} \leq \epsilon$$

reads

$$\inf_{\alpha>0} \underbrace{\alpha \left[\exp\{\alpha^{-1}w_0 + \frac{\alpha^{-2}}{2}\sum_{\ell=1}^{L} w_\ell^2\} - \epsilon\right]}_{f(\alpha)} \leq 0. \qquad (*)$$

♡ **Assuming** $w_1^2 + ... + w_L^2 > 0$**, we have** $f(\alpha) \to \infty$ **as** $\alpha \to +\infty$ **and as** $\alpha \to +0$

$\Rightarrow$ $(*)$ **is equivalent to** $\exists \alpha > 0 : \alpha^{-1}w_0 + \frac{\alpha^{-2}}{2} \sum\limits_{\ell=1}^{L} \omega_\ell^2 \leq \ln(\epsilon)$**, or, which is the same, to**

$$w_0 + \sqrt{2\ln(1/\epsilon)}\sqrt{\sum_{\ell=1}^{L} w_\ell^2} \leq 0. \qquad (+)$$

♡ **When** $w_1 = ... = w_L = 0$**,** $(*)$ **also is equivalent to** $(+)$**.**

♠ **Bottom line:** *When ζ_ℓ are independent zero mean random variables taking values in $[-1,1]$, the conic quadratic inequality*

$$w_0 + \sqrt{2\ln(1/\epsilon)}\sqrt{\sum_{\ell=1}^{L} w_\ell^2} \leq 0. \qquad (+)$$

is a safe tractable approximation of the chance constraint

$$\mathrm{Prob}\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell > 0\} \leq \epsilon$$

♠ **Observation:** *The constraint $(+)$ is the RC*

$$w_0 + \sum_{\ell=1}^{L} a_\ell w_\ell \leq 0 \ \forall a \in \mathcal{U}$$

of the uncertain constraint

$$w_0 + \sum_{\ell=1}^{L} a_\ell w_\ell \leq 0$$

in variables w, with the ball

$$\mathcal{U} = \{a : \|a\|_2 \leq \sqrt{2\ln(1/\epsilon)}\}$$

in the role of the uncertainty set.

$$\forall (P \in \mathcal{P}) : \mathrm{Prob}_{\zeta \sim P}\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell > 0\} \le \epsilon \qquad (!)$$

♣ **Discussion.** **The simplest RO-based way to safely approximate (!) is as follows:**

• **We choose as the uncertainty set $\mathcal{U}$ a convex compact set $\mathcal{U}^\epsilon$ which** "$(1-\epsilon)$-*supports*" **all distributions from $\mathcal{P}$:**

$$\forall (P \in \mathcal{P}) : \mathrm{Prob}_{\zeta \sim P}\{\zeta \notin \mathcal{U}^\epsilon\} \le \epsilon.$$

• **We set $\mathcal{U} = \mathcal{U}^\epsilon$, thus ensuring that**

$$\begin{array}{c} w_0 + \sum_{\ell=1}^{L} a_\ell w_\ell \le 0 \; \forall a \in \mathcal{U}^\epsilon \\ \Rightarrow \forall (P \in \mathcal{P}) : \mathrm{Prob}_{\zeta \sim P}\left\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell > 0\right\} \le \epsilon. \end{array}$$

♠ **Note:** The perturbation set

$$\mathcal{U} = \{a : \|a\|_2 \leq \Omega := \sqrt{2\ln(1/\epsilon)}\}$$

yielded by our approximation scheme is, for ϵ fixed and L large, *incomparably smaller* than any $(1-\epsilon)$-support $\mathcal{U}^\epsilon$ of the distributions from $\mathcal{P}$.

For example, with $\epsilon = 1.\text{e-}6$ **we have** $\Omega = 5.26$**, and**

- **When** $L > \Omega^2 = 27.63$ **and** P **is the uniform distribution on the vertices of** $[-1,1]^L$**, we have** $P\{\zeta \in \mathcal{U}\} = 0$**;**
- **When** $\epsilon < 1/2$**, the Euclidean diameter of every** $(1-\epsilon)$**-support of the distributions from** $\mathcal{P}$ **is at least** $2\sqrt{L}$**, while the Euclidean diameter of** $\mathcal{U}$ **is just** $2\Omega \approx 10.51$**;**
- **The ratio of the volume of an** $(1-\epsilon)$**-support of distributions from** $\mathcal{P}$ **to the volume of** $\mathcal{U}$ **exponentially grows with** L **when** $L \geq 60$**; for** $L = 256$ **this ratio is as large as** $2 \cdot 10^{44}$**.**

The Least Conservative Implementation: Conditional Value At Risk

$$\forall (P \in \mathcal{P}) : \mathrm{Prob}_{\zeta\sim P}\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell > 0\} \leq \epsilon \qquad (!)$$

$$\gamma(\cdot) : \text{ convex}, \gamma(-\infty) = 0, \gamma(0) \geq 1$$

$$\Rightarrow \Psi(w) = \sup_{P\in\mathcal{P}} \mathbf{E}_{\zeta\sim P}\left\{\gamma\left(w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell\right)\right\}$$

$$\Rightarrow G(w) := \inf_{\alpha>0} \alpha[\Psi(w/\alpha) - \epsilon] \leq 0 \qquad \text{(Appr)}$$

♣ **Question:** *What is the best choice of* $\gamma(\cdot)$ *?*

♠ **Answer:** *As far as the conservatism of* (Appr) *is concerned, the best choice of* $\gamma(\cdot)$ *is*

$$\gamma(s) = \max[1 + s, 0]$$

$$[\textit{or } \gamma(s) = \max[1 + \alpha s, 0] \textit{ with } \alpha > 0]$$

Indeed, when $\gamma(0) > 1$**, the conservatism is reduced by**

$$\gamma(\cdot) \leftarrow \gamma(\cdot)/\gamma(1).$$

Assuming $\gamma(0) = 1$ **and setting** $\alpha = \gamma'(+0)$**, we get** $\alpha > 0$ **and** $\gamma(s) \geq 1 + \alpha s$**. Since** $\gamma(s) \geq 0$ **(as a generator), we get** $\gamma(s) \geq \bar{\gamma}(s) = \max[1+\alpha s, 0]$**.** $\bar{\gamma}(\cdot)$ **is a legitimate generator, and since** $\gamma(\cdot) \geq \bar{\gamma}(\cdot)$**, passing from** γ **to** $\bar{\gamma}$ **can only reduce the conservatism of** (Appr)**.**

$$\forall(P\in\mathcal{P}):\mathrm{Prob}_{\zeta\sim P}\{w_0+\sum_{\ell=1}^L\zeta_\ell w_\ell>0\}\leq\epsilon \qquad (!)$$

$$\gamma(\cdot):\ \text{convex},\ \gamma(-\infty)=0,\gamma(0)\geq 1$$

$$\Rightarrow \Psi(w)=\sup_{P\in\mathcal{P}}\mathbf{E}_{\zeta\sim P}\left\{\gamma\left(w_0+\sum_{\ell=1}^L\zeta_\ell w_\ell\right)\right\}$$

$$\Rightarrow G(w):=\inf_{\alpha>0}\alpha[\Psi(w/\alpha)-\epsilon]\leq 0 \qquad (\mathrm{Appr})$$

♠ **On a closest inspection, when $\mathcal{P}=\{P\}$ the approximation** (Appr) **associated with** $\gamma(s)=\max[1+s,0]$ **reads**

$$\underbrace{\min_a\left[a+\frac{1}{\epsilon}\mathbf{E}\left\{\max[w_0-a+\sum_{\ell=1}^L\zeta_\ell w_\ell,0]\right\}\right]}_{=:\mathrm{CVaR}_\epsilon[w_0+\sum_{\ell=1}^L\zeta_\ell w_\ell]}\leq 0$$

This is the well-known *Conditional Value at Risk* **safe convex approximation of** (!) **originating from Rockafellar et al.**

♠ **Bad news on the** CVaR **approximation:** *This approximation typically is computationally intractable.*

Two basic exceptions are:

• $\zeta\sim\mathcal{N}(a,Q)$ — **of no interest: here** (!) **by itself is an explicit convex constraint on** w, **and no approximations are necessary**

• ζ **takes moderately many values** $\zeta^1,...,\zeta^N$ **with known probabilities** $\pi_1,...,\pi_N$.

♠ **When ζ takes moderately many values $\zeta^1,...,\zeta^N$ with known probabilities $\pi_1,...,\pi_N$, the CVaR approximation of the chance constraint reads**

$$\min_a \left\{ a + \frac{1}{\epsilon} \sum_i \pi_i \max\left[w_0 - a + \sum_{\ell=1}^{L} \zeta_\ell^i w_\ell, 0 \right] \right\} \leq 0$$

♡ **The RC form of this approximation is**

$$w_0 + \sum_{\ell=1}^{L} a_\ell w_\ell \leq 0 \ \forall a \in \mathcal{U} = \left\{ \sum_{i=1}^{N} u_i \zeta^i : \begin{array}{c} 0 \leq u_i \leq \frac{\pi_i}{\epsilon} \\ \sum_i u_i = 1 \end{array} \right\}$$

"Tractable Case": Bernstein Approximation

$$\forall (P \in \mathcal{P}) : \text{Prob}_{\zeta\sim P}\{w_0 + \sum\nolimits_{\ell=1}^{L} \zeta_\ell w_\ell > 0\} \leq \epsilon \qquad (!)$$

♣ **Assume that**

Brn.1 Random perturbations $\zeta_1, ..., \zeta_L$ *are independent* **and their distributions** P_ℓ **belong to given families** $\mathcal{P}_\ell$ **of probability distributions on the axis;**

Brn.2 We can point out *convex l.s.-c. upper bounds* Φ^+_ℓ **on logarithmic moment-generating functions of distributions from** $\mathcal{P}_\ell$**:**

$$\Phi^+_\ell(t) \geq \ln\left(\mathbf{E}_{\zeta_\ell\sim P_\ell}\{\exp\{t\zeta_\ell\}\}\right) \quad \forall P_\ell \in \mathcal{P}_\ell$$
$$\Rightarrow \Phi^+(w) = \sum\nolimits_{\ell=1}^{L} \Phi^+_\ell(\omega_\ell)$$

♠ **Taking** $\gamma(s) = \exp\{s\}$ **and applying our approximation scheme in logarithmic scale, we arrive at**

Theorem: *In the case of Brn.1-2,* $(!)$ *admits safe convex approximation given by the convex constraint*

$$H^+(w) := \inf_{\alpha>0} \alpha\left[w_0/\alpha + \Phi^+(w/\alpha) + \ln(1/\epsilon)\right] \leq 0$$

This approximation is tractable, provided that the l.s.-c. convex functions $\Phi^+_\ell(\cdot)$ *are efficiently computable. Assuming* $0 \in \text{intDom}\,\Phi^+$*, the approximation is of the RC form:*

$$H^+(w) \leq 0 \Leftrightarrow w_0 + a^T[w_1; ...; w_l] \leq 0\, \forall a \in \mathcal{U},$$

where $\mathcal{U}$ *is a nonempty convex compact set. When*

$$\Phi^+(w) = \sup_z \left[[w_1; ...; w_L]^T[Bz + b] - \phi(z)\right]$$

for a l.s.-c. convex function $\phi(\cdot)$ *with bounded level sets, one can take*

$$\mathcal{U} = \{Bz + b : \phi(z) \leq \ln(1/\epsilon)\}.$$

$$\forall (P \in \mathcal{P}) : \mathrm{Prob}_{\zeta\sim P}\left\{w_0 + \sum_{\ell=1}^L \zeta_\ell w_\ell > 0\right\} \leq \epsilon \quad (!)$$

Example 1: *$\zeta_\ell \sim \mathcal{N}(\mu_\ell, \sigma_\ell^2)$ are independent; we know lower and upper bounds $\underline{\mu}_\ell$, $\overline{\mu}_\ell$ on μ_ℓ and upper bounds $\overline{\sigma}_\ell$ on σ_ℓ.*

Building Bernstein approximation of (!):

$$\begin{array}{rl}
& \mathcal{P}_\ell = \{\mathcal{N}(\mu, \sigma^2) : \underline{\mu}_\ell \leq \mu \leq \overline{\mu}_\ell, \sigma \leq \overline{\sigma}_\ell\} \\
\Rightarrow & \Phi_\ell^+(t) := \sup\limits_{P_\ell \in \mathcal{P}_\ell} \ln\left(\mathbf{E}_{\zeta_\ell \sim P_\ell}\{\exp\{t\zeta_\ell\}\}\right) \\
& = \max[\underline{\mu}_\ell t, \overline{\mu}_\ell t] + \overline{\sigma}_\ell^2 t^2/2 \\
& = \max\limits_{0\leq\lambda\leq1}\left\{[\lambda\underline{\mu}_\ell + (1-\lambda)\overline{\mu}_\ell]t + \overline{\sigma}_\ell^2 t^2/2\right\} \\
& = \max\limits_{0\leq\lambda\leq1}\max\limits_{v}\left\{vt - [\lambda\underline{\mu}_\ell + (1-\lambda)\overline{\mu}_\ell - v]^2/(2\overline{\sigma}^2)\right\} \\
& = \max\limits_{v}\left\{vt - \min_{0\leq\lambda\leq1}[\lambda\underline{\mu}_\ell + (1-\lambda)\overline{\mu}_\ell - v]^2/(2\overline{\sigma}^2)\right\} \\
& = \max\limits_{v}\left\{vt - \mathrm{dist}^2(v, [\underline{\mu}_\ell, \overline{\mu}_\ell])/(2\overline{\sigma}_\ell^2)\right\}
\end{array}$$

$\Rightarrow$ **The Bernstein approximation is**

$$w_0 + \sum_{\ell=1}^L \max[\underline{\mu}_\ell w_\ell, \overline{\mu}_\ell w_\ell] + \sqrt{2\ln(1/\epsilon)}\sqrt{\sum_{\ell=1}^L \overline{\sigma}_\ell^2 w_\ell^2} \leq 0$$

♠ *The uncertainty set $\mathcal{U}$ participating in the RC representation of the Bernstein approximation is*

$$\mathcal{U} = \{a \in \mathbb{R}^L : \sum_{\ell=1}^L \mathrm{dist}^2(a_\ell, [\underline{\mu}_\ell, \overline{\mu}_\ell])/(2\overline{\sigma}_\ell^2) \leq \ln(1/\epsilon)\},$$

which is the arithmetic sum of the box $\{\underline{\mu} \leq a \leq \overline{\mu}\}$ and the ellipsoid $\left\{a : \sqrt{\sum_{\ell=1}^L a_\ell^2/\overline{\sigma}_\ell^2} \leq \sqrt{2\ln(1/\epsilon)}\right\}$.

♡ $\underline{\mu}_\ell = \overline{\mu}_\ell$ *for all $\ell \Rightarrow$ the uncertainty set is ellipsoid.*

$$\forall(P\in\mathcal{P}):\ \mathrm{Prob}_{\zeta\sim P}\left\{w_0+\sum_{\ell=1}^L\zeta_\ell w_\ell>0\right\}\leq\epsilon\qquad(!)$$

Example 2: *ζ_ℓ are independent, supported on $[-1,1]$ and with given bounds on expectations:* $\mathbf{E}\{\zeta_i\}\in[\mu_i^-,\mu_i^+]$ (by "scaling" $\zeta_\ell\leftarrow a_\ell+b_\ell\zeta_\ell$ this covers the case of independent ζ_ℓ with partially known means and known finite ranges).

♣ **The uncertainty set $\mathcal{U}$ associated with the RC form of Bernstein approximation of (!) is**

$$\mathcal{U}^{\mathrm{Entropy}}=\left\{\zeta:\sum_{\ell=1}^L\phi_\ell(\zeta_\ell)\leq 2\ln(1/\epsilon)\right\}$$

$$\phi_\ell(s)=\begin{cases}(1+s)\ln\left(\frac{1+s}{1+\mu_\ell^-}\right)+(1-s)\ln\left(\frac{1-s}{1-\mu_\ell^-}\right) & ,-1\leq s\leq\mu_\ell^-\\ 0 & ,\mu_\ell^-\leq s\leq\mu_\ell^+\\ (1+s)\ln\left(\frac{1+s}{1+\mu_\ell^+}\right)+(1-s)\ln\left(\frac{1-s}{1-\mu_\ell^+}\right) & ,\mu_\ell^+\leq s\leq 1\\ +\infty & ,|s|>1\end{cases}$$

• *The Entropy approximation of (!) is given by a system of explicit efficiently computable convex constraints.*

♣ **The Entropy uncertainty can be extended to the** *Ball-Box* **one:**

$$\mathcal{U}^{\mathrm{BB}}=\left[\{\mu^-\leq\zeta\leq\mu^+\}+\left\{\|\zeta\|_2\leq\sqrt{2\ln(1/\epsilon)}\right\}\right]\bigcap\{\|\zeta\|_\infty\leq 1\}$$

• *The BallBox approximation of (!) is given by a system of conic quadratic constraints.*

♣ **The BallBox uncertainty can be extended to** *Budgeted uncertainty*

$$\mathcal{U}^{\mathrm{Bdg}}=\left[\{\mu^-\leq\zeta\leq\mu^+\}+\left\{\|\zeta\|_1\leq\sqrt{2\ln(1/\epsilon)m}\right\}\right]\bigcap\{\|\zeta\|_\infty\leq 1\}$$

• *The Budgeted approximation of (!) is given by a system of linear constraints.*

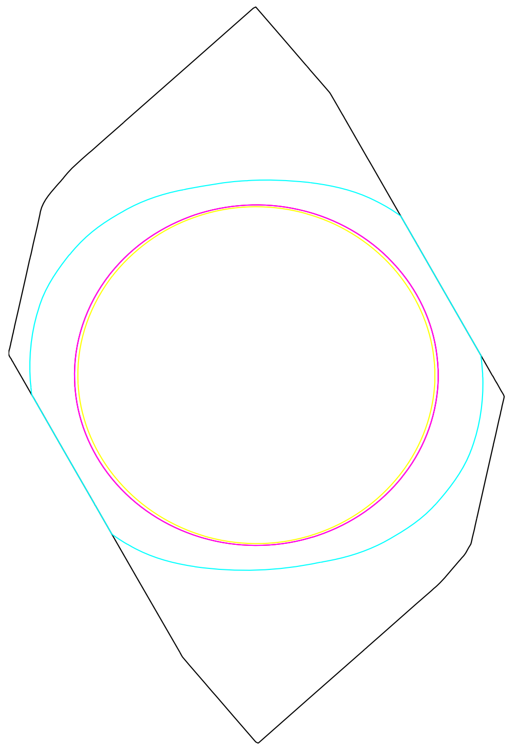
$$\forall (P \in \mathcal{P}) : \mathrm{Prob}_{\zeta\sim P}\left\{ w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell > 0 \right\} \leq \epsilon \qquad (!)$$

♣ **Finally, the Budgeted uncertainty can be extended to** *Box uncertainty*

$$\mathcal{U}^{\mathrm{Box}} = \{\zeta : \|\zeta\|_\infty \leq 1\}$$

which ignores any information on the distribution of ζ aside of its support.

• *The Box approximation of (!) is given by a system of linear constraints.*



Example 2: Random 2D central cross-sections of perturbation sets corresponding to various approximations, $L = 256$

- **black: Box approximation**
- **cyan: Budgeted approximation**
- **magenta: Ball-Box approximation**
- **yellow: Entropy approximation**

Illustration: Portfolio selection

There are $L = 200$ assets with independent random yearly returns r_ℓ. It is known that

- For $\ell \leq 199$, return r_ℓ has expectation $\mu_\ell = 1.05 + 0.3\frac{200-\ell}{199}$ and varies in $[\mu_\ell - \sigma_\ell, \mu_\ell + \sigma_\ell]$, $\sigma_\ell = 0.05 + 0.6\frac{200-\ell}{199}$;
- For $\ell = 200$, $r_\ell \equiv 1.05$ ["money in the bank"].

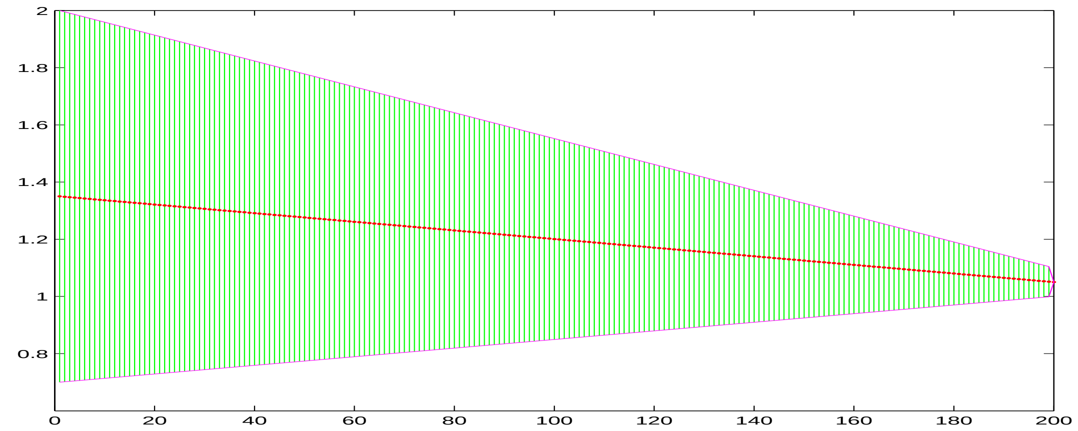
We want to distribute \$1 between the assets in order to maximize the Value-at-0.5%-Risk of the portfolio in a year from now, that is, we want to solve the chance constrained problem

$$\max_{t,x}\left\{t : \begin{array}{l}\mathrm{Prob}\left\{\sum_{\ell=1}^{L} r_\ell x_\ell < t\right\} \leq \epsilon = 0.005\\ x \geq 0, \sum_\ell x_\ell = 1\end{array}\right\}$$

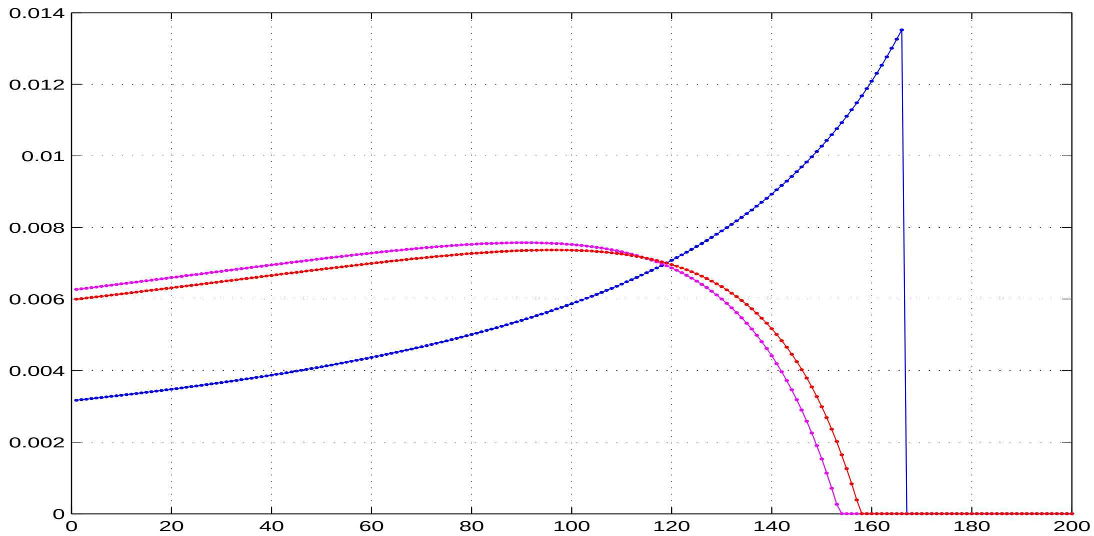
♠ **Setting $r_\ell = \mu_\ell + \sigma_\ell\zeta_\ell$, the random variables ζ_ℓ are independent zero mean and take values in $[-1,1]$. The problem of interest now reads**

$$\max_{t,x}\left\{t : \begin{array}{l}x \geq 0, \sum_\ell x_\ell = 1\\ \mathrm{Prob}\{w_0[x,t] + \sum_{\ell=1}^{n}\zeta_\ell w_\ell[x,t] > 0\} \leq \epsilon\end{array}\right\},$$
$$w_0[x,t] = t - \sum_{\ell=1}^{L}\mu_\ell x_\ell,\ w_\ell[x,t] = -\sigma_\ell x_\ell$$

Replacing the chance constraint with its safe tractable approximation and solving the resulting convex program, we get a *feasible* **suboptimal solution to the problem of interest.**



Means and ranges of returns



Distributions of \$1 between assets

- **blue: Budgeted approximation**
- **magenta: Ball-Box approximation**
- **red: Entropy approximation**
- **The Box approximation (not shown) leads to the solution "keep all money in the bank."**

Approx.	Box	Budgeted	Ball-Box	Entropy
Opt. Val.	1.0500	1.1012	1.1200	1.1209

Beyond the Scope of Affinely Perturbed Chance Constraints with Independent Perturbations: Lagrangian Relaxation of Chance Constraint

♣ **As far as *tractable* approximations are concerned, our construction imposes on ζ_ℓ the requirements to *be independent* and to *enter affinely* the body of the chance constraint.**

Consider a different situation, where

• **the random perturbations ζ_ℓ enter the body of chance constraint *quadratically* (decision variables still enter it linearly), and**

• **we have *partial* information on the *marginal distributions* P_ℓ of ζ_ℓ, on their *covariances*, and on the *domain* of ζ, but no more than that.**

♠ **Thus, from now on our chance constraint is**

$$\begin{array}{l}\forall (P \in \mathcal{P}):\\ \mathrm{Prob}_{\zeta\sim P}\{\underbrace{\zeta^T U[x]\zeta + 2\zeta^T v[x] + w[x]}_{=\mathrm{Tr}(W[x]Z[\zeta])} > 0\} \leq \epsilon\end{array}$$

$$W[x] = \left[\begin{array}{c|c} U[x] & v[x] \\ \hline v^T[x] & w[x]\end{array}\right],\ Z[\zeta] = \left[\begin{array}{c|c} \zeta\zeta^T & \zeta \\ \hline \zeta^T & 1\end{array}\right]$$

where $W[x]$ is affine in x.

$$\forall (P \in \mathcal{P}) : \mathrm{Prob}_{\zeta\sim P}\left\{\mathrm{Tr}(WZ[\zeta]) > 0\right\} \leq \epsilon \qquad (!)$$

As about $\mathcal{P}$, we assume that this family is comprised of all distributions on $\mathbb{R}^L$ such that

• *the marginal distributions P_ℓ belong to known families $\mathcal{P}_\ell$ of probability distributions on the axis,*

• *the covariance matrix $V = \mathbf{E}\{Z[\zeta]\}$ of ζ is known to belong to a given closed convex subset $\mathcal{V}$ of the positive semidefinite cone,*

• *ζ is supported on a set S given by a finite list of quadratic constraints:*

$$\mathrm{Tr}(A_i Z[\zeta]) \leq 0, \ \ i = 1, ..., I.$$

$$\forall (P \in \mathcal{P}) : \mathrm{Prob}_{\zeta\sim P}\left\{\mathrm{Tr}(WZ[\zeta]) > 0\right\} \le \epsilon \qquad (!)$$
$$\mathcal{P} : \zeta_\ell \sim P_\ell \in \mathcal{P}_\ell,\ \mathbf{E}_{\zeta\sim P}\{Z[\zeta]\} \in \mathcal{V}$$
$$\forall P \in \mathcal{P} : \mathrm{supp}(P) \subset S = \{z : \mathrm{Tr}(A_i Z[z]) \le 0,\ i \le I\}$$

♣ **The idea** of our approximation scheme inspired by Bertsimas, Popescu and Sethuraman, is as follows:

• *We build a mechanism which produces pairs* $(f(z) : \mathbb{R}^L \to \mathbb{R}, \lambda > 0)$ *such that*

A. $f(z) \ge 0$ *on* S

B. $f(z) \ge \lambda$ *whenever* $z \in S$ *and* $\mathrm{Tr}(WZ[z]) > 0$

C. *We can point out an upper bound* $\Psi[f]$ *on* $\sup\limits_{P\in\mathcal{P}} \mathbf{E}_{\zeta\sim P}\{f(\zeta)\}$.

Clearly, for a pair $(f(\cdot), \lambda)$ **produced by our mechanism for every** $P \in \mathcal{P}$ **we have**

$$\Psi[f] \ge \int f(z)dP(z) = \int_S f(z)dP(z) \ge \int\limits_{\substack{z\in S:\\ \mathrm{Tr}(WZ[z])>0}} f(z)dP(z)$$
$$\ge \lambda \mathrm{Prob}_{\zeta\sim P}\{\mathrm{Tr}(WZ[\zeta]) > 0\},$$

so that *the condition*

$$\Psi[f] \le \lambda\epsilon \qquad (+)$$

is sufficient for the validity of (!).

We impose on the "free parameters" of our construction the requirement to ensure (+), thus arriving at a safe approximation of (!)**.**

♠ **Implementation:**

C: The simplest way to ensure the possibility to bound from above $\sup\limits_{P\in\mathcal{P}} \int f(z)dP(z)$ **is to stick to "simple"** f**, e.g., those of the form**

$$f(z) = \sum_{\ell=1}^{L} f_\ell(z_\ell) + \mathrm{Tr}(FZ[z]), \qquad (*)$$

which allows to take

$$\Psi[f] = \sum_{\ell=1}^{L} \sup_{P_\ell\in\mathcal{P}_\ell} \int f_\ell(s)dP_\ell(s) + \sup_{V\in\mathcal{V}} \mathrm{Tr}(FV).$$

A: The simplest way to ensure that a function f **of the form** $(*)$ **is** ≥ 0 **on** $S = \{z : \mathrm{Tr}(A_i Z[z]) \leq 0, i \leq I\}$ **is to ensure that**

$$f_\ell(s) \geq a_\ell + 2b_\ell s + c_\ell s^2, \ 1 \leq \ell \leq L \qquad (a)$$

and that the quadratic form of z

$$\sum_{\ell=1}^{L} [a_\ell + 2b_\ell z_\ell + c_\ell z_\ell^2] + \mathrm{Tr}(FZ[z]) + \sum_{i=1}^{I} \mu_i \mathrm{Tr}(A_i Z[z])$$

where the parameters μ_i *are* ≥ 0**, is nonnegative everywhere, which amounts to the matrix inequality**

$$\left[\begin{array}{c|c} \mathrm{Diag}\{c_1,...,c_L\} & b \\ \hline b^T & \sum_\ell a_\ell \end{array}\right] + F + \sum_{i=1}^{I} \mu_i A_i \succeq 0$$

$$f(z)=\sum_{\ell=1}^{L} f_\ell(z_\ell)+\mathrm{Tr}(FZ[z]), \qquad (*)$$

B: The simplest way to ensure that a function f **of the form** $(*)$ **is** $\geq \lambda$ **on** $S=\{z:\mathrm{Tr}(A_iZ[z])\leq 0, i\leq I, \mathrm{Tr}(WZ[z])>0\}$ **is to ensure that**

$$f_\ell(s)\geq p_\ell+2q_\ell s+r_\ell s^2,\ 1\leq \ell\leq L \qquad (b)$$

and that the quadratic form of z

$$\sum_{\ell=1}^L[p_\ell+2q_\ell z_\ell+r_\ell z_\ell^2]+\mathrm{Tr}(FZ[z])+\sum_{i=1}^I \nu_i\mathrm{Tr}(A_iZ[z]) \\ -\mathrm{Tr}(WZ[z])-\lambda$$

where the parameters ν_i *are* ≥ 0, **is nonnegative everywhere, which amounts to the matrix inequality**

$$\left[\begin{array}{c|c}\mathrm{Diag}\{r_1,...,r_L\} & q\\ \hline q^T & \sum_\ell p_\ell-\lambda\end{array}\right]+F-W+\sum_{i=1}^I \nu_iA_i\succeq 0$$

$$\forall(P\in\mathcal{P}):\mathrm{Prob}_{\zeta\sim P}\{\mathrm{Tr}(WZ[\zeta])>0\}\leq\epsilon \qquad (!)$$
$$\mathcal{P}:\zeta_\ell\sim P_\ell\in\mathcal{P}_\ell,\ \mathbf{E}_{\zeta\sim P}\{Z[\zeta]\}\in\mathcal{V}$$
$$\forall P\in\mathcal{P}:\mathrm{supp}(P)\subset S=\{z:\mathrm{Tr}(A_iZ[z])\leq 0,\ i\leq I\}$$

♠ **We have arrived at the following**

Theorem: *The system of convex constraints in variables* $W,\ \lambda,\ \{a_\ell,b_\ell,c_\ell,p_\ell,q_\ell,r_\ell\}_{\ell=1}^L,\ \{\mu_i,\nu_i\}_{i=1}^I,\ F\in\mathbf{S}^{L+1},\ \theta$

$$\begin{array}{ll}
(a) & \left[\begin{array}{c|c}\mathrm{Diag}\{c_1,...,c_L\} & b\\ \hline b^T & \sum_\ell a_\ell\end{array}\right]+F+\sum_{i=1}^I\mu_iA_i\succeq 0\\
(b) & \left[\begin{array}{c|c}\mathrm{Diag}\{r_1,...,r_L\} & q\\ \hline q^T & \sum_\ell p_\ell-\lambda\end{array}\right]+F\\
& \qquad\qquad -W+\sum_{i=1}^I\nu_iA_i\succeq 0\\
(c) & \left[\begin{array}{c|c}\lambda & 1\\ \hline 1 & \theta\end{array}\right]\succeq 0\ [\text{says that }\lambda>0]\\
(d) & \sum_{\ell=1}^L\sup_{P_\ell\in\mathcal{P}_\ell}\int\max[a_\ell+2b_\ell s+c_\ell s^2,p_\ell+2q_\ell s+r_\ell s^2]dP_\ell(s)\\
& \qquad\qquad +\max_{V\in\mathcal{V}}\mathrm{Tr}(FV)\leq\lambda\epsilon\\
(e) & \mu_i\geq 0,\ \nu_i\geq 0
\end{array}$$

is a safe convex approximation of the chance constraint (!). *This approximation is tractable, provided that the suprema in* (d) *are efficiently computable.*

♣ How it works? Portfolio Selection revisited.

There are L assets with random yearly returns $r_\ell = 1 + \mu_\ell + \sigma_\ell \zeta_\ell$, $1 \leq \ell \leq L$, where $\mu_\ell \geq 0$ and $\sigma_\ell \geq 0$ are known expected gains and their variabilities, and ζ_ℓ are random perturbations taking values in $[-1, 1]$. Given partial information on the distribution of $\zeta = [\zeta_1; ...; \zeta_L]$, we want to distribute \$1 between the assets in order to maximize the guaranteed value-at-ϵ-risk of the profit $\sum_\ell [\mu_\ell + \sigma_\ell \zeta_\ell] x_\ell$:

$$\max_{x,t} \left\{ t : \begin{array}{c} \mathrm{Prob}_{\zeta \sim P} \left\{ \sum_{\ell=1}^{15} [\mu_\ell + \sigma_\ell \zeta_\ell] x_\ell < t \right\} \leq 0.01 \, \forall P \in \mathcal{P} \\ x \geq 0, \sum_{\ell=1}^{15} x_\ell = 1 \end{array} \right\}$$

♡ **In the experiments, we set** $L=15,\ \epsilon=0.01$**,**

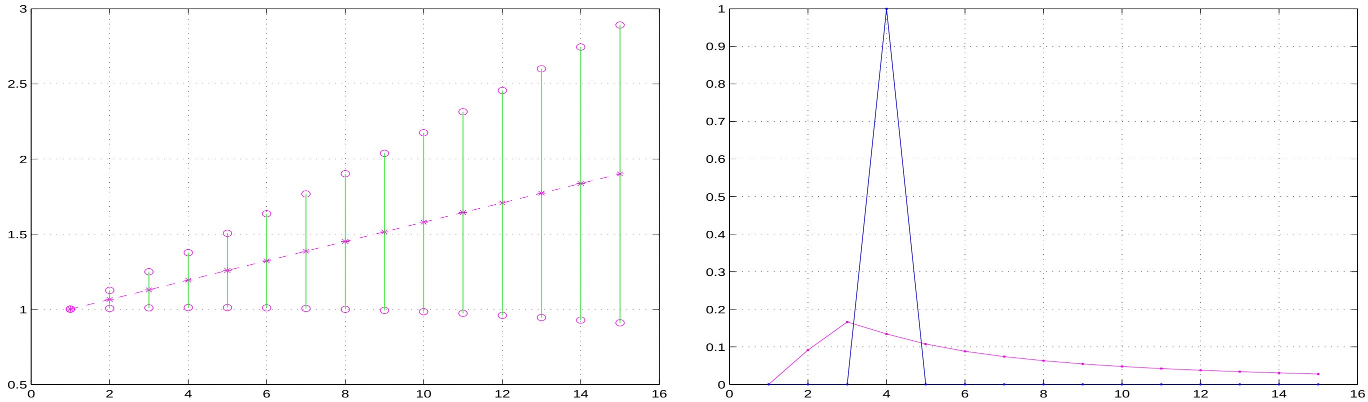
$$\mu_\ell = 0.001 + 0.9\frac{\ell-1}{L-1},\ \sigma_\ell = \left(0.9+0.2\frac{\ell-1}{L-1}\right)\mu_\ell$$

and consider 3 concurrent hypotheses on ζ**:**

A: $\zeta_1,...,\zeta_{15}\in[-1,1]$ *are zero mean and independent*

B: $\zeta_1,...,\zeta_{15}\in[-1,1]$ *are zero mean and uncorrelated*

C: $\zeta_1,...,\zeta_{15}\in[-1,1]$ *are zero mean*



● Expectations and ranges of returns ● Portfolios: **A** (magenta), **B**, **C** (blue)

Hypothesis	Approximation	Guaranteed value-at-1%-risk
A	Bernstein	0.0552
B	Lagrangian	0.0101
C	Lagrangian	0.0101

♣ **Facts:**

• **The single-asset portfolio given in the cases of B, C by Lagrangian approximation is** *exactly optimal* **in these cases**

• **Diversified portfolio given in the case of A by Bernstein approximation can exhibit** *negative profit* **if the true case is B**

Explanation: A, B, C postulate that $\zeta_i \in [-1, 1]$ are zero mean. A postulates that ζ_i are independent, and B – that ζ_i are uncorrelated. There exists a distribution P_{bad} of ζ *compatible with* **B where "crisis" $\zeta = [-1; -1; ...; -1]$ happens with probability $> 1\%$.**

$\Rightarrow$ *Under hypotheses* **B**, **C**, *the guaranteed value-at-1%-risk for any portfolio cannot be better than its profit in the case of crisis. As a matter of fact, in our problem Lagrangian approximation maximizes the profit of a portfolio in the case of crisis.*

Note: Assuming yearly returns of assets i.i.d. over time, *it takes > 100 years* **to distinguish, with reliability 99%, between A and P_{bad}.**

Uncertain Conic Quadratic and Semidefinite Optimization

♣ "Canonical" Conic problem:

$$\min_x \left\{ c^T x + d : \begin{array}{c} A_1 x + b_1 \in \mathbf{K}_1 \\ \cdots \\ A_I x + b_I \in \mathbf{K}_I \end{array} \right\} \qquad \text{(CP)}$$

• x: decision vector

• $\mathbf{K}_i$: simple cone: *nonnegative orthant* $\mathbb{R}^m_+$, or *Lorentz cone* $\mathbf{L}^m = \{y \in \mathbb{R}^m : y_m \geq \sqrt{y_1^2 + ... + y_{m-1}^2}\}$, or *Semidefinite cone* $\mathbf{S}^m_+$ *comprised of positive semidefinite symmetric* $m \times m$ *matrices*, with $m = m_i$.

• **Structure** of (CP): the collection of cones $\mathbf{K}_1, ..., \mathbf{K}_I$

• **Data** of (CP): $c, d, A_1, b_1, ..., A_I, b_I$

$$\min_x \left\{ c^T x + d : \begin{array}{c} A_1 x + b_1 \in \mathbf{K}_1 \\ \cdots \\ A_I x + b_I \in \mathbf{K}_I \end{array} \right\} \tag{CP}$$

♠ **Uncertain canonical conic problem $\mathcal{P}$:** *a collection of canonical conic programs* ("instances") *with common structure and with the data running through a given uncertainty set $\mathcal{U}$.*

♠ *We always assume that the data is affinely parameterized by a perturbation vector ζ running through perturbation set $\mathcal{Z}$:*

$$\mathcal{U} = \Big\{ (c[\zeta], d[\zeta], A_1[\zeta], ..., b_I[\zeta]) := (c^0, d^0, A_1^0, ..., b_I^0) + \sum_{\ell=1}^L \zeta_\ell (c^\ell, d^\ell, A_1^\ell, ..., b_I^\ell) : \zeta \in \mathcal{Z} \Big\}$$

♠ **Robust Counterpart** *of uncertain canonical conic problem $\mathcal{P}$: the problem*

$$\min_{t,x} \left\{ t : \left. \begin{array}{rl} t - c[\zeta]^T x - d[\zeta] & \in \mathbb{R}_+ \\ A_1[\zeta]x + b_1[\zeta] & \in \mathbf{K}_1 \\ \cdots & \\ A_I[\zeta]x + b_I[\zeta] & \in \mathbf{K}_I \end{array} \right\} \forall \zeta \in \mathcal{Z} \right\} \tag{RC}$$

$$
\min_x \left\{ c^T x + d : \begin{array}{c} A_1 x + b_1 \in \mathbf{K}_1 \\ \cdots \\ A_I x + b_I \in \mathbf{K}_I \end{array} \right\}_{(c,d,A_1,\dots,b_I)\in\mathcal{U}}
$$

$$
\begin{array}{rcl}
\mathcal{U} &=& \Big\{ (c[\zeta], d[\zeta], A_1[\zeta], ..., b_I[\zeta]) := (c^0, d^0, A_1^0, ..., b_I^0) \\
&& \quad + \sum_{\ell=1}^{L} \zeta_\ell (c^\ell, d^\ell, A_1^\ell, ..., b_I^\ell) : \zeta \in \mathcal{Z} \Big\}
\end{array}
\tag{UCP}
$$

♠ **Same as in the LO case**, we w.l.o.g. can assume the objective to be certain (and then skip d), in which case the RC becomes

$$
\min_x \left\{ c^T x : \left. \begin{array}{rcl} A_1[\zeta]x + b_1[\zeta] & \in & \mathbf{K}_1 \\ & \dots & \\ A_I[\zeta]x + b_I[\zeta] & \in & \mathbf{K}_I \end{array} \right\} \forall \zeta \in \mathcal{Z} \right\}
\tag{RC}
$$

Note: Same as in LO, the RC remains intact when extending $\mathcal{Z}$ to its closed convex hull

$\Rightarrow$ *From now on, we always assume $\mathcal{Z}$ to be convex and closed.*

♠ **Same as in the LO case**, the RC of (UCP) with certain objective is a constraint-wise construction, and we can focus on the RC

$$
A[\zeta]x + b[\zeta] \in \mathbf{K}
\tag{RC}
$$

of a *single* uncertain conic constraint.

$$\begin{array}{c} A[\zeta]x + b[\zeta] \in \mathbf{K}\ \forall \zeta \in \mathcal{Z} \\ [A[\zeta], b[\zeta]] = [A^0, b^0] + \sum_{\ell=1}^{L} \zeta_\ell [A^\ell, b^\ell] \end{array} \tag{RC}$$

♣ **Questions of primary importance:**

• When the semi-infinite conic constraint (RC) is computationally tractable?

• What to do when (RC) is intractable?

♠ **Fact:** Tractability of (RC) depends on "tradeoff" between the geometries of the perturbation set $\mathcal{Z}$ and the cone $\mathbf{K}$: the simpler is $\mathcal{Z}$, the more complicated can be $\mathbf{K}$.

• When $\mathbf{K}$ is as simple as possible – just a nonnegative orthant (uncertain LO), $\mathcal{Z}$ can be a whatever computationally tractable convex set, e.g., one given by well-structured conic representation.

• When $\mathcal{Z}$ is as simple as possible:

$$\mathcal{Z} = \mathrm{Conv}\left\{\zeta^1, ..., \zeta^N\right\}$$

(scenario uncertainty), (RC) is equivalent to

$$A[\zeta^i]x + b[\zeta^i] \in \mathbf{K},\ i = 1, ..., N,$$

that is, (RC) is tractable whenever the cone $\mathbf{K}$ is so (as is the case for nonnegative orthants, Lorentz and Semidefinite cones we are interested in).

$$\begin{array}{c} A[\zeta]x + b[\zeta] \in \mathbf{K}\ \forall \zeta \in \mathcal{Z} \\ [A[\zeta], b[\zeta]] = [A^0, b^0] + \sum_{\ell=1}^{L} \zeta_\ell [A^\ell, b^\ell] \end{array} \qquad \text{(RC)}$$

• In the "in-between" situations, tractability of (RC) is a "rare commodity:"

• When $\mathbf{K}$ is the Lorentz cone, (RC) is tractable when $\mathcal{Z}$ is an ellipsoid, and is intractable when $\mathcal{Z}$ is a box.
Indeed, checking feasibility of $x = 0$ for the semi-infinite Least Squares inequality

$$\begin{array}{c} \|Ax + b\|_2 \leq 1\ \forall b \in \mathcal{U} = \{B\zeta : \|\zeta\|_\infty \leq 1\} \\ \Leftrightarrow [Ax + B\zeta; 1] \in \mathbf{L}^{m+1}\ \forall \zeta, \|\zeta\|_\infty \leq 1 \end{array}$$

reduces to checking whether $\|B\zeta\|_2 \leq 1$ for all ζ with $\|\zeta\|_\infty \leq 1$, or, equivalently, to the maximization of the positive semidefinite quadratic form $\zeta^T[B^TB]\zeta$ over the unit box, which is NP-hard even when accuracy of 4% is sought.

• When $\mathbf{K}$ is a Semidefinite cone, (RC) is intractable when $\mathcal{Z}$ is either an ellipsoid or a box.

♠ As a result, *In Robust Conic Optimization the goals of primary importance are*

• *To discover special cases where* (RC) *is tractable*, and

• *To build tight safe tractable approximations to* (RC) *when* (RC) *"'as it is" is intractable.*

$$\begin{array}{c} A[\zeta]x + b[\zeta] \in \mathbf{K} \ \forall \zeta \in \mathcal{Z} \\ [A[\zeta], b[\zeta]] = [A^0, b^0] + \sum_{\ell=1}^{L} \zeta_\ell [A^\ell, b^\ell] \end{array} \qquad \text{(RC)}$$

Definition. A system $\mathcal{S}$ of efficiently computable convex constraints in variables x and, perhaps, additional variables u is called a *safe tractable approximation* of (RC) if the projection $X[\mathcal{S}]$ of the feasible set of the system on the plane of x-variables is contained in the feasible set of (RC), that is,

$$(x, u) \text{ is feasible for } \mathcal{S} \ \Rightarrow \ x \text{ is feasible for (RC)}$$

♡ Replacing the RC's of the conic constraints in an uncertain conic problem $\mathcal{P}$ with their safe tractable approximations, we end up with a *computationally tractable* problem such that (the x-components of) its feasible solutions are *feasible for the RC* of the uncertain problem.

♠ **Question:** How to quantify "tightness" of a safe approximation?

♠ **Answer:** Assume, as it usually is the case, that $0 \in \mathcal{Z}$ ($\zeta = 0$ corresponds to the nominal data).

$\Rightarrow$ We can embed (RC) in a parametric family of RC's

$$A[\zeta]x + b[\zeta] \in \mathbf{K} \ \forall \zeta \in \mathcal{Z}_\rho := \rho\mathcal{Z} \qquad \text{(RC}[\rho])$$

As the *uncertainty level* ρ grows, the perturbation set $\mathcal{Z}_\rho$ extends, and the feasible set X_ρ of (RC[ρ]) shrinks.

$$A[\zeta]x + b[\zeta] \in \mathbf{K}\ \forall \zeta \in \mathcal{Z}_\rho := \rho\mathcal{Z} \qquad (\mathrm{RC}[\rho])$$

Definition: A *safe tractable approximation* of (RC[ρ]) is a system $\mathcal{S}[\rho]$ of efficiently computable convex constraints, depending on $\rho \geq 0$ as on a parameter, in variables x and additional variables u such that the projection $\boldsymbol{\mathcal{X}}_\rho$ of the feasible set of $\mathcal{S}[\rho]$ on the plane of x-variables is, for every $\rho \geq 0$, contained in the feasible set X_ρ of (RC[ρ]).
Such an approximation is called *ϑ-tight*, if

$$\forall \rho \geq 0 : X_\rho \supset \boldsymbol{\mathcal{X}}_\rho \supset X_{\vartheta\rho}.$$

Equivalently: $\mathcal{S}[\cdot]$ is ϑ-tight safe approximation of (RC[$\cdot$]), if, for every $\rho \geq 0$,

• [safety] whenever x can be extended to a feasible solution to $\mathcal{S}[\rho]$, x is feasible for (RC[ρ]), and

• [tightness] whenever x cannot be extended to a feasible solution to $\mathcal{S}[\rho]$, x is not feasible for (RC[$\vartheta\rho$]).

• We call an approximation scheme *tight*, if its tightness factor is independent of the numerical values of the data specifying $\mathcal{Z}$.

Example: The semi-infinite Least Squares inequality with box uncertainty

$$\begin{array}{c}\forall \zeta, \|\zeta\|_\infty \leq \rho : \|A[\zeta]x + b[\zeta]\|_2 \leq \tau \\ \Updownarrow \\ [A[\zeta]x + b[\zeta]; \tau] \in \mathbf{L}^{m+1}\ \forall \zeta, \|\zeta\|_\infty \leq \rho \qquad (\mathrm{RC}[\rho]) \\ [A[\zeta], b[\zeta]] = [A^0, b^0] + \sum_{\ell=1}^L \zeta_\ell [A^\ell, b^\ell]\end{array}$$

is, in general, computationally intractable. It, however, admits a *tight within the factor* $\pi/2$ safe tractable approximation

$$\begin{array}{l}\left[\begin{array}{c|c}\tau & [A^0x + b^0]^T \\ \hline A^0x + b^0 & \tau I\end{array}\right] - \rho \sum_{\ell=1}^L Y_\ell \succeq 0 \\ Y_\ell \succeq \pm \left[\begin{array}{c|c} & [A^\ell x + b^\ell]^T \\ \hline A^\ell x + b^\ell & \end{array}\right] \succeq 0,\ 1 \leq \ell \leq L\end{array} \qquad \mathcal{S}[\rho]$$

(variables are x and symmetric matrices $Y_1, ..., Y_L$).

♡ As far as the x-components of feasible solutions are concerned, the above system can be replaced with a system with a much smaller number of variables, namely

$$\left[\begin{array}{c|c|ccc}\tau - \sum_{\ell=1}^L \lambda_\ell & [A^0x + b^0]^T & & & \\ \hline A^0x + b^0 & \tau I_m & A^1x + b^1 & \dots & A^Lx + b^L \\ \hline & [A_1x + b_1]^T & \lambda_1 & & \\ & \vdots & & \ddots & \\ & [A^Lx + b^L]^T & & & \lambda_L\end{array}\right] \succeq 0$$

(variables are x and $\lambda_1, ..., \lambda_L$).

Note: Let $\mathcal{P}$ be an uncertain canonical conic problem with certain objective. Assume that the RC's of all conic constraints of $\mathcal{P}$ admit ϑ-tight safe tractable approximations. Then the optimal value $\mathrm{Opt}_{\mathrm{Appr}}(\rho)$ of the resulting safe tractable approximation of $\mathcal{P}$, treated as a function of the uncertainty level ρ, satisfies

$$\mathrm{Opt}_{\mathcal{P}}(\rho) \leq \mathrm{Opt}_{\mathrm{Appr}}(\rho) \leq \mathrm{Opt}_{\mathcal{P}}(\vartheta\rho).$$

Tractable Reformulations and Tight Tractable Approximations of Semi-Infinite Conic Quadratic Inequalities

♣ The fact that a vector $[Ax+b; c^Tx+d] \in \mathbb{R}^{m+1}$ affinely depending on x belongs to the Lorentz cone $\mathbf{L}^{m+1}$ can be equivalently represented by *conic quadratic inequality* (c.q.i.)

$$\|Ax+b\|_2 \leq c^Tx+d \tag{CQI}$$

When $[A,b;c^T,d]$ are affinely parameterized by a perturbation vector ζ:

$$\begin{array}{l}[A,b;c^T,d]=[A[\zeta],b[\zeta];c^T[\zeta],d[\zeta]]\\ =[A^0,b^0;[c^0]^T,d^0]+\sum_{\ell=1}^L \zeta_\ell[A^\ell,b^\ell;[c^\ell]^T,d^\ell]\end{array}$$

and we want the inclusion $[Ax+b;c^Tx+d] \in \mathbf{L}^{m+1}$ (or, equivalently, the c.q.i. (CQI)) to hold true for all $\zeta \in \mathcal{Z}$, we end up with *semi-infinite* c.q.i.

$$\|A[\zeta]x+b[\zeta]\|_2 \leq c^T[\zeta]x+d[\zeta]\ \forall \zeta \in \mathcal{Z} \tag{RC}$$

Note: Convex quadratic constraint

$$x^TA^TAx+2b^Tx+c \leq 0,$$

the data being A,b,c, is equivalent to the c.q.i.

$$\|[2Ax;1+2b^Tx+c]\|_2 \leq 1-2b^Tx-c$$

$\Rightarrow$ *What follows covers, in particular, the RC's of uncertainty-affected convex quadratic constraints.*

$$\|A[\zeta]x + b[\zeta]\|_2 \leq c^T[\zeta]x + d[\zeta] \ \forall \zeta \in \mathcal{Z} \qquad \text{(RC)}$$

♠ Aside of the trivial case of scenario-generated perturbation set $\mathcal{Z}$, essentially the only known generic case when (RC) is computationally tractable/admits tight computationally tractable approximation *independently of any assumptions on how the affine perturbations enter the problem* is the case of *an ellipsoid* $\mathcal{Z}$.

All other tractability results known to us deal with the case of *semi-infinite Least Squares inequality*

$$\|A[\zeta]x + b[\zeta]\|_2 \leq \tau \ \forall \zeta \in \mathcal{Z}$$

in variables x, τ, or, which is the same, with the case when the right hand side in (RC) is not affected by the uncertainty, since in this case (RC) is equivalent to

$$\|A[\zeta]x + b[\zeta]\|_2 \leq \tau \ \forall \zeta \in \mathcal{Z} \ \& \ \tau \leq c^T x + d$$

and tractable reformulation/building tight safe tractable approximation of the latter system reduces to the same problem for the "troublemaking" semi-infinite Least Squares inequality.

$$\|A[\zeta]x + b[\zeta]\|_2 \leq c^T[\zeta]x + d[\zeta] \ \forall \zeta \in \mathcal{Z} \qquad \text{(RC)}$$

♡ **Note:** (RC) with *side-wise* uncertainty, where the perturbation set $\mathcal{Z}$ is the product $\mathcal{Z}^{\text{left}} \times \mathcal{Z}^{\text{right}}$ of sets of perturbations affecting the left- and the right hand side data, reduces to the system of semi-infinite constraints

$$\begin{array}{ll} (a) & \|A[\zeta^{\text{left}}]x + b[\zeta^{\text{left}}]\|_2 \leq \tau \ \forall \zeta^{\text{left}} \in \mathcal{Z}^{\text{left}}, \\ (b) & c^T[\zeta^{\text{right}}]x + d[\zeta^{\text{right}}] \geq \tau \ \forall \zeta^{\text{right}} \in \mathcal{Z}^{\text{right}} \end{array}$$

in variables x, τ.

• Whenever the right hand side perturbation set is tractable, the semi-infinite constraint (b) is computationally tractable

$\Rightarrow$ *All we need to process the RC efficiently is to build a tractable reformulation/tight tractable approximation of the semi-infinite Least Squares inequality* (a).

$$\|A[\zeta]x + b[\zeta]\|_2 \leq c^T[\zeta]x + d[\zeta] \ \forall \zeta \in \mathcal{Z} \qquad \text{(RC)}$$

Tractable Case I: $\mathcal{Z}$ is an Ellipsoid

♣ W.l.o.g. we can assume that the ellipsoid $\mathcal{Z}$ is just the unit ball centered at the origin:

$$\mathcal{Z} = \{\zeta \in \mathbb{R}^L : \|\zeta\|_2 \leq 1\}.$$

♠ A well-structured representation of (the feasible set of) (RC) was recently found by R. Hildenbrandt
(Highly nontrivial construction with highly nontrivial justification.)

♠ We restrict ourselves with an easy demonstration that $\mathcal{Z}$ *being the ball, the feasible set X of (RC)* (which clearly is a closed convex set) *admits an efficient Separation oracle – a routine which, given on input x, reports whether $x \in X$, and if it is not the case, returns a separator - a linear form e such that*

$$e^T x > \max_{x' \in X} e^T x'.$$

Note: Given an efficient Separation oracle for X, we can use, say, the Ellipsoid method to solve efficiently convex problems of the form

$$\min_{x \in X} \{f(x) : g_i(x) \leq 0, \, i = 1, ..., m\}$$

provided that f and g_i are efficiently computable.

$$X = \left\{x : \|A[\zeta]x + b[\zeta]\|_2 \le c^T[\zeta]x + d[\zeta] \ \forall \zeta, \|\zeta\|_2 \le 1\right\} \tag{RCBall}$$

♠ Since $A[\zeta]$,..., $d[\zeta]$ are affine in ζ, (RCBall) reads

$$X = \{x : \|\alpha[x]\zeta + \beta[x]\|_2 \le \gamma^T[x]\zeta + \delta[x] \ \ \forall \zeta, \|\zeta\|_2 \le 1\}$$

with $\alpha[x]$,...,$\delta[x]$ affine in x.

♣ Clearly, $x \in X$ *iff*

A. $\|\zeta\|_2 \le 1 \to \gamma^T[x]\zeta + \delta[x] \ge 0 \Leftrightarrow \|\gamma[x]\|_2 \le \delta[x]$

and

B. *The quadratic form of* ζ*:* $\zeta^T P[x]\zeta + 2p^T[x]\zeta + q[x]$ $\equiv [\gamma^T[x]\zeta + \delta[x]]^2 - \|\alpha[x]\zeta + \beta[x]\|_2^2$ *is* ≥ 0 *on the domain* $\zeta^T\zeta \le 1$*.*

♣ **Miracle # 1: $\mathcal{S}$-Lemma:** *Consider two quadratic forms*

$$f(z) = z^T Az + 2a^T z + \alpha, \ \ g(z) = z^T Bz + 2b^T z + \beta$$

and let the set $\{z : g(z) > 0\}$ *be nonempty. Then the implication*

$$g(z) \ge 0 \Rightarrow f(z) \ge 0$$

holds true iff $\exists \lambda \ge 0 : f(z) \ge \lambda g(z) \ \forall z$*, that is, iff*

$$\exists \lambda \ge 0 : \left[\begin{array}{c|c} A - \lambda B & a - \lambda b \\ \hline [a - \lambda b]^T & \alpha - \lambda\beta \end{array}\right] \succeq 0.$$

$$X = \{x : \|\alpha[x]\zeta + \beta[x]\|_2 \leq \gamma^T[x]\zeta + \delta[x] \;\; \forall \zeta, \|\zeta\|_2 \leq 1\}$$

♠ $x \in X$ iff the following two properties take place:

A. $\|\gamma[x]\|_2 \leq \delta[x]$

B. $g(z) = 1 - z^T z \geq 0 \Rightarrow f(z) := z^T P[x]z + 2p^T[x]z + q[x]$

$\Leftrightarrow \exists \lambda > 0 : \left[\begin{array}{c|c} P[x] + \lambda I & p[x] \\ \hline p^T[x] & q[x] - \lambda \end{array}\right] \succeq 0$

♡ Given x, we can easily verify **A** and **B**; if both the properties hold true, we report that $x \in X$.

• Now let *either* **A**, *or* **B**, *or both do not take place.* On a closest inspection, here we can find efficiently $\bar{\zeta}$, $\|\bar{\zeta}\|_2 \leq 1$, such that the vector

$$\bar{y} = [\alpha[x]\bar{\zeta} + \beta[x]; \gamma^T[x]\bar{\zeta} + \delta[x]]$$

does *not* belong to $\mathbf{L}^{m+1}$, and thus can be easily separated from $\mathbf{L}^{m+1}$, That is, we can efficiently build $\eta \in \mathbb{R}^{L+1}$ such that

$$\eta^T \bar{y} > s := \sup_{y \in \mathbf{L}^{m+1}} \eta^T y.$$

$\Rightarrow$ The affine form $e[\xi] = \eta^T[\alpha[\xi]\bar{\zeta} + \beta[\xi]; \gamma^T[\xi]\bar{\zeta} + \delta[\xi]]$ separates x and X, and we can return this form as a required separator.

Tractable Case II: Semi-Infinite Least Squares Inequality with Unstructured Norm-Bounded Uncertainty

$$\|A[\zeta]x + b[\zeta]\|_2 \leq \tau \ \forall \zeta \in \mathcal{Z} \tag{RC}$$

Definition: We say that the uncertainty in (RC) is *unstructured norm-bounded*, if

• $\mathcal{Z}$ is the set of $p \times q$ matrices of matrix norm $\|\cdot\|_{2,2}$ not exceeding 1

• $A[\zeta]x + b[\zeta] \equiv A^0x + b^0 + L^T[x]\zeta R[x]$ with matrices $L[x]$, $R[x]$ of appropriate sizes affinely depending on x and such that *either $L[x]$, or $R[x]$ is constant*.

♠ **Example: $\mathcal{Z}$ is an ellipsoid.** W.l.o.g. we can assume that $\mathcal{Z} = \{\zeta \in \mathbb{R}^L : \|\zeta\|_2 \leq 1\}$, that is, $\mathcal{Z}$ is comprised of $L \times 1$ matrices ζ of norm ≤ 1. We have

$$A[\zeta]x + b[\zeta] \equiv \alpha[x]\zeta + \beta[x]$$

with affine in x $\alpha[x]$ and $\beta[x]$, and we can set

$$A^0x + b^0 = \beta[x],\ L^T[x] = \alpha[x],\ R[x] = 1.$$

$$\begin{array}{c}\|A[\zeta]x + b[\zeta]\|_2 \leq \tau \ \forall \zeta \in \mathbb{R}^{p\times q} : \|\zeta\|_{2,2} \leq 1 \\ A[\zeta]x + b[\zeta] \equiv A^0 x + b^0 + L^T[x]\zeta R[x]\end{array} \qquad \text{(RC)}$$

Theorem: *When $R[x]$ is independent of x,* (RC) *can be represented equivalently by the Linear Matrix Inequality*

$$\left[\begin{array}{c|c|c} \tau - \lambda R^T R & [A^0 x + b^0]^T & \\ \hline A^0 x + b^0 & \tau I_m & L^T[x] \\ \hline & L[x] & \lambda I_p \end{array}\right] \succeq 0,$$

in variables x, λ.
When $L[x]$ is independent of x, (RC) *can be represented equivalently by the LMI*

$$\left[\begin{array}{c|c|c} \tau & [A^0 x + b^0]^T & R^T[x] \\ \hline A^0 x + b^0 & \tau I_m - \lambda L^T L & \\ \hline R[x] & & \lambda I_q \end{array}\right] \succeq 0$$

in variables x, λ.
Proof: to be given later.

Example [Robust Linear Estimation] *Given noisy observation*

$$w = [A + B\Delta C]v + \xi \qquad [\xi \sim \mathcal{N}(0, \Sigma)]$$

of unknown signal v known to belong to a given ellipsoid $\{v : v^T Q v \leq R^2\}$, estimate the value at v of a given linear form $\langle f, \cdot \rangle$.

Here: • A, B, C are given matrices, • Δ is unknown perturbation known to be norm-bounded: $\|\Delta\|_{2,2} \leq \rho$ with a given ρ.

We restrict ourselves with linear in w estimates $x^T w$ (x is the weight vector to be found) *and want to minimize the worst-case expected squared recovery error.* Thus, our problem is

$$\min_x \left\{ \max_{\substack{v: v^T Q v \leq R^2 \\ \Delta: \|\Delta\|_{2,2} \leq \rho}} \sqrt{\mathbf{E}\left\{ \left(f^T v - x^T [[A + B\Delta C]v + \xi]\right)^2 \right\}} \right\}$$

We have

$$\begin{array}{l}
\mathbf{E}\left\{ \left(f^T v - x^T [[A + B\Delta C]v + \xi]\right)^2 \right\} = \left[v^T [f - [A^T + C^T \Delta^T B^T]x]\right]^2 + x^T \Sigma x \\
\Rightarrow \max\limits_{v: v^T A Q v \leq R^2} \mathbf{E}\left\{ \left(f^T v - x^T [[A + B\Delta C]v + \xi]\right)^2 \right\} \\
\qquad = R^2 [f - [A^T + C^T \Delta^T B^T]x]^T Q^{-1} [f - [A^T + C^T \Delta^T B^T]x] + x^T \Sigma x.
\end{array}$$

$\Rightarrow$ *the problem of interest is*

$$\min_{x,\tau,s,r} \left\{ r : \begin{array}{c} \|\Sigma^{1/2} x\|_2 \leq s, \sqrt{R^2 \tau^2 + s^2} \leq r \\ \|Q^{-1/2} [f - [A^T + C^T \Delta^T B^T]x]\|_2 \leq \tau \\ \forall \Delta, \|\Delta\|_{2,2} \leq \rho \end{array} \right\}.$$

$$\min_{x,\tau,s,r}\left\{r:\ \begin{array}{r}\|\Sigma^{1/2}x\|_2\leq s,\sqrt{R^2\tau^2+s^2}\leq r\\ \|Q^{-1/2}[f-[A^T+C^T\Delta^TB^T]x]\|_2\leq\tau\\ \forall\Delta,\|\Delta\|_{2,2}\leq\rho\end{array}\right\}.$$

♠ The only "troublemaker" is the semi-infinite Least Squares inequality

$$\begin{array}{l}\|Q^{-1/2}[f-[A^T+C^T\Delta^TB^T]x]\|_2\\ \equiv\|\overbrace{Q^{-1/2}[f-A^Tx]}^{A^0x+b^0}+\overbrace{[-Q^{-1/2}C^T\Delta^TB^Tx]}^{L^T\zeta R[x],\zeta=\Delta^T/\rho}\|_2\leq\tau\\ \hfill\forall\zeta:\|\zeta\|_{2,2}\leq 1\end{array}$$

Passing to its tractable reformulation, the problem of interest becomes an explicit canonical conic program

$$\min_{x,\tau,r,s,\lambda}\left\{r:\ \begin{array}{c}\|\Sigma^{1/2}x\|_2\leq s,\sqrt{R^2\tau^2+s^2}\leq r\\ \left[\begin{array}{c|c|c}\tau & [Q^{-1/2}[f-A^Tx]]^T & \rho x^TB\\ \hline Q^{-1/2}[f-A^Tx] & \tau I-\lambda Q^{-1/2}C^TCQ^{-1/2} & \\ \hline \rho B^Tx & & \lambda I\end{array}\right]\succeq 0\end{array}\right\}$$

Tight Approximation of Semi-Infinite Least Squares Inequality with Structured Norm-Bounded Uncertainty

$$\|A[\zeta]x + b[\zeta]\|_2 \le \tau \ \forall \zeta \in \rho\mathcal{Z} \tag{RC}$$

Definition: We say that the uncertainty in (RC) is *structured norm-bounded*, if

• $\mathcal{Z}$ is the set of collections of L $p_\ell \times q_\ell$ matrices ζ^ℓ of matrix norm $\|\cdot\|_{2,2}$ not exceeding 1

• $A[\zeta]x + b[\zeta] \equiv A^0x + b^0 + \sum_{\ell=1}^{L} L_\ell^T[x]\zeta^\ell R_\ell[x]$ with matrices $L_\ell[x]$, $R_\ell[x]$ of appropriate sizes affinely depending on x and such that *for every ℓ, either $L_\ell[x]$, or $R_\ell[x]$ is constant.*

♠ **Example: $\mathcal{Z}$ is a box.** W.l.o.g. we can assume that $\mathcal{Z} = \{\zeta \in \mathbb{R}^L : \|\zeta\|_\infty \le 1\}$, that is, $\mathcal{Z}$ is comprised of L 1×1 matrices $\zeta^\ell = \zeta_\ell$ of norm ≤ 1. We have

$$A[\zeta]x + b[\zeta] \equiv \sum_{\ell=1}^{L} \zeta_\ell \alpha_\ell[x] + \beta[x]$$

with affine in x $\alpha_\ell[x]$ and $\beta[x]$, and we can set

$$A^0x + b^0 = \beta[x],\ L_\ell^T[x] = \alpha_\ell[x],\ R_\ell[x] = 1.$$

$$
\begin{array}{l}
\|A[\zeta]x+b[\zeta]\|_2 \\
\equiv \|A^0x+b^0+\sum_{\ell=1}^L L_\ell^T[x]\zeta^\ell R_\ell[x]\|_2 \leq \tau \ \forall \zeta\in\rho\mathcal{Z}, \qquad \text{(RC)} \\
\mathcal{Z}=\{\zeta=(\zeta^1,...,\zeta^L): \zeta^\ell\in\mathbb{R}^{p_\ell\times q_\ell}, \|\zeta^\ell\|_{2,2}\leq 1\}.
\end{array}
$$

Theorem: *Semi-infinite Least Squares inequality with structured norm-bounded uncertainty admits safe tractable approximation given by an explicit system of LMIs and tight within the factor* $\pi/2$. *This approximation is precise when* $L=1$ (i.e., in the case of unstructured norm-bounded perturbation).

Explicit representation of the approximation and its justification will be presented later.

Example: Antenna Design with Least Squares fit. When speaking on Robust LO, we have considered the Antenna Design problem with uniform fit. Now consider similar problem with Least Squares fit:

$$\min_{x,\tau}\left\{\tau : \|D_* - \sum_{\ell=1}^{10} x_\ell D_\ell\|_{2,w} \le \tau\right\}$$

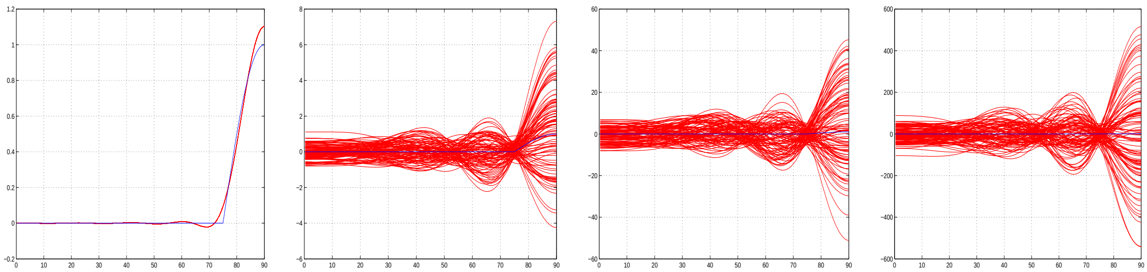
$$\left[\begin{array}{c} x_\ell \mapsto (1+\zeta_\ell)x_\ell,\ \zeta_\ell \in [-\rho,\rho] \\ \Updownarrow \\ D_\ell \mapsto (1+\zeta_\ell)D_\ell,\ \zeta_\ell \in [-\rho,\rho] \end{array}\right]$$

- $D_*,\ D_1,...,D_{10}$: restrictions onto the grid $\delta_i = i\pi/480$, $1 \le i \le 240$ of altitude angles of the target diagram and the diagrams of 10 antenna elements.
- $\|D\|_{2,w}^2 = \frac{1}{240}\sum_{i=1}^{240}\cos(\theta_i)D^2(\theta_i)$

Origin of weights: Physically, diagrams in question are functions on the upper hemisphere S depending solely of the altitude angle, and our weighted 2-norm mimics the standard norm of $L_2(S)$.

$$\min_{\tau,x}\left\{\tau : \|D_* - \sum_{\ell=1}^{10} x_\ell(1+\zeta_\ell)D_\ell\|_{2,w} \le \tau\right\}$$
$$-\rho \le \zeta_\ell \le \rho,\ 1 \le \ell \le 10$$

♠ **Nominal optimal design**, same as in the case of uniform fit, is a complete disaster when implementation errors are present:



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

"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
$\|\cdot\|_{2,w}$ distance to target	0.011	0.077	0.424	0.957	1.177	4.687	9.711	8.709	45.15	109.5
energy concentration	99.4%	0.23%	20.1%	77.7%	0.70	19.5%	61.5%	0.53%	18.9%	61.5%

Quality of nominal antenna design. Data over 100 samples of actuation errors per each value of ρ.

$$\min_{\tau,x}\left\{\tau : \|D_* - \sum_{\ell=1}^{10} x_\ell(1+\zeta_\ell)D_\ell\|_{2,w} \leq \tau \forall \zeta, \|\zeta\|_\infty \leq \rho\right\}_{\text{RC}}$$

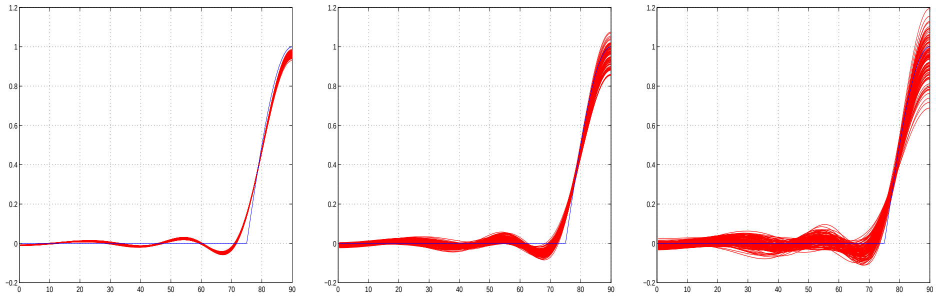
♠ We are in the case of box (and thus – structured norm-bounded) uncertainty, whence (RC) admits a tight, within the factor $\pi/2$, tractable approximation. The approximation reads

$$\begin{array}{ll}\min\limits_{\tau,x,\gamma} & \tau \\ \text{s.t.} & \\ \multicolumn{2}{l}{\left[\begin{array}{c|c|c} \tau - \sum_{\nu=1}^{L}\gamma_\nu & [WDx-b]^T & \\ \hline WDx-b & \tau I & \rho WD\mathrm{Diag}\{x\} \\ \hline & \rho[WD\mathrm{Diag}\{x\}]^T & \mathrm{Diag}\{\gamma_1,...,\gamma_{10}\} \end{array}\right] \succeq 0}\end{array}$$

where:

- $D = [D_{i\ell} = D_\ell(\theta_i)]_{\substack{1\leq\ell\leq 10,\\ 1\leq i\leq 240}}$,
- $W = \mathrm{Diag}\{\cos(\theta_1), ..., \cos(\theta_{240})\}/\sqrt{240}$
- $b = WD_*$

Setting $\rho = 0.01$, we end up with *robust design* which withstands implementation errors incomparably better than the nominal one:



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

"Dream and reality," robust optimal design: samples of 100 of 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
$\|\cdot\|_{2,w}$ distance to target	0.021	0.021	0.021	0.021	0.023	0.030	0.021	0.030	0.048
energy concentration	96.5%	96.7%	96.9%	93.0%	95.8%	96.8%	80.6%	92.9%	96.7%

Quality of robust antenna design. Data over 100 samples of actuation errors per each value of ρ.

For comparison: For nominal design, with $\rho = 0.001$, the average $\|\cdot\|_{2,w}$-distance of the actual diagram to target is as large as 4.69, and the expected energy concentration is as low as 19.5%.

♠ How conservative is our safe tractable approximation?

• The robust design was obtained from *safe tractable approximation* of the true RC rather than from the RC itself $\Rightarrow$ the guaranteed value $\mathrm{ApprOpt}(0.01) = 0.0212$ of the objective can be larger than the true optimal value $\mathrm{RobOpt}(0.01)$ of the RC at the uncertainty level 0.01. How large is the loss in optimality?

• Our approximation is tight within the factor $\pi/2$. This means only that

$$\mathrm{RobOpt}(0.01) \leq \mathrm{ApprOpt}(0.01) \leq \mathrm{RobOpt}(0.01 \cdot \pi/2)$$

and does not allow to make any conclusion on how far is $\mathrm{ApprOpt}(0.01)$ from $\mathrm{RobOpt}(0.01)$.

• The nominal optimal value $\mathrm{NomOpt} = 0.011$ is a lower bound on $\mathrm{RobOpt}(0.01) \Rightarrow \mathrm{ApprOpt}(0.01) = 0.0212$ is within 90% of the true robust optimal value.

• Fortunately, our perturbation set is a box in $\mathbb{R}^{10}$ and thus can be treated as a set given by a large, but not prohibitively so, number $2^{10} = 1024$ of scenarios. This allows to compute the true robust optimal value *exactly*, *and it turns out to be by just 0.2% worse than its upper bound* $\mathrm{ApprOpt}(0.01)$.

Tight Tractable Approximation of Semi-Infinite Least Squares Inequality with $\bigcap$-Ellipsoidal Perturbation Set

$$\|A[\zeta]x - b[\zeta]\|_2 \equiv \|\alpha[x]\zeta + \beta[x]\|_2 \leq \tau \;\forall \zeta \in \rho\mathcal{Z} \qquad \text{RC}$$

♣ Consider the case of $\bigcap$-*ellipsoidal perturbations*

$$\mathcal{Z} = \{\zeta \in \mathbb{R}^L : \zeta^T Q_j \zeta \leq 1,\, 1 \leq j \leq J\}$$

$$[Q_j \succeq 0, \sum_j Q_j \succ 0]$$

Geometrically: $\mathcal{Z}$ is bounded and is the intersection of J ellipsoids/elliptic cylinders centered at the origin.

Examples: • Unit ball; • A polytope $\mathcal{Z}$, $0 \in \mathrm{int}\mathcal{Z}$, symmetric w.r.t. the origin (e.g., a box centered at the origin). Indeed, such a polytope can be represented as $\mathcal{Z} = \{\zeta : (q_j^T\zeta)^2 \leq 1,\, 1 \leq j \leq J\}$.

♠ **Deriving the approximation.** We ask when

$$\{\zeta^T Q_j \zeta \leq \rho^2,\, 1 \leq j \leq J\} \Rightarrow \|\alpha[x]\zeta + \beta[x]\|_2^2 \leq \tau^2,$$

or, equivalently, when

$$\begin{array}{c}\{\zeta^T Q_j \zeta \leq \rho^2,\, 1 \leq j \leq J, t^2 \leq 1\} \\ \Rightarrow \|\alpha[x]\zeta + t\beta[x]\|_2^2 \leq \tau^2\end{array} \qquad (!)$$

An evident *sufficient* condition for (!) is:

$$\exists \{\lambda_j \geq 0\}_{j=0}^J : \left\{\begin{array}{l}\lambda_0 + \rho^2\lambda_1 + \ldots + \rho^2\lambda_J \leq \tau^2 \\ \lambda_0 t^2 + \lambda_1 \zeta^T Q_1 \zeta + \ldots + \lambda_J \zeta^T Q_J \zeta \\ \qquad \geq \|\alpha[x]\zeta + t\beta[x]\|_2^2 \;\forall (t,\zeta).\end{array}\right.$$

$$\begin{array}{r}\|A[\zeta]x-b[\zeta]\|_2\equiv\|\alpha[x]\zeta+\beta[x]\|_2\le\tau\\ \forall\zeta:\zeta^TQ_j\zeta\le\rho^2,1\le j\le J\end{array}\qquad\text{RC}$$

♡Assuming $\tau>0$ and setting $\mu_j=\lambda_j/\tau$, the above sufficient condition for the validity of (RC) reads

$$\begin{array}{ll}\exists\mu_j\ge0,\ 0\le j\le J: & \\ (a) & \mu_0+\rho^2\sum_{j=1}^J\mu_j\le\tau\\ (b) & \left[\begin{array}{c|c}\mu_0 & \\ \hline & \sum_{j=1}^J\mu_jQ_j\end{array}\right]-\frac{1}{\tau}[\beta[x],\alpha[x]]^T[\beta[x],\alpha[x]]\succeq0\end{array}$$

Now let us use

Miracle # 2: Schur Complement Lemma: *A symmetric block matrix* $P=\left[\begin{array}{cc}A & B\\ B^T & C\end{array}\right]$ *with* $C\succ0$ *is* $\succeq0$ *iff* $A-BC^{-1}B^T\succeq0.$

Proof: $P\succeq0$ iff $[u;v]^TP[u;v]\ge0$ for all u, v, that is, iff $\min\limits_v[u;v]^TP[u;v]\ge0$ for all u. Since $C\succ0$, the latter minimum is $u^T[A-BC^{-1}B^T]u$. □

♡ Thus, a sufficient condition for $(x,\tau>0)$ to satisfy (RC) is

$$\exists\{\mu_j\ge0\}_{j=0}^J:\ \begin{array}{c}\left[\begin{array}{c|c|c}\mu_0 & & \beta^T[x]\\ \hline & \sum_{j=1}^J\mu_jQ_j & \alpha^T[x]\\ \hline \beta[x] & \alpha[x] & \tau I\end{array}\right]\succeq0\\ \mu_0\le\tau-\rho^2\sum_{j=1}^J\mu_j\end{array}$$

$$\begin{array}{r}\|A[\zeta]x - b[\zeta]\|_2 \equiv \|\alpha[x]\zeta + \beta[x]\|_2 \leq \tau \\ \forall \zeta : \zeta^T Q_j \zeta \leq \rho^2, 1 \leq j \leq J\end{array} \qquad \text{RC}$$

♡ We have proved the first part of the following

Theorem. (i) *The explicit system of convex constraints*

$$\left[\begin{array}{c|c|c} \tau - \rho^2 \sum_{j=1}^J \mu_j & & \beta^T[x] \\ \hline & \sum_{j=1}^J \mu_j Q_j & \alpha^T[x] \\ \hline \beta[x] & \alpha[x] & \tau I \end{array}\right] \succeq 0 \qquad (+)$$

in variables $x, \tau, \mu_1, ..., \mu_J$ *is a safe tractable approximation of* (RC) — *whenever* (x, τ) *can be extended to a feasible solution of* $(+)$, *this pair is feasible for* (RC).

(ii) *The approximation is precise when* $J = 1$, *and is tight within the factor* $O(1)\sqrt{\ln(J)}$ *otherwise.*

Tractable Reformulations/Tight Safe Tractable Approximations of Semi-Infinite Linear Matrix Inequalities

♣ Here we are interested in tractable reformulations/tight safe tractable approximations of a semi-infinite Linear Matrix Inequality

$$\mathcal{A}[\zeta,x]\succeq 0\ \forall\zeta\in\rho\mathcal{Z} \tag{RC}$$

where $\mathcal{A}[\zeta,x]$ is a symmetric matrix which is bi-affine in x and in ζ:

$$\mathcal{A}[\zeta,x]\equiv A_0[\zeta]+\sum_{j=0}^{n}x_jA_j[\zeta]\equiv\alpha_0[x]+\sum_{i=1}^{d}\zeta_i\alpha_i[x],$$

where the symmetric matrices $A_j[\zeta]$ and $\alpha_i[x]$ are affine in their arguments.

♠ As a matter of fact, aside of the "universally tractable" (and trivial) case of scenario-generated $\mathcal{Z}$, in the LMI case no "universally good" – leading to tractable or "nearly so" (i.e., admitting tight safe tractable approximations) – geometries of $\mathcal{Z}$ are known.

♠ All known to us generic tractability results impose structural restrictions on how the uncertainty enters the body of the LMI, specifically, they assume *norm-bounded* , *structured or not*. *model of perturbations*

$$\mathcal{A}[\zeta,x]\succeq 0\ \forall \zeta\in\rho\mathcal{Z} \tag{RC}$$

Definition: We say that (RC) is with *norm-bounded perturbations*, if

A. $\mathcal{Z}$ is comprised of collections $\zeta=\{\zeta^1,...,\zeta^L\}$ of matrices $\zeta^\ell\in\mathbb{R}^{p_\ell\times q_\ell}$ of the spectral norm not exceeding 1.

In addition, prescribed part of the matrices ζ^ℓ — those with indices from a given set $\mathcal{I}_{\rm s}$ — are marked as "scalar perturbation blocks" and should be *scalar* – proportional to the unit matrix (for those ℓ, of course, $p_\ell=q_\ell$). The remaining "full perturbation blocks" ζ^ℓ in ζ can be arbitrary $p_\ell\times q_\ell$ matrices of norm ≤ 1. Thus,

$$\mathcal{Z}=\left\{\zeta=\{\zeta^1,...,\zeta^L\}:\begin{array}{l}\zeta^\ell\in\mathbb{R}^{p_\ell\times q_\ell}\,\forall\ell\\ \|\zeta^\ell\|_{2,2}\leq 1\,\forall\ell\\ \zeta^\ell\in\mathbb{R}\cdot I_{p_\ell},\ell\in\mathcal{I}_{\rm s}\end{array}\right\}$$

B. The body $\mathcal{A}[\zeta,x]$ of (RC) is

$$\mathcal{A}[\zeta,x]=\mathcal{A}^0[x]+\sum_{\ell=1}^{L}\left[L_\ell^T[x]\zeta^\ell R_\ell[x]+R_\ell^T[x][\zeta^\ell]^T L_\ell[x]\right],$$

where $L_\ell[x]$, $R_\ell[x]$ are affine in x and for every ℓ at least one of these matrices is constant.

Note: W.l.o.g. we assume that $R_\ell[x]\equiv R_\ell$ for all ℓ.

• Norm-bounded uncertainty is called *structured*, if $L>1$, and *unstructured* otherwise.

♣ **Relation to norm-bounded uncertainty in c.q.i.**

♠ The Lorentz cone $\mathbf{L}^m$ can be represented as the intersection of the semidefinite cone $\mathbf{S}^m_+$ and a linear subspace. Specifically, by the Schur Complement Lemma

$$[u;t] \in \mathbf{L}^m \Leftrightarrow \mathrm{Arrow}(u,t) := \left[\begin{array}{c|c} t & u^T \\ \hline u & tI_{m-1} \end{array}\right] \succeq 0$$

♠ Consequently, a semi-infinite c.q.i. can be reformulated equivalently as a semi-infinite LMI. In particular, a semi-infinite Least Squares inequality is equivalent to an "arrow" semi-infinite LMI with uncertainty-affected off-diagonal entries of the first row and column:

$$\begin{array}{c} \|A[\zeta]x + b[\zeta]\|_2 \leq \tau \, \forall \zeta \in \rho\mathcal{Z} \\ \Leftrightarrow \mathcal{A}[\zeta,x] := \left[\begin{array}{c|c} \tau & [A[\zeta]x+b]^T \\ \hline A[\zeta]x+b & \tau I \end{array}\right] \succeq 0 \; \forall \zeta \in \rho\mathcal{Z} \end{array}$$

With this correspondence, *norm-bounded uncertainty in the semi-infinite Least Squares inequality:*

$$A[\zeta]x + b[\zeta] = A^0 x + b^0 + \sum_{\ell=1}^{L} L_\ell^T[x]\zeta^\ell R_\ell[x]$$

induces norm-bounded uncertainty with no scalar perturbation blocks in the semi-infinite LMI:

$$\begin{array}{rcl} \mathcal{A}[\zeta,x] & = & \mathrm{Arrow}(A^0x + b, \tau) \\ & & + \sum_{\ell=1}^{L} \left[\mathcal{L}_\ell^T[x]\zeta^\ell \mathcal{R}_\ell[x] + \mathcal{R}_\ell^T[x][\zeta^\ell]^T \mathcal{L}_\ell[x]\right]. \end{array}$$

♠ As a consequence, *every result on tractability/tight safe tractable approximation of semi-infinite LMI with unstructured or structured norm-bounded uncertainty induces similar results on semi-infinite Least Squares inequalities*, and this is how the results of the latter type we have mentioned in the "c.q.i.-part" were obtained.

Derivation of Tight Safe Tractable Approximation of Semi-Infinite LMI with Norm-Bounded Uncertainty

$$
\begin{array}{l}
A_0[x] + \sum_{\ell=1}^{L}[L_\ell^T[x]\zeta^\ell R_\ell + R_\ell^T[\zeta^\ell]^T L_\ell[x]] \succeq 0 \\
\forall \{\zeta^\ell \in \mathbb{R}^{p_\ell \times q_\ell}\} : \|\zeta^\ell\|_{2,2} \leq \rho \forall \ell,\ \zeta^\ell \in \mathbb{R}\cdot I_{p_\ell} \forall \ell \in \mathcal{I}_{\rm s}
\end{array}
\qquad \text{(RC)}
$$

The idea is pretty simple: an evident *sufficient* condition for x to be feasible for (RC) is existence of matrices Y_ℓ such that

$$
\begin{array}{ll}
(a) & Y_\ell \succeq [L_\ell^T[x]\zeta^\ell R_\ell + R_\ell^T[\zeta^\ell]^T L_\ell[x]] \\
& \qquad \forall \zeta^\ell = \lambda_\ell I_{p_\ell} : \|\zeta_\ell\|_{2,2} \leq 1,\ \ell \in \mathcal{I}_{\rm s} \\
(b) & Y_\ell \succeq [L_\ell^T[x]\zeta^\ell R_\ell + R_\ell^T[\zeta^\ell]^T L_\ell[x]] \\
& \qquad \forall \zeta^\ell : \|\zeta_\ell\|_{2,2} \leq 1,\ \ell \notin \mathcal{I}_{\rm s} \\
(c) & A_0[x] - \rho \sum_{\ell=1}^{L} Y_\ell \succeq 0
\end{array}
\qquad (!)
$$

• "Semi-infinite" LMIs (a) are equivalent to finite system of LMIs

$$Y_\ell \succeq \pm[L_\ell^T[x]R_\ell + R_\ell^T L_\ell[x]],\ \ell \in \mathcal{I}_{\rm s} \qquad (A)$$

• From $\mathcal{S}$-Lemma one can derive without much thought that semi-infinite LMIs (b) can be represented equivalently by finite system of LMIs

$$\left[\begin{array}{c|c} Y_\ell - \lambda_\ell R_\ell^T R_\ell & L_\ell^T[x] \\ \hline L_\ell[x] & \lambda I_{p_\ell} \end{array}\right] \succeq 0,\ \ \ell \notin \mathcal{I}_{\rm s} \qquad (B)$$

$$\begin{array}{l} A_0[x] + \sum_{\ell=1}^{L}[L_\ell^T[x]\zeta^\ell R_\ell + R_\ell^T[\zeta^\ell]^T L_\ell[x]] \succeq 0 \\ \forall \{\zeta^\ell \in \mathbb{R}^{p_\ell \times q_\ell}\} : \|\zeta^\ell\|_{2,2} \leq \rho \forall \ell,\, \zeta^\ell \in \mathbb{R} \cdot I_{p_\ell} \forall \ell \in \mathcal{I}_{\rm s} \end{array} \qquad \text{(RC)}$$

We have arrived at the first part of the following

Theorem: **(i)** *The explicit system of LMIs*

$$\begin{array}{ll} (a) & Y_\ell \succeq \pm[L_\ell^T[x]R_\ell + R_\ell^T L_\ell[x]],\ \forall \ell \in \mathcal{I}_{\rm s} \\ (b) & \left[\begin{array}{c|c} Y_\ell - \lambda_\ell R_\ell^T R_\ell & L_\ell^T[x] \\ \hline L_\ell[x] & \lambda I_{p_\ell} \end{array}\right] \succeq 0,\ \ell \not\in \mathcal{I}_{\rm s} \\ (c) & A_0[x] - \rho \sum_{\ell=1}^{L} Y\ell \succeq 0 \end{array}$$

in variables x, $\lambda_\ell \in \mathbb{R}$, $\ell \not\in \mathcal{I}_{\rm s}$, Y_ℓ, $1 \leq \ell \leq L$, *is a safe tractable approximation of* (RC).

(ii) *This approximation is exact when* $L = 1$ (unstructured norm-bounded perturbations) *and is tight within factor* $\pi/2$ *when* $L > 1$, *provided that there are no nontrivial (with* $p_\ell > 1$*) scalar perturbation blocks.*

When there are nontrivial scalar perturbation blocks, the tightness factor of the approximation is a universal function $\vartheta(\mu)$ *of* $\mu = 2\max_{\ell \in \mathcal{I}_{\rm s}} p_\ell$ *such that*

$$\vartheta(\mu) \leq \sqrt{\pi\mu/2}.$$

♣ Example: Robust Truss Topology Design

• A *truss* is a mechanical construction, like electric mast, railroad bridge, or Eiffel Tower, comprised of thin elastic *bars* linked to each other at *nodes*.

• Under external load, the truss deforms until the internal forces caused by the deformation compensate the external forces. At the resulting static equilibrium, the truss capacitates certain energy, called *compliance*. Compliance is a natural measure of the rigidity of the truss w.r.t. a load: the less the compliance, the more rigid is the truss.

• In a Truss Topology Design problem, one is given

A. A *list of tentative nodes* – an m-point grid in $\mathbb{R}^d$ ($d = 2/d = 3$), along with *boundary conditions* which declare some nodes partially or completely fixed by supports, and thus define for every $i \leq m$ a linear subspace $V_i \subset \mathbb{R}^d$ of *virtual displacements* of node i. A *virtual displacement of the nodal set* is a collection of allowed displacements of the nodes, and these virtual displacements form a linear space $\mathcal{V} = \bigoplus\limits_{i=1}^{m} V_i$;

B. A collection of n *tentative bars* – pairs of nodes which can be linked by a bar;

C. A finite set $\mathcal{F}$ of *loading scenarios* – vectors $f \in \mathcal{V}$ comprised of external forces acting at the nodes and representing the load in question.

The goal in a TTD problem is *to assign the with non-negative volumes* $t = [t_1; ...; t_n] \in \mathcal{T}$ *in a way which minimizes the worst, over loading scenarios, compliance of the construction w.r.t. a scenario.* Here $\mathcal{T} \subset \mathbb{R}^n_+$ is a given polytope of *admissible designs*, most typically – just the simplex $\{t \geq 0, \sum_i t_i \leq w\}$, where w is an a priori upper bound on the weight of the construction.

♠ Mathematically, the fact that the compliance of a truss t w.r.t. a load $f \in \mathcal{V}$ is $\leq \tau$ is expressed as

$$\left[\begin{array}{c|c} 2\tau & f^T \\ \hline f & A(t) \end{array}\right] \succeq 0,$$

where

$$A(t) = \sum_{i=1}^{n} t_i b_i b_i^T$$

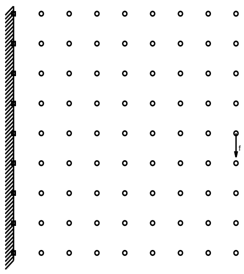
is the *stiffness matrix* of the truss; $b_i \in \mathcal{V}$ are readily given by the geometry of the nodal set.

Thus, the TTD problem is the semidefinite program

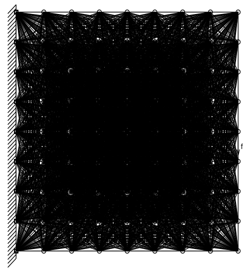
$$\min_{\tau,t} \left\{ \tau : \left[\begin{array}{c|c} 2\tau & f^T \\ \hline f & A(t) \end{array}\right] \succeq 0,\ f \in \mathcal{F} \right\}.$$

♠ In TTD, one starts with a dense nodal grid and allows for all pair connections of nodes with bars. At optimality, most of the bars get zero volume, thus revealing the *optimal topology* of the construction, along with its *optimal sizing*.

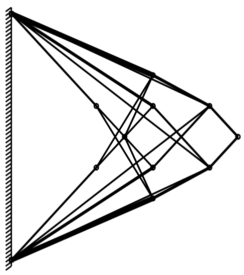
♠ The set of loading scenarios $\mathcal{F}$ usually is comprised of small (1-2-3) number of “loads of interest.” In reality, the truss will be subject to small “unforeseen occasional loads” which can crush it:



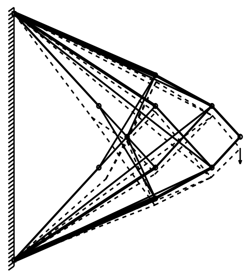
(a): 9×9 nodal grid with most left nodes fixed and the load of interest. $M =$ 144 degrees of freedom.



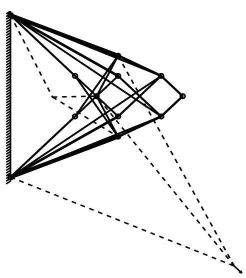
(b): 2,039 tentative bars



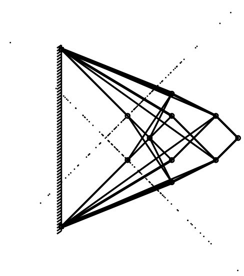
(c): Single-load optimal design, 12 nodes, 24 bars. Compliance w.r.t. load of interest 1.00.



(d): Deformation of nominal design under the load of interest f.



(e): Deformation of nominal design under “occasional” load 10^8 times less than f.



(f): “Dotted lines”: positions of nodes in deformed nominal design, sample of 100 loads $\sim \mathcal{N}(0, 10^{-16} I_{20})$

♠ To avoid potential instability of the designed truss w.r.t. small occasional loads, it makes sense to control its compliance w.r.t. both loads of actual interest and *all "occasional" loads of norm* $\leq \rho$.

• **Question:** *Where should the occasional loads be applied?* Usually, most of the nodes from the original grid are not used in the resulting construction; why to bother about forces acting at *nonexisting* nodes?

Possible answer: When "robustifying" the nominal design, we can choose, as the nodal set, the nodes actually used in this design and to allow the occasional loads to act at all these nodes.

• **Question:** *How to choose* ρ*?*

Possible answer: Let τ_* be the nominal optimal value in the TTD problem, and let $\tau^+ > \tau_*$ be the compliance "we are ready to tolerate." When robustifying the nominal design, we can look for the design which ensures, *for as large* ρ *as possible*, that the compliance w.r.t. the loads of interest *and all occasional loads* g *with* $\|g\|_2 \leq \rho$ is $\leq \tau^+$.

♠ The resulting Robust TTD problem reads

$$\rho_* = \max_{\rho,t}\left\{\rho : \begin{array}{l} \left[\begin{array}{c|c} 2\tau^+ & f^T \\ \hline f & A(t)\end{array}\right] \succeq 0 \ \forall f\in\mathcal{F} \\ \left[\begin{array}{c|c} 2\tau^+ & \zeta^T \\ \hline \zeta & A(t)\end{array}\right] \succeq 0 \ \forall \zeta, \|\zeta\|_2\leq\rho \\ t\in\mathcal{T}\end{array}\right\}. \qquad (*)$$

Note: Let $M = \dim\zeta$ be the number of degrees of freedom in the (reduced) nodal set. The body of the semi-infinite LMI in $(*)$ is

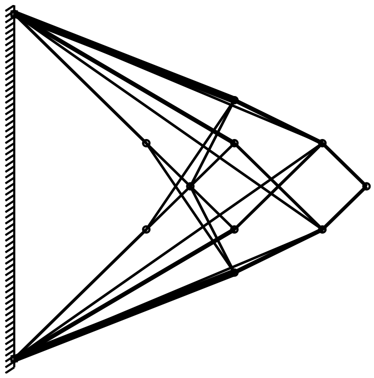
$$\left[\begin{array}{c|c} 2\tau^+ & f^T \\ \hline f & A(t)\end{array}\right] = \left[\begin{array}{c|c} 2\tau^+ & \\ \hline & A(t)\end{array}\right] + \overbrace{[0_{M\times 1}, I_M]^T}^{L^T}\zeta\overbrace{[1, 0_{1\times M}]}^{R} + R^T\zeta^T L,$$

with $\zeta\in\mathbb{R}^{m\times 1}$, so that $\|\zeta\|_{2,2}=\|\zeta\|_2$

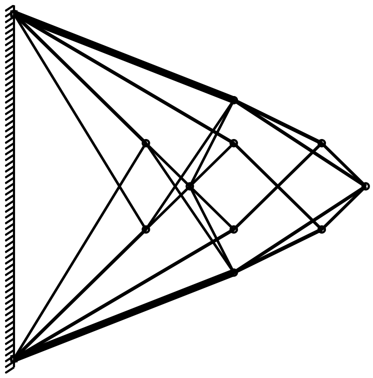
$\Rightarrow$ *We are in the case of unstructured norm-bounded uncertainty and thus $(*)$ admit a tractable reformulation*, namely,

$$\max_{\rho,t,s}\left\{\rho : \begin{array}{l} \left[\begin{array}{c|c} 2\tau^+ & f^T \\ \hline f & A(t)\end{array}\right] \succeq 0 \ \forall f\in\mathcal{F} \\ A(t)\succeq sI_M, \left[\begin{array}{c|c} 2\tau^+ & \rho \\ \hline \rho & s\end{array}\right]\succeq 0 \\ t\in\mathcal{T}\end{array}\right\}. \qquad (!)$$

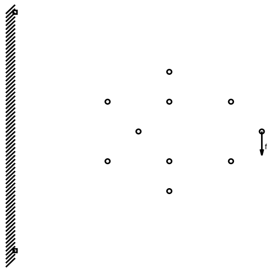
♠ Applying the outlined approach in the console design, we set $\tau^+ = 1.025\tau_* = 1.025$ and end up with $\rho_* = 0.362$. The resulting design, being only marginally inferior to the nominally optimal one as far as the load of interest is concerned, is incomparably more rigid w.r.t. occasional loads.



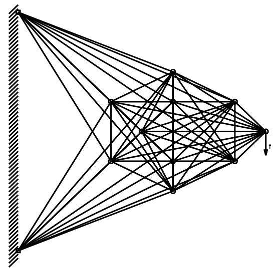
nominal design



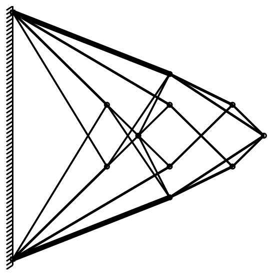
robust design



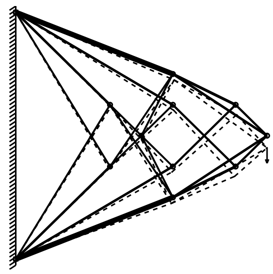
(a): reduced 12-node set with most left nodes fixed and the load of interest. $M = 20$ degrees of freedom.



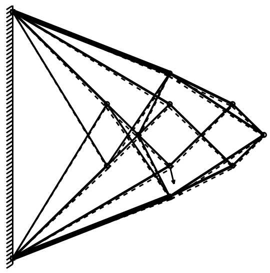
(b): 54 tentative bars



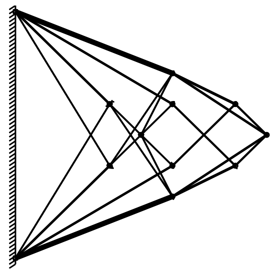
(c): Robust optimal design, 12 nodes, 24 bars. Compliance w.r.t. load of interest 1.025.



(d): Deformation of robust design under the load of interest.



(e): Deformation of robust design under "occasional" load 10 times less than the load of interest.



(f): "Bold dots": positions of nodes in deformated robust design over 100 loads $\sim \mathcal{N}(0, 10^{-2} I_{20})$

Robust design of a console

Applications in Robust Control

♣ Semi-infinite LMIs arise on numerous occasions in *Robust Control*.

Example: Lyapunov Stability Analysis. Consider an *uncertain time-varying linear dynamic system*

$$\dot{x}(t) = A_t x(t),$$

where A_t for all t is known to belong to a given convex compact set $\mathcal{U}$. How to certify that the system is *stable*, that is, all trajectories of (all realizations of) the system go to 0 as $t \to \infty$?

♠ The standard *sufficient* stability condition is existence of *Lyapunov Stability Certificate:* a matrix X such that

$$X \succ 0 \ \& \ A^T X + XA \prec 0 \ \forall A \in \mathcal{U}.$$

Indeed, given and LSC X, by compactness of $\mathcal{U}$

$$\exists \alpha > 0 : A^T X + XA \preceq -\alpha X \ \forall A \in \mathcal{A}$$

whence for every trajectory

$$\frac{d}{dt}[x^T(t)Xx(t)] = x^T(t)[A_t^T X + XA_t]x(t) \leq -\alpha[x^T(t)Xx(t)]$$
$$\Downarrow$$
$$x^T(t)Xx(t) \leq \exp\{-\alpha t\}x^T(0)Xx(0) \Rightarrow x(t) \to 0, \ t \to \infty.$$

♠ **Note:** Existence of an LSC is equivalent to the feasibility of the semi-infinite *Lyapunov LMI*

$$X \succeq I, \ \ A^T X + XA \preceq -I \ \forall A \in \mathcal{U}$$

$$X \succeq I,\ A^T X + XA \preceq -I\ \forall A \in \mathcal{U}$$

♠ In many Control applications, $\mathcal{U}$ is given by norm-bounded perturbations:

$$\mathcal{U} = \mathcal{U}_\rho = A^{\mathrm{n}} + \rho \mathcal{Z},$$

$$\mathcal{Z} = \left\{ A = \sum_{\ell=1}^L P_\ell \zeta^\ell Q_\ell : \begin{array}{c} \zeta^\ell \in \mathbb{R}^{p_\ell \times q_\ell} \\ \|\zeta^\ell\|_{2,2} \leq 1\ \forall \ell \\ \zeta^\ell \in \mathbb{R}I,\ \ell \in \mathcal{I}_{\mathrm{s}} \end{array} \right\}$$

Example 1: Interval uncertainty. In this case

$$\mathcal{U}_\rho = \{A : |A_{ij} - A^{\mathrm{n}}_{ij}| \leq \rho d_{ij}\},$$

that is,

$$\mathcal{Z} = \{Z = \sum_{i,j} d_{ij} e_i \zeta^{ij} e_j^T : |\zeta^{ij}| \leq 1\, \forall i, j\}$$

(full 1×1 perturbation blocks).

Example 2: Closed loop dynamical system. Consider a time invariant Linear Dynamical system "closed" by a linear feedback control:

$$\begin{array}{rcl}\dot{x}(t) & = & Px(t) + Qu(t) + T\xi(t) \text{ [state equations]}\\ y(t) & = & Rx(t) + U\xi(t) \text{ [observed outputs]}\\ u(t) & = & Sy(t) \text{ [feedback]}\\ \Rightarrow & & \\ \dot{x}(t) & = & [P + QSR]x(t) + [T + QSU]\xi(t) \text{ [closed loop]}\end{array}$$

When one (or more) of the matrices P, Q, R, S drifts with time around its nominal value, the system becomes uncertain. Assuming norm-bounded uncertainty in P, Q, R, S:

$$\|\zeta^P := P_t - P\|_{2,2} \leq \rho d_P, ..., \|\zeta^S := S_t - S\|_{2,2} \leq \rho\delta_S$$

we can approximate the range $\mathcal{U}_\rho$ of the matrix

$$A_t = P_t + Q_t S_t R_t$$

of the closed loop system by the norm-bounded perturbation set:

$$\begin{array}{c}\mathcal{U}_\rho \approx [P + QSR] + \rho\mathcal{Z},\\ \mathcal{Z} \;=\; \{\zeta^P + \zeta^Q SR + Q\zeta^S R + QS\zeta^R : \\ \|\zeta^P\|_{2,2} \leq 1, ..., \|\zeta^S\|_{2,2} \leq 1\}\end{array}$$

(this approximation is exact when only one of the matrices P, Q, R, S is subject to drift).

$$\mathcal{A}[A,X] := -I - A^T X - XA \succeq 0 \;\forall A \in \mathcal{U}_\rho \tag{L}$$

$$\begin{array}{c} \mathcal{U}_\rho = A^{\rm n} + \rho\mathcal{Z}, \\ \mathcal{Z} = \left\{ \sum_{\ell=1}^L P_\ell \zeta^\ell Q_\ell : \begin{array}{c} \zeta^\ell \in \mathbb{R}^{p_\ell \times q_\ell} \\ \|\zeta^\ell\|_{2,2} \leq 1 \;\forall \ell \\ \zeta^\ell \in \mathbb{R}I,\; \ell \in \mathcal{I}_{\rm s} \end{array} \right\} \end{array} \tag{NB}$$

Observation: *Norm-bounded uncertainty* (NB) *induces norm-bounded uncertainty in the Lyapunov LMI* (L), *the number, sizes and types (full/scalar) of the perturbation blocks being preserved.*

Corollary: *When $L = 1$* (unstructured norm-bounded uncertainty, case A), (L) *admits a tractable reformulation, otherwise:*

— *when there are no nontrivial ($p_\ell > 1$) scalar perturbations blocks* (case B), (L) *admits tight within the factor $\pi/2$ safe tractable approximation,*

— *otherwise* (case C) (L) *admits a safe tractable approximation tight within factor $\vartheta(\mu) \leq \sqrt{\pi\mu/2}$,* $\mu = 2\max\limits_{\ell \in \mathcal{I}_{\rm s}} p_\ell$.

♡ In particular, *The Lyapunov Stability Radius of* $(A^{\rm n}, \mathcal{Z})$ – *the largest ρ for which* (L) *has a positive definite solution X* – *admits an efficiently computable lower bound which is*

— *exact in case A,*

— *tight within the factor $\pi/2$ in case B,*

— *tight within the factor $\vartheta(\mu)$ in case C.*

Globalized Robust Counterparts

♣ We are about to reconsider two of the basic assumptions on the "decision environment" we made, namely, that

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 are *hard* — we cannot tolerate violations of constraints, even small ones, when the data is in $\mathcal{U}$.

E.g., the mail traffic (or shopping) around Christmas is much higher than during the rest of the year. When following A.2-3 literally, when designing the corresponding service capacities, we should

— either completely ignore Christmas and orient ourselves towards the most-of-the-year load on the system,

— or design the system as if the Christmas load could happen every day.

• Both these extremes hardly are wise.

♠ **What we want now is to "immunize" against data uncertainty in the case when**

— the data are allowed to run out of the uncertainty set, and

— we allow for *controlled* **violation of the constraints when it happens.**

♠ **Pursuing this goal, we, as always, can focus on a** *single* **uncertainty-affected conic constraint, which now is convenient to write down in the form of**

$$A[\zeta]x + b[\zeta] \in \mathbf{Q} = \left\{ y : \begin{array}{c} P_i y + p_i \in \mathbf{K}_i, \\ 1 \leq i \leq I \end{array} \right\} \subset F \qquad \text{(UCC)}$$

$A[\zeta]$, $b[\zeta]$ **are affine in their arguments.**

♣ **We assume that the "physically possible" perturbations ζ run through the set**

$$\mathcal{Z} + \mathcal{L} \subset E = \mathbb{R}^L, \qquad \text{(Pert)}$$

- $\mathcal{Z}$**: closed convex "normal range" of ζ**
- $\mathcal{L}$**: closed convex cone.**

Definition: x is robust feasible for (UCC), (Pert) *with global sensitivity* α, **if**

$$\mathrm{dist}(A[\zeta]x + b[\zeta], \mathbf{Q}) \leq \alpha \mathrm{dist}(\zeta, \mathcal{Z}|\mathcal{L}) \ \forall \zeta \in \mathcal{Z} + \mathcal{L}$$

$$\left[\begin{array}{c} \mathrm{dist}(y, \mathbf{Q}) = \min\limits_{y' \in \mathbf{Q}} \|y - y'\|_F; \\ \mathrm{dist}(z, \mathcal{Z}|\mathcal{L}) = \inf\limits_{z'} \{\|z - z'\|_E : z' \in \mathcal{Z}, z - z' \in \mathcal{L}\} \end{array} \right]$$

$$A[\zeta]x + b[\zeta] \in \mathbf{Q} \tag{UCC}$$

x is robust feasible for (UCC) with global sensitivity α
$\Leftrightarrow \mathrm{dist}(A[\zeta]x + b[\zeta], \mathbf{Q}) \leq \alpha \mathrm{dist}(\zeta, \mathcal{Z}|\mathcal{L})\ \forall \zeta \in \mathcal{Z} + \mathcal{L}$

Clearly, *if x is robust feasible for* (UCC) *with a whatever global sensitivity, x is robust feasible for* (UCC), *the uncertainty set being $\mathcal{Z}$.*

♠ **Given an uncertain conic problem with certain objective**

$$\mathcal{P} = \left\{ \min_x \left\{ c^T x : A_i[\zeta]x + b_i[\zeta] \in \mathbf{Q}_i,\ i = 1, ..., m \right\}_\zeta \right\}$$

and a *perturbation structure*, **that is, $\mathcal{Z}$, $\mathcal{L}$ and norms used to measure the participating distances, the** *Globalized Robust Counterpart* **of the uncertain problem is the semi-infinite conic problem**

$$\min_x \left\{ c^T x : \begin{array}{r} \mathrm{dist}(A_i[\zeta]x + b_i[\zeta], \mathbf{Q}_i) \leq \alpha_i \mathrm{dist}(\zeta, \mathcal{Z}|\mathcal{L}) \\ \forall \zeta \in \mathcal{Z} + \mathcal{L} \end{array} \right\} \tag{GRC}$$

where $\alpha_i \geq 0$ are given parameters.

♡**Alternatively, we can treat x** *and* **α_i as the decision variables and to optimize a new convex objective, depending both on x and α_i, under the same semi-infinite constraints as in (GRC), and, perhaps, additional certain constraints on x and α_i.**

♠ **Sometimes it makes sense to "add some structure" to the perturbations, specifically, to assume that**

$$\zeta = [\zeta^1;...;\zeta^k] \in \underbrace{[\mathcal{Z}_1 \times ... \times \mathcal{Z}_K]}_{\mathcal{Z}} + \underbrace{[\mathcal{L}_1 \times ... \times \mathcal{L}_K]}_{\mathcal{L}}$$

($\mathcal{Z}_k$ **are closed convex sets,** $\mathcal{L}_k$ **are closed convex cones) and to define the GRC of an uncertain conic constraint**

$$A[\zeta]x + b[\zeta] \in \mathbf{Q}$$

as

$$\forall \zeta \in \mathcal{Z} + \mathcal{L} : \mathrm{dist}(A[\zeta]x + b[\zeta], \mathbf{Q}) \leq \sum_{k=1}^{K} \alpha_k \mathrm{dist}(\zeta^k, \mathcal{Z}_k | \mathcal{L}_k).$$

GRC of Scalar Linear Inequality

♣ **Consider the GRC of an uncertain** *scalar linear inequality*

$$\mathrm{dist}(a^T[\zeta]x + b[\zeta], \mathbb{R}_-) \leq \alpha \mathrm{dist}(\zeta, \mathcal{Z}|\mathcal{L}) \;\forall \zeta \in \mathcal{Z} + \mathcal{L} \tag{ULC}$$

♠ **Since** $a[\zeta]$**,** $b[\zeta]$ **are affine in** ζ**, we have**

$$a[\zeta]x + b[\zeta] \equiv \omega^T[x]\zeta + \gamma[x]$$

with $\omega[x]$, $\gamma[x]$ *affine in* x.

Theorem: Semi-infinite inequality (ULC) is equivalent to the pair of semi-infinite inequalities

$$\begin{array}{ll} (a) & \omega^T[x]\zeta + \gamma[x] \leq 0 \;\forall \zeta \in \mathcal{Z} \\ (b) & \omega^T[x]\zeta \leq \alpha \;\forall \zeta \in \mathcal{L}_1 = \{\zeta \in \mathcal{L} : \|\zeta\|_E \leq 1\}. \end{array}$$

In particular, (GRC) is computationally tractable, provided that $\mathcal{Z}$, $\mathcal{L}$ and $\|\cdot\| = \|\cdot\|_E$ are so.

♠ **Illustration: $\|\cdot\|_\infty$ Antenna Design via GRC.**

• **$\|\cdot\|_\infty$ Antenna Design is the uncertain LO**

$$\left\{\min_{x,\tau}\left\{\tau : -\tau\mathbf{1} \leq d - D(I+\mathrm{Diag}\{\zeta\})x \leq \tau\mathbf{1}\right\} : \|\zeta\|_\infty \leq \rho\right\}$$
$$[D : m\times n]$$

♠ **The** *robust performance* **of a design x is**

$$F_x(\rho) = \max_{\zeta:\|\zeta\|_\infty\leq\rho} \|d - D(I+\mathrm{Diag}\{\zeta\})x\|_\infty.$$

This is a convex nondecreasing function of ρ.

♠ **Let us fix an uncertainty level $\bar\rho \geq 0$ and set $\mathcal{Z} = \{\zeta : \|\zeta\|_\infty \leq \bar\rho\}$, $\mathcal{L} = \mathbb{R}^n$, $\|\zeta\|_E \equiv \|\zeta\|_\infty$.**

Note: *A pair (x,τ) is robust feasible, with global sensitivity α, for the Antenna Design problem iff $\tau \geq F_x(\bar\rho)$ and*

$$\alpha \geq \alpha(x) := \lim_{\rho\to\infty}\frac{d}{d\rho}F_x(\rho) = \max_{1\leq i\leq m}\sum\nolimits_{j=1}^n |D_{ij}||x_j|$$

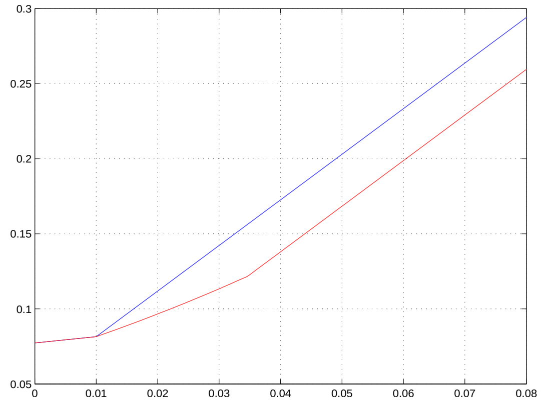
$\Rightarrow$ $F_x(\bar\rho)$, $\alpha(x)$ **imply a "global" upper bound on the robust performance of x:**

$$\forall \rho \geq 0 : F_x(\rho) \leq F_x(\bar\rho) + \alpha(x)\max\left[0, \rho - \bar\rho\right].$$

Invoking the easily computable quantity $F_x(0)$, we can improve this bound to

$$F_x(\rho) \leq \begin{cases} \frac{\bar\rho-\rho}{\bar\rho}F_x(0) + \frac{\rho}{\bar\rho}F_x(\bar\rho), & 0 \leq \rho < \bar\rho \\ F_x(\bar\rho) + \alpha(x)[\rho - \bar\rho], & \rho \geq \bar\rho \end{cases}$$

$$F_x(\rho) \leq \begin{cases} \frac{\bar{\rho}-\rho}{\bar{\rho}} F_x(0) + \frac{\rho}{\bar{\rho}} F_x(\bar{\rho}), & 0 \leq \rho < \bar{\rho} \\ F_x(\bar{\rho}) + \alpha(x)[\rho - \bar{\rho}], & \rho \geq \bar{\rho} \end{cases} \qquad \text{(UB)}$$



Robust performance of a design
and its upper bound (UB) with $\bar{\rho} = 0.01$

♠ **When solving Antenna Design problem, our ideal goal would be to optimize in x the robust performance $F_x(\rho)$** *for all ρ simultaneously*, **which of course is impossible.**

♠ **With the usual RC approach** *we fix the uncertainty level $\rho = \bar{\rho}$ and optimize the robust performance at this level.* **This makes sense when we know reasonably well the uncertainty level we intend to work with.**

♠ *When the range of possible uncertainty levels is wide, it can be more to the point to look for a "good global upper bound"* (UB) *on the robust performance*.

Example: Given a "reference uncertainty level" $\bar{\rho}$, we can act as follows:

- **We solve the RC of the problem, thus finding the best robust performance** $\phi(\bar{\rho}) = \min_x F_x(\bar{\rho})$ **at the reference uncertainty level** $\rho = \bar{\rho}$**;**
- **We then** *allow for controlled deterioration of the robust performance at the reference uncertainty level and choose among the corresponding designs the one with the smallest global sensitivity*, **i.e. solve the problems**

$$x_\delta \in \operatorname*{Argmin}_x \{\alpha(x) : F_x(\bar{\rho}) \leq (1+\delta)\phi(\bar{\rho})\} \qquad (P_\delta)$$

for several values of δ**.** *The larger δ, the worse is the robust performance of x_δ at the uncertainty level $\bar{\rho}$*, **and** *the smaller is the* (upper bound on the) *rate at which the robust performance deteriorates when $\rho \geq \bar{\rho}$ grows.*

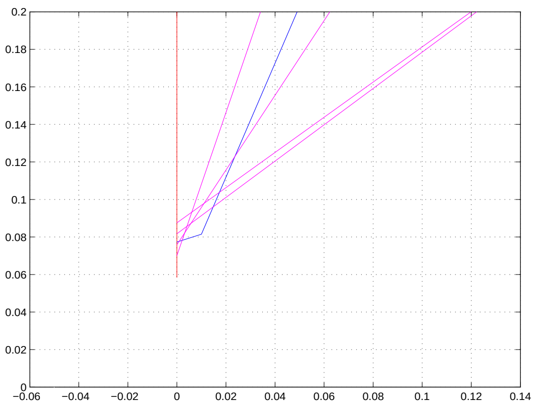
- **The arising tradeoff "robust performance at the reference uncertainty level" vs. "deterioration of robust performance as the uncertainty level grows" can be resolved by the end-user.**

Illustration: Set $\bar\rho = 0$. **Here** (P_δ) **is a simple LO program:**

$$x_\delta \in \mathop{\mathrm{Argmin}}_x \left\{ \max_{1\le i\le m} \sum_{j=1}^n |D_{ij}||x_j| : \right.$$
$$\left. \|d - Dx\|_\infty \le (1+\delta)\min_u \|d - Du\|_\infty,\ \forall i \right\}$$

♡ **Here are the optimal values in** (P_δ) **(≡ global sensitivities of the designs** x_δ**):**

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



Red and magenta: upper bounds $F_{x_\delta}(0) + \alpha(x_\delta)\rho$ on $F_{x_\delta}(\rho)$

Blue: upper bound on the robust performance of the robust optimal design associated with $\rho = 0.01$.

Globalized Robust Counterparts of Conic Constraints

♣ **Entity of interest:** semi-infinite conic constraint

$$\forall \zeta \in \mathcal{Z} + \mathcal{L} : \mathrm{dist}(A[\zeta]x + b[\zeta], \mathbf{Q}) \leq \sum_{k=1}^{K} \alpha_k \mathrm{dist}(\zeta^k, \mathcal{Z}_k | \mathcal{L}_k). \tag{GRC}$$

- $\mathbf{Q} = \{y : P_i y + q_i \in \mathbf{K}_i,\ 1 \leq i \leq I\}$
  $[\mathbf{K}_i : \mathbb{R}^{m_i}_+ / \mathbf{L}^{m_i} / \mathbf{S}^{m_i}_+]$
- $\mathcal{Z}_k \subset \mathbb{R}^{L_k}$: closed convex set
- $\mathcal{L}_k \subset \mathbb{R}^{L_k}$: closed convex cone

Note: $A[\zeta]x + b[\zeta] \equiv \sum_{k=1}^{K} \Omega_k[x]\zeta^k + \gamma[x]$ with $\Omega_k[x]$, $\gamma[x]$ affine in x.

♠ **Equivalent reformulation of** (GRC)

- **Recessive cone** $\mathrm{Rec}(X)$ **of a closed convex set** X**:** the set of all directions d such that $\bar{y} + td \in X\ \forall t > 0$ for some (and then for all) $\bar{y} \in X$.

Examples: A. X **is bounded** $\Rightarrow \mathrm{Rec}(X) = \{0\}$

B. X **is a cone** $\Rightarrow \mathrm{Rec}(X) = X$

Theorem: $(x, \alpha \geq 0)$ *is feasible for* (GRC) *if and only if*

$$\begin{array}{ll} (a) & A[\zeta]x + b[\zeta] \in \mathbf{Q}\ \forall \zeta \in \mathcal{Z}_1 \times \ldots \times \mathcal{Z}_K \\ (b) & \Psi_k(x) := \sup\limits_{\zeta^k \in \mathcal{L}_1^k} \mathrm{dist}(\Omega_k[x]\zeta^k, \mathrm{Rec}(\mathbf{Q})) \leq \alpha_k,\ k \leq K \\ & [\mathcal{L}_1^k = \{\zeta^k \in \mathcal{L}^k : \|\zeta^k\|_{(k)} \leq 1\}] \end{array}$$

♠ **We have reduced the Globalized Robust Counterpart** (GRC) **of an uncertain conic constraint to a pair of semi-infinite conic constraints** (a), (b), **and such a system not necessarily is tractable.** *What to do when* (a) - (b) *is intractable?*

As in the case of RC, assume that $0 \in \mathcal{Z}_k$, $k \le K$, **and embed** (GRC) **into a single-parametric family of semi-infinite conic constraints**

$$\begin{array}{l} \forall \zeta \in \rho\mathcal{Z} + \mathcal{L}: \\ \mathrm{dist}(\sum_{k=1}^K \Omega_k[x]\zeta^k + \gamma[x], \mathbf{Q}) \le \sum_{k=1}^K \alpha_k \mathrm{dist}(\zeta^k, \rho\mathcal{Z}_k | \mathcal{L}_k) \end{array} \quad \text{(GRC)}$$

$$\Updownarrow$$

$$\begin{array}{ll} (a_\rho) & A[\zeta]x + b[\zeta] \in \mathbf{Q} \;\; \forall \zeta \in \rho\mathcal{Z}_1 \times \ldots \times \rho\mathcal{Z}_K \\ (b) & \Psi_k(x) := \sup\limits_{\zeta^k \in \mathcal{L}_1^k} \mathrm{dist}(\Omega_k[x]\zeta^k, \mathrm{Rec}(\mathbf{Q})) \le \alpha_k, \; k \le K \\ & [\mathcal{L}_1^k = \{\zeta^k \in \mathcal{L}^k : \|\zeta^k\|_{(k)} \le 1\}] \end{array}$$

and let us look for *tight tractable approximations* **of** (a_ρ) **and** (b)**.**

$$
\begin{array}{ll}
(a_\rho) & A[\zeta]x + b[\zeta] \in \mathbf{Q}\,\forall \zeta \in \rho\mathcal{Z}_1 \times ... \times \rho\mathcal{Z}_K \\
(b) & \Psi_k(x) := \sup\limits_{\zeta^k \in \mathcal{L}_1^k} \mathrm{dist}(\Omega_k[x]\zeta^k, \mathrm{Rec}(\mathbf{Q})) \le \alpha_k,\ \forall k
\end{array}
\qquad (\mathrm{GRC}_\rho)
$$

• **We already know what a ϑ-tight safe tractable approximation of (a_ρ) is:** a system $\mathcal{S}_\rho$ of efficiently computable convex constraints on x and additional variables such that

— if $\rho \ge 0$ and x are such that x can be extended to a feasible solution of $\mathcal{S}_\rho$, x is feasible for (a_ρ);

— if $\rho \ge 0$ and x are such that x cannot be extended to a feasible solution of $\mathcal{S}_\rho$, x is not feasible for $(a_{\vartheta\rho})$

• *By definition,* **a** *safe κ-tight tractable approximation of (b) is a collection of efficiently computable convex upper bounds Φ_k on Ψ_k such that $\Phi_k(x) \le \kappa\Psi_k(x)$ for every x.*

♣ Replacing (a_ρ) with a ϑ-tight s.t.a. $\mathcal{S}_\rho$, and (b) with a κ-tight s.t.a. $\Phi_k(x) \le \alpha_k$, $k \le K$, we get a tractable system $\mathcal{S}_\rho^+$ of convex constraints on $x, \alpha_1, ..., \alpha_K$ and additional variables such that

— if ρ and (x, α) are such that (x, α) can be extended to a feasible solution of $\mathcal{Z}_\rho^+$, (x, α) is feasible for (GRC_ρ).

— if ρ and (x, α) are such that (x, α) cannot be extended to a feasible solution of $\mathcal{Z}_\rho^+$, $(x, \kappa^{-1}\alpha)$ is not feasible for $(\mathrm{GRC}_{\vartheta^{-1}\rho})$.

Tight Approximations of Ψ_k

♣ **Situation:** We are given

- Euclidean space F with closed convex cone K^F and norm $\|\cdot\|_F$
- Euclidean space E with closed convex cone K^E and norm $\|\cdot\|_E$
- A linear mapping $x \mapsto Ax : E \to F$.

♠ **Goal:** to build a tight efficiently computable convex upper bound $\Phi[A]$ on the function

$$\Psi[A] = \max_e \left\{ \mathrm{dist}_{\|\cdot\|_F}(Ae, K^F) : e \in K^E, \|e\|_E \leq 1 \right\}$$

Fact: The resulting problem admits a kind of *duality*.

♡ **Definition:** Let B be a Euclidean space with a norm $\|\cdot\|$. The conjugate to $\|\cdot\|$ norm is defined as

$$\|z\|_* = \max_{y:\|y\|\leq 1} \langle z, y \rangle$$

In fact, $\|\cdot\|_*$ is the smallest norm on B such that

$$\langle x, y \rangle \leq \|z\|_* \|y\| \ \forall y, z.$$

Example 1: When B is $\mathbb{R}^n$ with the standard inner product, $p \in [1, \infty]$ and $\|\cdot\| = \|\cdot\|_p$, one has $\|\cdot\|_* = \|\cdot\|_q$ with $\frac{1}{p} + \frac{1}{q} = 1$.

Example 2: When B is $\mathbb{R}^{m\times n}$ with the standard inner product, $p \in [1, \infty]$ and $\|Y\| = \|\sigma(Y)\|_p$, $\sigma(Y)$ being the vector of singular values of Y, one has $\|Z\|_* = \|\sigma(Z)\|_q$ with $\frac{1}{p} + \frac{1}{q} = 1$.

Fact: $(\|\cdot\|_*)_* = \|\cdot\|$.

$$\max_e \left\{\mathrm{dist}_{\|\cdot\|_F}(Ae, K^F) : e \in K^E, \|e\|_E \le 1\right\} \qquad (*)$$

Observe that

$$\mathrm{dist}_{\|\cdot\|_F}(Ae, K^F) = \max_f \left\{-\langle Ae, f\rangle_F : f \in K_*^F, \|f\|_{F,*} \le 1\right\}$$

$\Rightarrow$

$$\max_e \left\{\mathrm{dist}_{\|\cdot\|_F}(Ae, K^F) : e \in K^E, \|e\|_E \le 1\right\}$$

$$= \max_{e,f} \left\{ -\underbrace{\langle Ae, f\rangle_F}_{=\langle A^* f, e\rangle_E} : \left\{ \begin{array}{l} e \in K^E, \|e\|_E \le 1, \\ f \in K_*^F, \|f\|_{F,*} \le 1 \end{array} \right. \right\}$$

$$= \max_f \left\{\mathrm{dist}_{\|\cdot\|_{E,*}}(A^* f, K_*^E) : f \in K_*^F, \|f\|_{F,*} \le 1\right\}$$

Thus: *Quantity* $(*)$ *remains intact when one carries out the following transformation of its data:*

- $A \leftarrow A_+ := A^*$
- $E \leftarrow E_+ := F$, $K^E \leftarrow K_+^E := K_*^F$, $\|\cdot\|_E \leftarrow \|\cdot\|_E^+ := \|\cdot\|_{F,*}$
- $F \leftarrow F_+ := E$, $K^F \leftarrow K_+^F := K_*^E$, $\|\cdot\|_F \leftarrow \|\cdot\|_F^+ := \|\cdot\|_{E,*}$.

Conclusion: *"Good cases" of* $(*)$ — tuples

$$(E, K^E, \|\cdot\|_E, F, K^F, \|\cdot\|_F)$$

allowing for efficient computation (of a tight upper bound on) the quantity $(*)$ as a function of A — *are met in "conjugate pairs" with members of a pair linked to each other by the above transformation.*

$$E, K^E, B_E = \{\|e\|_E \leq 1\} \;\&\; F, K^F, B_F = \{\|f\|_F \leq 1\}$$
$$\Downarrow$$
$$\Psi[A] = \max_{e \in B_E \cap K^E} \mathrm{dist}_{\|\cdot\|_F}(Ae, K^F)$$

♣ From now on we assume that K^E, B_E, K^F, B_F are computationally tractable, and ask when $\Psi[\cdot]$ is efficiently computable (or admits tight computable upper bound).

0) [Trivial case] $\boxed{\begin{array}{c} K^E = \{0\} \\ \hline K^F = F \end{array}} \Rightarrow \Psi[A] \equiv 0$

1) $\boxed{\begin{array}{l} K^E = E, B_E = \mathrm{Conv}\{e_1, ..., e_N\} \\ \hline K^F = \{0\}, \|f\|_F = \max\limits_{i \leq N} |\langle f, f_i \rangle_F| \end{array}}$

$$\Rightarrow \boxed{\begin{array}{l} \Psi[A] = \max\limits_i \mathrm{dist}_{\|\cdot\|_F}(Ae_i, K^F) \\ \hline \Psi[A] = \max\limits_{i,e} \left\{ \langle Ae, f_i \rangle_F : \|e\|_E \leq 1, e \in K^E \right\} \end{array}}$$

Note: In the GRC context, K^E is the cone $\mathcal{L}$ (or $\mathcal{L}_k$), and K^F is the recessive cone of $\mathbf{Q}$. Therefore **1)** implies, e.g., that *the GRC of uncertain conic constraint is tractable when*

— $\mathcal{L} = E$ *and* $\|\cdot\|_E = \|\cdot\|_1$, *or when*

— $\mathbf{Q}$ *is bounded and* $\|\cdot\|_F = \|\cdot\|_\infty$

2) $K^E = E, K_F = \{0\}$ [self-conjugate case]. In the GRC context this relates to $\mathcal{L} = E$, $\mathbf{Q}$ bounded. Here

$$\Psi[A] = \max_e \{\|Ae\|_F : \|e\|_E \leq 1\} =: \|A\|_{E,F}.$$

♡ We know 3 generic cases when $\|A\|_{E,F}$ is efficiently computable:

- $B_E = \mathrm{Conv}\{e_1, ..., e_N\}$
  $\Rightarrow \|A\|_{E,F} = \max_i \|Ae_i\|_F$
- $B_F = \{f : \langle f, f_i \rangle_F \leq 1, i \leq N\}$
  $\Rightarrow \|A\|_{E,F} = \max\limits_{i,e} \{\langle A^* f_i, e\rangle_E : \|e\|_E \leq 1\}$
- $\|\cdot\|_E = \|\cdot\|_2$, $\|\cdot\|_F = \|\cdot\|_2$
  $\Rightarrow \|A\|_{E,F} = \sqrt{\lambda_{\max}(A^*A)}$

♡ When $\|\cdot\|_E = \|\cdot\|_p$, $\|\cdot\|_F = \|\cdot\|_r$, computing $\|A\|_{E,F}$ is provably NP-hard, provided that $p > r$. However, we have

Theorem [Nesterov] *When $p \geq 2 \geq r$, $\|A\|_{E,F}$ admits the efficiently computable upper bound*

$$\min_{\mu,\nu} \left\{ \frac{\|\mu\|_{\frac{p}{p-2}} + \|\nu\|_{\frac{r}{2-r}}}{2} : \begin{bmatrix} \mathrm{Diag}\{\mu\} & A^T \\ A & \mathrm{Diag}\{\nu\} \end{bmatrix} \succeq 0 \right\},$$

and this bound is tight within the factor $\left[\frac{2\sqrt{3}}{\pi} - \frac{2}{3}\right]^{-1} \approx 2.2936$.

• Depending on p, r, the tightness factor can be improved; e.g., *when $p = \infty, r = 2$, it reduces to* $\sqrt{\pi/2}$.

3)

$$F=\mathbb{R}^m,\ \|\cdot\|_F=\|\cdot\|_\infty,\ K^F=\left\{u:\ \begin{array}{l} u_i\geq 0, i\in I_+\\ u_i\leq 0, i\in I_-\\ u_i=0, i\in I_0\end{array}\right\}$$

$$E=\mathbb{R}^n,\ \|\cdot\|_E=\|\cdot\|_1,\ K^E=\left\{v:\ \begin{array}{l} v_j\geq 0, j\in J_+\\ v_j\leq 0, j\in J_-\\ v_j=0, j\in J_0\end{array}\right\}$$

$\Rightarrow$

$$\Psi[A]=\max_{u\in U}\max_{e\in K^E,\|e\|_E\leq 1} u^Te,$$

$$U=\{-f_i\}_{i\in I_+}\cup\{f_i\}_{i\in I_-}\cup\{\pm f_i\}_{i\in I_0}$$

[f_i: basic orths in F]

$$\Psi[A]=\max_{v\in V}\mathrm{dist}_{\|\cdot\|_F}(Av,K^F),$$

$$V=\{e_j\}_{j\in J_+}\cup\{-e_j\}_{j\in J_-}\cup\{\pm e_j\}_{j\notin(J_+\cup J_-\cup J_0)}$$

[e_j: basic orths in E]

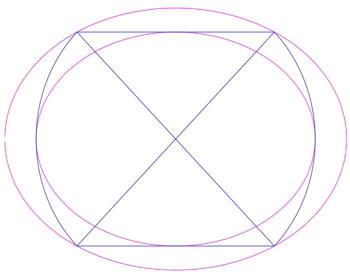
4)
$$\boxed{\begin{array}{l} F=\mathbb{R}^m,\ K^F=\mathbf{L}^m,\ \|\cdot\|_F=\|\cdot\|_2,\\ K^E=E=\mathbb{R}^n,\ \|\cdot\|_E=\|\cdot\|_2 \\ \hline F=\mathbb{R}^n,\ K^F=\{0\},\ \|\cdot\|_F=\|\cdot\|_2,\\ E=\mathbb{R}^m,\ K^E=\mathbf{L}^m,\ \|\cdot\|_E=\|\cdot\|_2 \end{array}}$$

$$\Rightarrow \boxed{\begin{array}{l} \Psi[A]\leq \|A^*\mathrm{Diag}\{\sqrt{\tfrac{3}{2}}I_{m-1},\tfrac{\sqrt{3}}{2}\}\|_{2,2}\leq \sqrt{3/2}\Psi[A] \\ \hline \Psi[A]\leq \|A\mathrm{Diag}\{\sqrt{\tfrac{3}{2}}I_{m-1},\tfrac{\sqrt{3}}{2}\}\|_{2,2}\leq \sqrt{3/2}\Psi[A] \end{array}}$$

Explanation: Assuming $E=\mathbb{R}^m$, $\|\cdot\|_E=\|\cdot\|_2$, $K^E=\mathbf{L}^m$, $F=\mathbb{R}^n$, $K^F=\{0\}$, $\|\cdot\|_F=\|\cdot\|_2$, **we have**

$$\Psi[A]=\max_v\left\{\|Av\|_2:\begin{array}{c}v\in\mathbf{L}^m,\\ \|v\|_2\leq 1\end{array}\right\}=\max_v\left\{\|Av\|_2:v\in B\right\},$$
$$B:=\mathrm{Conv}\{\{v\in\mathbf{L}^m,\|v\|_2\leq 1\}\cup\{v\in-\mathbf{L}^m,\|v\|_2\leq 1\}\},$$

$\Rightarrow$ $\Psi[A]$ *is the operator norm of* A *induced by the norm* $\|\cdot\|^E$ *with the unit ball* B *in* E *and the Euclidean norm in* F. **The norm** $\|v\|^E$ **can be approximated within the factor** $\sqrt{3/2}$ **by the Euclidean norm** $\sqrt{v_1^2+...+v_{m-1}^2+2v_m^2}$**:**



Blue: set B

and the result follows.

♠ **Illustration: Least Squares Antenna Design via GRC.**

• $\|\cdot\|_2$ **Antenna Design is the uncertain Least Squares problem with instances**

$$\min_{x,\tau}\left\{\tau : \underbrace{[h - H(I+\mathrm{Diag}\{\zeta\})x;\tau] \in \mathbf{Q} \equiv \mathbf{L}^{m+1}}_{\Leftrightarrow \|h-H(I+\mathrm{Diag}\{\zeta\})x\|_2 \le \tau}\right\}$$

• $H : m \times n$ • ζ: data perturbation

♠ **The** *robust performance* **of a design** x **is**

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

This is a convex nondecreasing function of ρ**.**

♠ **Let us fix an uncertainty level** $\bar\rho \ge 0$ **and set** $\|\cdot\|_F = \|\cdot\|_2$, $\mathcal{Z} = \{\zeta : \|\zeta\|_\infty \le \bar\rho\}$, $\mathcal{L} = \mathbb{R}^n$, $\|\cdot\|_E \equiv \|\cdot\|_\infty$.

Note: *A pair* (x,τ) *is robust feasible, with global sensitivity* α*, for the Least Squares Antenna Design problem, iff* $\tau \ge F_x(\bar\rho)$ *and*

$$\alpha \ge \alpha(x) := \max_{\|\zeta\|_\infty \le 1} \mathrm{dist}_{\|\cdot\|_2}([D[x]\zeta;0], \mathbf{L}^{m+1})$$
$$D[x] = H\mathrm{Diag}\{x\}$$

Observation: $\mathrm{dist}_{\|\cdot\|_2}([u;0];\mathbf{L}^{m+1}) = 2^{-1/2}\|u\|_2$.

Observation:

$$\lim_{\rho\to\infty} \frac{d}{d\rho}F_x(\rho) = \lim_{\rho\to\infty} F_x(\rho)/\rho = \max_{\|\zeta\|_\infty \le 1} \|D[x]\zeta\|_2.$$

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

Conclusion: similarly to the LO case, $F_x(0)$, $F_x(\bar\rho)$ and $\alpha(x)$ produce a *global upper bound* on $F_x(\cdot)$:

$$F_x(\rho) \le \begin{cases} \frac{\bar\rho-\rho}{\bar\rho}F_x(0)+\frac{\rho}{\bar\rho}F_x(\bar\rho), & 0\le\rho<\bar\rho \\ F_x(r\bar{ho})+2^{1/2}\alpha(x)[\rho-\bar\rho], & \rho\ge\bar\rho \end{cases}$$

$\Rightarrow$ when the uncertainty level $\bar\rho$ we should work with is only vaguely known, we can use the GRC methodology to optimize, to some extent, a *global* upper bound on the robust performance, similarly to what we did in the $\|\cdot\|_\infty$ Antenna Design.

Tractability Issues

♣ **In contrast to the LO case, now neither $F_x(\rho)$, nor $\alpha(x)$ is easy to compute. However,** *these quantities admit tight tractable upper bounds:*

- $$F_x(\rho) = \max_{\|\zeta\|_\infty} \| \underbrace{h - H(I+\mathrm{Diag}(\zeta))x}_{h-Hx+D[x]\zeta} \|_2 = \max_{\eta,t:\|[\eta;t]\|_\infty\leq 1} \|\rho D[x]\eta + t[h-Hx]\|_2 = \|[\rho D[x], h-Hx]\|_{\infty,2}.$$

By Nesterov's Norm Bound Theorem, the efficiently computable quantity

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

is a tight within the factor $\sqrt{\pi/2}$ upper bound on $F_x(\rho)$.

Note: same as $F_x(\rho)$, $\widehat{F}_x(\rho)$ is a convex nondecreasing function of $\rho \geq 0$.

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

$\Rightarrow$ the efficiently computable quantity

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

is a tight within the factor $\sqrt{\pi/2}$ upper bound on $\alpha(x)$.

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

♣ **Summary:** *the efficiently computable quantities* $\widehat{F}_x(\rho)$, $\widehat{\alpha}(x)$ *are tight, within the factor* $\sqrt{\pi/2}$, *upper bounds on* $F_\rho(x)$, $\alpha(x)$, *respectively*.

Since $\widehat{F}_x(\rho)$ **is convex and nondecreasing in** ρ **and** $\widehat{\alpha}(x) = 2^{-1/2}\lim\limits_{\rho\to\infty}\frac{d}{d\rho}\widehat{F}_x(\rho)$, **we have**

$$F_x(\rho) \le \begin{cases} \frac{\bar{\rho}-\rho}{\bar{\rho}}\widehat{F}_x(\bar{\rho}) + \frac{\rho}{\bar{\rho}}\widehat{F}_x(0), & 0 \le \rho < \bar{\rho} \\ \widehat{F}_x(\rho) + 2^{1/2}\widehat{\alpha}(x)[\rho-\bar{\rho}], & \rho \ge \bar{\rho} \end{cases}$$

and we can optimize, to some extent, the right hand side in order to ensure a desirable robust performance of our design in a wide range of values of ρ**.**

♠ **Illustration:** Setting $\bar{\rho} = 0$, we solve the problems

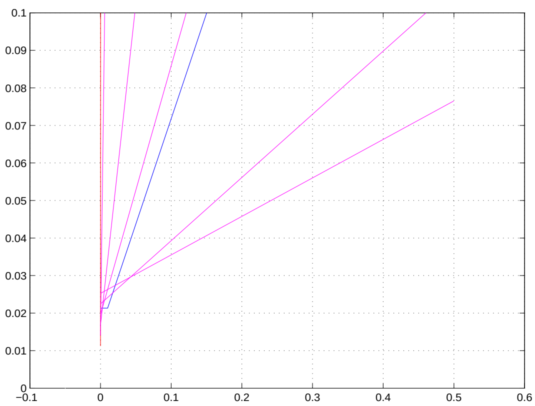
$$\beta(\delta) = \min_x \left\{ \widehat{\alpha}(x) : \widehat{F}_x(0) \equiv \|h - Hx\|_2 \leq (1+\delta) \min_u \|h - Hu\|_2 \right\}$$

for several values of δ, thus getting global upper bounds

$$\widehat{F}_{x_\delta}(0) + 2^{1/2}\beta(\delta)\rho \qquad (*)$$

on the robust performances of the resulting designs. The end-user could then choose the design he finds the most appropriate.

δ	0	0.25	0.50	0.75	1.00	1.25
$\beta(\delta)$	9441.4	14.883	1.7165	0.6626	0.1684	0.1025



Red and Magenta: bounds $(*)$ for the values of δ from the table.
Blue: Bound on global performance of the RC design corresponding to $\rho = 0.01$ (global sensitivity ≤ 0.3962).

Intermediate Summary

♣ **So far,**

♠ **We have defined the notion of** *uncertain conic problem* **– a family $\mathcal{P}$ of instances**

$$\min_x \left\{ c^T x : \begin{array}{c} A_1[\zeta]x + b_1[\zeta] \in \mathbf{K}_1 \\ \dots\dots\dots\dots \\ A_m[\zeta]x + b_m[\zeta] \in \mathbf{K}_m \end{array} \right\}, \zeta \in \mathcal{Z}$$

with the *data* $A_1, b_1, ..., A_m, b_m$ **affinely parameterized by the** *perturbation* ζ **running through a given** *perturbation set* $\mathcal{Z}$**. Here $\mathbf{K}_i$ are simple cones, specifically,** *nonnegative rays* **(uncertain LO), or** *Lorentz/Semidefinite* **cones (uncertain CQO/SDO), and $\mathcal{Z}$ w.l.o.g. can be assumed convex and closed.**

♠ **We associated with uncertain problem $\mathcal{P}$ its** *Robust Counterpart* **– the** *semi-infinite* **convex problem**

$$\min_x \left\{ c^T x : \left. \begin{array}{c} A_1[\zeta]x + b_1[\zeta] \in \mathbf{K}_1 \\ \dots\dots\dots\dots \\ A_m[\zeta]x + b_m[\zeta] \in \mathbf{K}_m \end{array} \right\} \forall \zeta \in \mathcal{Z} \right\} \tag{RC}$$

and treated the optimal solution of (RC) as the best "uncertainty-immunized" solution to the uncertain problem of interest.

$$\min_x \left\{ c^T x : \left. \begin{array}{c} A_1[\zeta]x + b_1[\zeta] \in \mathbf{K}_1 \\ \dots\dots\dots \\ A_m[\zeta]x + b_m[\zeta] \in \mathbf{K}_m \end{array} \right\} \forall \zeta \in \mathcal{Z} \right\} \qquad \text{(RC)}$$

♠ **The major theoretical issue we focused on was the one of** *computational tractability* **of the RC. We have seen that this crucial property**

— *always takes place in uncertain LO*, **provided that the perturbation set $\mathcal{Z}$ is computationally tractable,**

— **takes place in the case of** *scenario uncertainty* $\mathcal{Z} = \mathrm{Conv}\{\zeta^1, ..., \zeta^N\}$,

— *sometimes* **takes place in Uncertain CQO/SDO.**

In the case of Uncertain CQO/SDO, we have listed known "solvable cases" (same as "nearly solvable" ones – those where the RC admits a tight safe tractable approximation), and we have seen that these (nearly) solvable cases cover a reasonably wide variety of interesting and important applications.

• I believe building tractable reformulations/tight safe tractable approximations of the RC (or, which is the same, of semi-infinite conic constraints) is a rich, challenging and nontrivial research area.

Challenges: Adjustable Robust Optimization

♣ **Aside of applications of the RO methodology in various subject areas, an important venue of the RO-related research is** *extending the RO methodology beyond the scope of the RC approach as presented so far.*

The most important in this respect is, we believe, passing to *Adjustable Robust Optimization, where the decision variables are allowed to "adjust themselves"*, **to come extent,** *to the true values of the uncertain data.*

♠ **One of the central assumptions which led us to the notion of Robust Counterpart reads:**

> **A.1.** All decision variables in uncertain problem 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."

While being adequate to many decision making situations, A.1 is not a "universal truth."

♠ **In some cases,** *not all decision variables represent "here and now" decisions.* **In dynamical decision making some of the variables represent "wait and see" decisions and as such can depend on the portion of the true data which "reveals itself" before the moment when the decision is being made.**

Example: In an inventory affected by uncertain demand, there are no reasons to specify all replenishment orders in advance; the true time to specify the replenishment order of period t is the beginning of this period, and thus we can allow this order to depend on the actual demands in periods $1, ..., t-1$.

♠ **Usually,** *not all decision variables represent actual decisions;* **there exist also "analysis" (or slack) variables which do not represent decisions at all and are used to convert the problem into a desired form, e.g., one of a LO problem.** *Since the analysis variables do not represent actual decisions, why not to allow them to depend on the entire true data?*

Example: The convex constraint $\sum_i |a_i^T x - b_i| \leq \tau$ can be represented by a system of linear constraints

$$-y_i \leq a_i^T x - b_i \leq y_i, \sum_i y_i \leq \tau. \qquad (*)$$

When the data a_i, b_i are uncertain and x_j represent "here and now" decisions and thus should be assigned values independent of the true data, *there are absolutely no reasons to impose the same restriction on the slack variables y.* **To see the difference,**

• **The "true" RC of the uncertain constraint** $\sum_i |a_i^T[\zeta]x - b_i[\zeta]| \leq \tau$, $\zeta \in \mathcal{Z}$ **is**

$$\sum_i |a_i^T[\zeta]x - b_i[\zeta]| \leq \tau \ \forall \zeta \in \mathcal{Z},$$

and the "true" robust feasible set is

$$\{x : \forall \zeta \in \mathcal{Z} : \exists y : -y_i \leq a_i^T[\zeta]x - b_i[\zeta] \leq y_i, \sum_i y_i \leq \tau\} \qquad (1)$$

• **The RC of the uncertain system** $(*)$ **is**

$$-y_i \leq a_i^T[\zeta]x - b_i[\zeta] \leq y_i, \sum_i y_i \leq \tau \ \forall \zeta \in \mathcal{Z},$$

and the robust feasible set is

$$\{x : \exists y : \forall \zeta \in \mathcal{Z} : -y_i \leq a_i^T[\zeta]x - b_i[\zeta] \leq y_i, \sum_i y_i \leq \tau\} \qquad (2)$$

(2) **is smaller than** (1)**, and the difference can be dramatic:**

$$|x + \zeta| + |x - \zeta| \leq 2, \ \zeta \in [-1, 1] \Rightarrow \begin{cases} (1) = \{-1 \leq x \leq 1\} \\ (2) = \{0\} \end{cases}$$

$$\mathcal{P} = \left\{ \min_x \{ c^T[\zeta]x + d[\zeta] : \sum_{j=1}^{n} x_j A_j[\zeta] \leq b[\zeta] \} : \zeta \in \mathcal{Z} \right\}$$

Adjustable and Affinely Adjustable Robust Counterpart

♣ **In order to allow for the decision variables in $\mathcal{P}$ to "adjust themselves," to some extent, to the true values of the uncertain data, we could act as follows:**

• **We fix matrices P_j and allow the decision variable x_j to be an arbitrary function of the "portion" $P_j\zeta$ of the true data: $x_j = X_j(P_j\zeta)$**

• **We plug the** *decision rules* $X_j(P_j\zeta)$ **into $\mathcal{P}$ and require them to be** *robust feasible*:

$$\sum_{j=1}^{n} X_j(P_j\zeta) A_j[\zeta] \leq b[\zeta] \;\forall \zeta \in \mathcal{Z}$$

• **We associate with uncertain problem $\mathcal{P}$ its** *Adjustable Robust Counterpart*

$$\min_{t, X_j(\cdot)} \left\{ t : \begin{array}{l} \sum_j c_j[\zeta] X_j(P_j\zeta) + d[\zeta] \leq t \\ \sum_j X_j(P_j\zeta) A_j[\zeta] \leq b[\zeta] \end{array} \right\} \forall \zeta \in \mathcal{Z} \right\} \quad \text{(ARC)}$$

Note: When the decision rules $X_j(\cdot)$ are restricted to be constant, (ARC) recovers the usual RC of $\mathcal{P}$.

$$\mathcal{P} = \left\{ \min_x \{c^T[\zeta]x + d[\zeta] : \sum\nolimits_{j=1}^n x_j A_j[\zeta] \leq b[\zeta]\} : \zeta \in \mathcal{Z} \right\} \Rightarrow$$

$$\min_{t, X_j(\cdot)} \left\{ t : \begin{array}{l} \sum_j c_j[\zeta] X_j(P_j\zeta) + d[\zeta] \leq t \\ \sum_j X_j(P_j\zeta) A_j[\zeta] \leq b[\zeta] \end{array} \right\} \forall \zeta \in \mathcal{Z} \Bigg\} \quad \text{(ARC)}$$

♠ **While perfectly well suited to capture the adjustability, if any, of decision variables to the true data, (ARC) has a severe built-in drawback:** *it is a "genuine" infinite-dimensional problem, and as such is, in general, severely computationally intractable.* **It is unclear even how to represent candidate decision rules – which are functions of many variables! – in a computer.** Seemingly the *only* techniques allowing to handle ARC are offered by Dynamic Programming, and thus suffer from the "curse of dimensionality."

♠ **Remedy:** to restrict ourselves with *parametric* decision rules, specifically, with *affine* ones:

$$X_j(P_j\zeta) \equiv \xi_j + \eta_j^T P_j \zeta.$$

♡ **Restricted to affine decision rules, the ARC becomes a** *finite-dimensional* **semi-infinite problem**

$$\min_{t, \xi_j, \eta_j} \left\{ t : \begin{array}{l} \sum_j c_j[\zeta][\xi_j + \eta_j^T P_j \zeta] + d[\zeta] \leq t \\ \sum_j [\xi_j + \eta_j^T P_j \zeta] A_j[\zeta] \leq b[\zeta] \end{array} \right\} \forall \zeta \in \mathcal{Z} \Bigg\}$$

called the *Affinely Adjustable RC* **of the uncertain problem $\mathcal{P}$.**

$$\min_{t,\xi_j,\eta_j}\left\{t:\ \left.\begin{array}{l}\sum_j c_j[\zeta][\xi_j+\eta_j^TP_j\zeta]+d[\zeta]\leq t\\ \sum_j[\xi_j+\eta_j^TP_j\zeta]A_j[\zeta]\leq b[\zeta]\end{array}\right\}\forall\zeta\in\mathcal{Z}\right\} \tag{AARC}$$

Definition: We say that $\mathcal{P}$ is *with fixed recourse*, if the coefficients $c_j[\zeta]$, $A_j[\zeta]$ of every *adjustable* (i.e., with $P_j\neq 0$) variable x_j are in fact certain.

Observation: Under fixed recourse, (AARC) is of the same structure as the RC of $\mathcal{P}$, specifically, is a problem with linear objective and *bi-affine* in $\zeta\in\mathcal{Z}$ and in (x,t) constraints.

We have arrived at the following

Theorem 1: *The AARC of an uncertain LO problem with fixed recourse is computationally tractable, provided the perturbation set $\mathcal{Z}$ is so.*

Theorem 2: *The AARC of an uncertain LO problem with non-fixed recourse and with $\cap$-ellipsoidal perturbation set $\mathcal{Z}=\mathcal{Z}_\rho=\{\zeta:\zeta^TQ_j\zeta\leq\rho^2, 1\leq j\leq J\}$ $[Q_j\succeq 0,\sum_j Q_j\succ 0]$ admits a safe tractable approximation tight within the factor $O(1)\sqrt{\ln(J+1)}$. When $J=1$, the approximation is exact.*

Note: the conclusion of Theorem 2 remains valid when the decision rules for "fixed recourse" x_j (those which enter the problem solely with certain coefficients) are allowed to be quadratic in $P_j\zeta$, and the decision rules for the "non-fixed-recourse" variables are allowed to be affine in $P_j\zeta$.

♣ **How it works: Inventory Problem.** A single-product inventory comprised of a warehouse and I factories evolves over time horizon $1, ..., N$. The inventory is affected by uncertain demand $d = [d_1; ...; d_N]$ varying in a given domain D. No backlogged demand is allowed. Let

- x_t be the inventory level at the beginning of period t,
- w_{it} be the amount of product, ordered from and delivered by factory # i in period t.

Given the initial state x_1 of the inventory, bounds on the inventory levels and on the instant and cumulative replenishment orders, we want to minimize the worst, over the demand trajectories from D, overall ordering cost:

$$
\begin{array}{lll}
\min\limits_{C,x,w} & C & \text{[total ordering cost]} \\
\textbf{s.t.} & C \geq \sum_{t=1}^{N} \sum_{i=1}^{I} c_{it} w_{it} & \text{[cost description]} \\
& x_{t+1} = x_t + \sum_{i=1}^{I} w_{it} - d_t & \text{[state equations]} \\
& \underline{X} \leq x_t \leq \overline{X} & \text{[bounds on states]} \\
& 0 \leq w_{it} \leq W_{it} & \text{bounds on orders]} \\
& 0 \leq \sum_{\tau=1}^{N} w_{it} \leq \widehat{W}_i & \text{[bounds on accumulated orders]}
\end{array}
$$

♠ **Applying the AARC approach, we**

- **allow to our actual "wait and see decisions"** w_{it} **to depend** *affinely* **on** $P_t d \equiv [d_1; ...; d_{t-1}]$**:** $w_{it} = p_{it} + \sum_{\tau<t} q_{it}^{\tau} d_{\tau}$
- **allow to the "analysis variables"** x_2**,...,**x_{N+1} **to be arbitrary** *affine* **functions of** d**:** $x_t = \xi_t + \eta_t^T d$**;**
- **treat** C **as the only non-adjustable variable.**

We plug the decision rules into the model and require the constraints to be satisfied for all $d \in D$, thus ending up with the semi-infinite LO problem

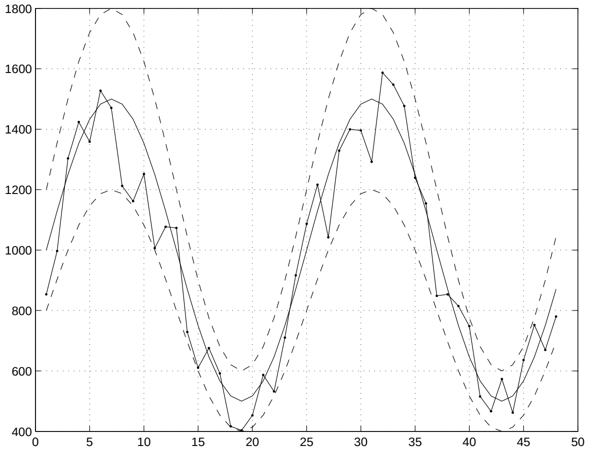
$$\begin{array}{ll}
\min\limits_{\substack{C,\xi_t\eta_t,\\ p_{it},q_{it}^\tau}} & C \\
\textbf{s.t.} & C \geq \sum_{t=1}^N \sum_{i=1}^I c_{it}\left[p_{it} + \sum_{\tau<t} q_{it}^\tau d_\tau\right] \\
& \xi_{t+1} + \eta_{t+1}^T d = \xi_t + \eta_t^T d + \sum_{i=1}^I [p_{it} + \sum_{\tau<t} q_{it}^\tau d_\tau] - d_t \\
& \underline{X} \leq \xi_t + \eta_t^T d \leq \overline{X} \\
& 0 \leq \sum_{i=1}^I [p_{it} + \sum_{\tau<t} q_{it}^\tau d_\tau] \leq W_{it} \\
& 0 \leq \sum_{t=1}^N [p_{it} + \sum_{\tau<t} q_{it}^\tau d_\tau] \leq \widehat{W}_i
\end{array}$$

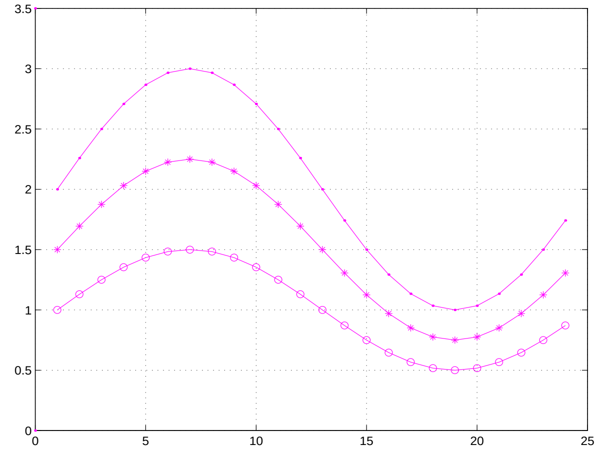
where the constraints should be satisfied for all $d \in D$.

• **Note:** We are in the situation of fixed recourse $\Rightarrow$ the AARC is tractable, provided that D is so.

• **Note:** We could handle easily a much more complicated problem (many products, additional components in the cost function, probabilistic constraints, etc., etc.) *All what matters is that the underlying problem is uncertain LO with uncertainty affecting the right hand sides of constraints only (and thus not affecting coefficients of adjustable variables).*

♣ **Numerical illustration:** $N = 48$, $I = 3$, D **is a box:**





Demand box $0.8d^* \leq d \leq 1.2d^*$ and a sample demand trajectory

Ordering prices c_{it} (periodic with period 24)

♡

We ran several hundreds of simulations and compared the average replenishment cost incurred by optimal *affine* **decision rules with** *utopian cost* **we could pay when** *knowing in advance* **the demand trajectory and** *optimizing accordingly our policy*. **This comparison (biased** *against* **affine decision rules) shows a surprisingly high quality of these rules:**

	AARC-based cost		Utopian cost	
Uncertainty	Mean	Std	Mean	Std
2.5%	33974	190	33878(-0.3%)	194
5%	34063	432	33864(-0.6%)	454
10%	34471	595	34009(-1.6%)	621
20%	35121	1458	33958(-3.4%)	1541

Note: With our setup, the RC is infeasible already at the 5% uncertainty level.

♡ **There is significant evidence that affine decision rules indeed work well in multi-stage Inventory problems.**

♣ **Recent (difficult!) result of Bertsimas, Iancu and Parrilo states that** *For a single-product Inventory with the cost function* $\sum_{t=1}^{N}[c_t w_t + h_t(x_t)]$ *(*$c_t > 0$, $h_t(\cdot)$ *are convex), state equations*

$$x_t = x_{t-1} + w_t - d_t$$

bounds $\underline{W}_t \leq w_t \leq \overline{W}_t$ *on replenishment orders and* a box *D in the role of the set of uncertain demand trajectories, the ARC is equivalent to the AARC, that is, the optimal, in terms of the worst-case management cost, decision rules can be chosen to be affine in the respective parts of the demand trajectory.*

While this theoretical result cannot be extended to a more general settings of the Inventory problem (say, it fails to be true when bounds on accumulated orders are added, and/or the box uncertainty set is replaced with a more general one), AARC, practically speaking, seems to be a good technique for *worst-case oriented* Inventory management.

Note: *When passing from minimizing the worst-case management cost to minimizing the average one, affine decision rules become by far non-optimal already for pretty simple Inventory models.*

♣ **Assume we want to solve a "restricted ARC" on an Uncertain LO problem with fixed recourse, that is, its ARC where the decision rules are restricted to reside in a given class (e.g., to be affine, or quadratic, or polynomial,... in their arguments).**

• When restricted to affine decision rules, the ARC becomes easy. *Is the affinity an actual restriction here?*

♠ **Assume that instead of affine decision rules** $x_j = \xi_j + \eta_j^T P_j\zeta$ **we indent to use rules from a general parametric family:**

$$x_j = \sum\nolimits_\ell \eta_{j\ell} f_{j\ell}(P_j\zeta) \qquad (*)$$

• $f_{j\ell}(\cdot)$**: "basic functions"** • $\eta_{j\ell}$**: free parameters.**

♡Augmenting the perturbation ζ by the entries $\zeta_{j\ell} = f_{j\ell}(P_j\zeta)$, that is, extending ζ to the new perturbation vector

$$\widehat{\zeta}[\zeta] = [\zeta; \{f_{j\ell}(P_j\zeta)\}_{j,\ell}],$$

decision rules $(*)$ become affine in the new perturbation. Thus, for all practical purposes *all parametric decision rules can be thought of as affine ones.*

!!! Bottleneck: The AARC of an Uncertain LO with fixed recourse is easy due to both affinity of the decision rules *and the assumption* (which we always made) *that the perturbation set $\mathcal{Z}$ is tractable*. Passing from ζ to its nonlinear transform $\widehat{\zeta}[\zeta]$, the perturbation set becomes $\widehat{\mathcal{Z}} = \mathrm{Conv}\{\widehat{\zeta}[\zeta] : \zeta \in \mathcal{Z}\}$ and can easily lose tractability.

Good case: Quadratic decision rules, $\mathcal{Z}$ is an ellipsoid. Here $\widehat{\zeta}[\zeta]=\left[\begin{array}{c|c} & \zeta^T\\ \hline \zeta & \zeta\zeta^T\end{array}\right]$. **Assuming** $\mathcal{Z}=\{\zeta:\|\zeta\|_2\leq 1\}$, $\widehat{\mathcal{Z}}:=\mathrm{Conv}\{\widehat{\zeta}[\zeta]:\zeta\in\mathcal{Z}\}$ **is computationally tractable:**

$$\widehat{\mathcal{Z}}=\left\{\left[\begin{array}{c|c} & \zeta^T\\ \hline \zeta & Z\end{array}\right]:\left[\begin{array}{c|c}1 & \zeta^T\\ \hline \zeta & Z\end{array}\right]\succeq 0,\ \mathrm{Tr}(Z)\leq 1\right\}$$

Semi-Good case: Quadratic decision rules, $\cap$-ellipsoidal uncertainty. Here

$$\mathcal{Z}=\{\zeta:\zeta^TQ_j\zeta\leq 1,\,1\leq j\leq J\}$$
$$[J>1,Q_j\succeq 0,\textstyle\sum_j Q_j\succ 0]$$

and $\widehat{\zeta}[\cdot]$ **is as above. Now the set**

$$\widehat{\mathcal{Z}}=\mathrm{Conv}\{\widehat{\zeta}[\zeta]:\zeta\in\mathcal{Z}\}$$

can be intractable, but it admits an outer tractable approximation:

$$\widehat{\mathcal{Z}}\subset\widetilde{\mathcal{Z}}=\left\{\left[\begin{array}{c|c} & \zeta^T\\ \hline \zeta & Z\end{array}\right]:\left[\begin{array}{c|c}1 & \zeta^T\\ \hline \zeta & Z\end{array}\right]\succeq 0,\ \mathrm{Tr}(ZQ_j)\leq 1,1\leq j\leq J\right\}$$

which is tight within factor $\vartheta=O(1)\ln(J)$**:**

$$\vartheta^{-1}\widetilde{\mathcal{Z}}\subset\widehat{\mathcal{Z}}\subset\widetilde{\mathcal{Z}}.$$

Generic Application: Synthesis of Linear Controllers

♣ **Consider time-varying discrete time linear dynamical system**

$$
\begin{array}{rcll}
x_0 &=& z & \text{[initial state]} \\
x_{t+1} &=& A_t x_t + B_t u_t + R_t d_t & \left[\begin{array}{l}\text{state equations} \\ \bullet\ x_t\text{: state}\quad \bullet\ u_t\text{: control} \\ \bullet\ d_t\text{: external disturbance}\end{array}\right] \\
y_t &=& C_t x_t + D_t d_t & \text{[observed output]}
\end{array}
$$

"closed" by *affine output-based control law*

$$u_t = g_t + \sum_{\tau=0}^{t} G_t^\tau y_\tau. \qquad (*)$$

♠ **Given finite time horizon $0 \leq t \leq N$, we want to specify a control law $(*)$ which ensures that** *the state-control trajectory $w = [x_0; ...x_{N+1}; u_0; ...; u_N]$* **satisfies given design specifications**

$$Aw \leq b \qquad (!)$$

robustly **w.r.t. the "perturbation"** $\zeta = [z; d_0; ...; d_N]$ **running through a given set $\mathcal{Z}$.**

Good news: by linearity of the system and the control law, the trajectory is affine in ζ.

$\Rightarrow$ The *Analysis problem: check whether a given control law $(*)$ robustly meets the design specifications* reduces to verifying whether a system of affine constraints on ζ is satisfied by all $\zeta \in \mathcal{Z}$. This is easy, provided $\mathcal{Z}$ is tractable.

$$\begin{array}{rcl} x_0 &=& z \\ x_{t+1} &=& A_t x_t + B_t u_t + R_t d_t \\ y_t &=& C_t x_t + D_t d_t \end{array} \qquad (S)$$

$$u_t = g_t + \sum\nolimits_{\tau=0}^{t} G_t^\tau y_\tau \qquad (*)$$

Bad news: the trajectory is highly nonlinear in the parameters $\gamma = \{g_t, G_t^\tau\}$ of the control law $(*)$

$\Rightarrow$ The *Synthesis problem: find control law* $(*)$, *if it exists, which robustly meets the design specifications* seems to be intractable.

Remedy: pass to affine purified-output-based control laws.

♠ **Consider, along with system** (S) **"closed" by some control law, its** *model*

$$\begin{array}{rcl} \widehat{x}_0 &=& 0 \\ \widehat{x}_{t+1} &=& A_t \widehat{x}_t + B_t u_t \\ \widehat{y}_t &=& C_t \widehat{x}_t \end{array} \qquad (M)$$

which we "feed" by the same controls u_t **as** (S)**. We can run the model in an on-line fashion, and thus at time** t**, before the decision on** u_t **should be made, we have in our disposal** *purified output* $v_t = y_t - \widehat{y}_t$

Observation: *purified outputs are independent on the control law known in advance affine functions of* ζ*.*

Indeed, setting $\Delta_t = x_t - \widehat{x}_t$, we clearly have

$$v_t = C_t \Delta_t + D_t d_t,\ \Delta_0 = z,\ \Delta_{t+1} = A_t \Delta_t + R_t d_t.$$

<table>
<tr><td>System:
$$\begin{array}{rcl} x_0 &=& z\\ x_{t+1} &=& A_t x_t + B_t u_t + R_t d_t \\ y_t &=& C_t x_t + D_t d_t \end{array} \quad (S)$$</td><td>Model:
$$\begin{array}{rcl} \widehat{x}_0 &=& 0\\ \widehat{x}_{t+1} &=& A_t \widehat{x}_t + B_t u_t \\ \widehat{y}_t &=& C_t \widehat{x}_t \end{array} \quad (M)$$</td></tr>
<tr><td colspan="2">Purified outputs: $v_t = y_t - \widehat{y}_t$</td></tr>
<tr><td colspan="2">$$u_t = \begin{cases} g_t + \sum_{\tau=0}^t G_t^\tau y_\tau & \text{[output-based affine law]} \quad (*)\\ h_t + \sum_{\tau=0}^t H_t^\tau v_t & \text{[purified-output-based affine law]} \quad (+) \end{cases}$$</td></tr>
</table>

Facts:

♡ *Purified-Output-Bbased (POB) affine laws are equivalent to the output-based affine laws: every mapping $\zeta \to w$ which can be obtained when "closing" (S) by a law $(*)$, can be obtained by closing (S) by a law $(+)$, and vice versa.*

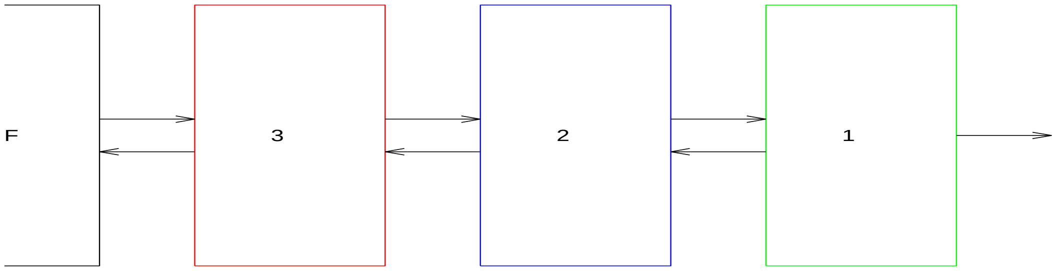
♡ *When (S) is closed by a purified-output-based affine control law $(+)$, the trajectory $w = W[\zeta, \eta]$ becomes bi-affine in ζ and in the parameters $\eta = \{h_t, H_t^\tau\}$ of the control law.*

♡ **As a result,** *Sticking to purified-output-based control laws, the Synthesis problem*

> *Given design specifications $Aw \leq b$ on the state-control trajectory, find a control law, if one exists, which meets these specifications robustly w.r.t. $\zeta = [z; d_0; ...; d_N] \in \mathcal{Z}$*

becomes an efficiently solvable system of semi-infinite affinely perturbed linear constraints on η.

How it Works: Control of 3-Level Serial Inventory



3–LEVEL SERIAL INVENTORY

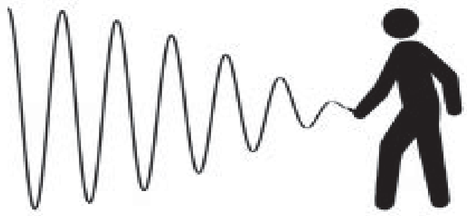
- Level 1 supplies extrnal demand
- Level 2 supplies Level 1
- Level 3 supplies Level 2 and is supplied from Factory
- There is 2-period delay in executing replenishment orders

The Inventory can be modeled as the 9-state LDS

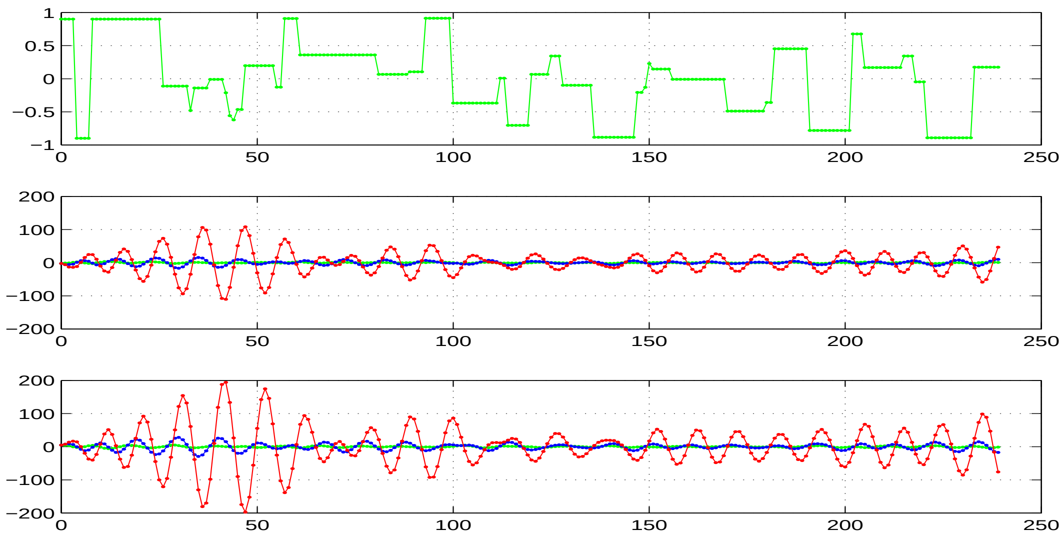
$$
\begin{array}{rcll}
x_1(t+1) & = & x_1(t) + x_2(t) & -d_t \\
x_2(t+1) & = & x_3(t) & \\
x_3(t+1) & = & & u_1(t) \\
x_4(t+1) & = & x_4(t) + x_5(t) & -u_1(t) \\
x_5(t+1) & = & x_6(t) & \\
x_6(t+1) & = & & u_2(t) \\
x_7(t+1) & = & x_7(t) + x_8(t) & -u_2(t) \\
x_8(t+1) & = & x_9(t) & \\
x_9(t+1) & = & & u_3(t) \\
\hline
y(t) & = & x(t) &
\end{array}
$$

- x_1, x_2, x_3 — **inventory levels**
- u_i — **replenishment orders**
- d_t — **demands**

♣ **It is well known that serial inventories with delays suffer from *bullwhip effect*: variations in external demand result in *much larger* variations in the inventory levels, especially in the one closest to the factory, thus badly affecting the production:**



♠ **This is what happens with "naive" feedback:**

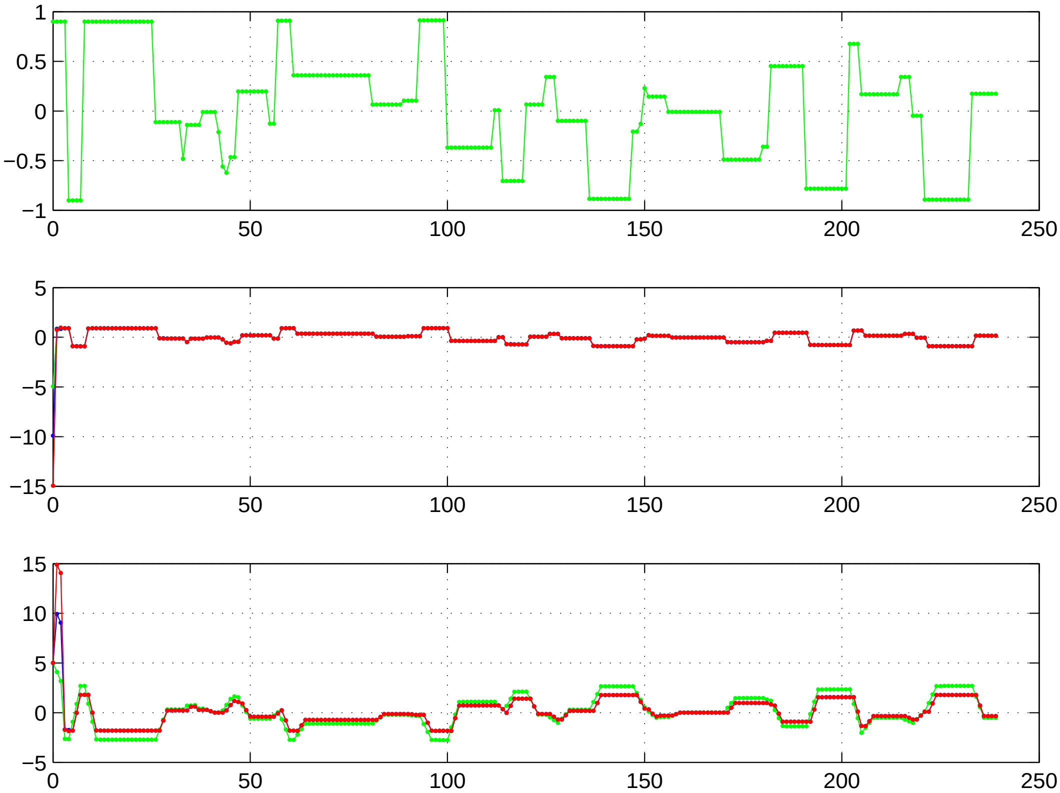


Bullwhip effect

Top: time-dependent demand varying in $[-1, 1]$
Middle: replenishment orders $u_1(t)$, $u_2(t)$, $u_3(t)$
Bottom: inventory levels (green: #1, blue: #2, red: #3)

♠ **Note:** variations of the demand in the range $[-1, 1]$ result in huge (hundreds!) oscillations in the level #3 and in the replenishment orders.

♡ **To reduce the bullwhip effect, we can look for the best — with the largest decay rate as certified by Lyapunov Stability Certificate — linear feedback. With this control, the picture looks much better:**

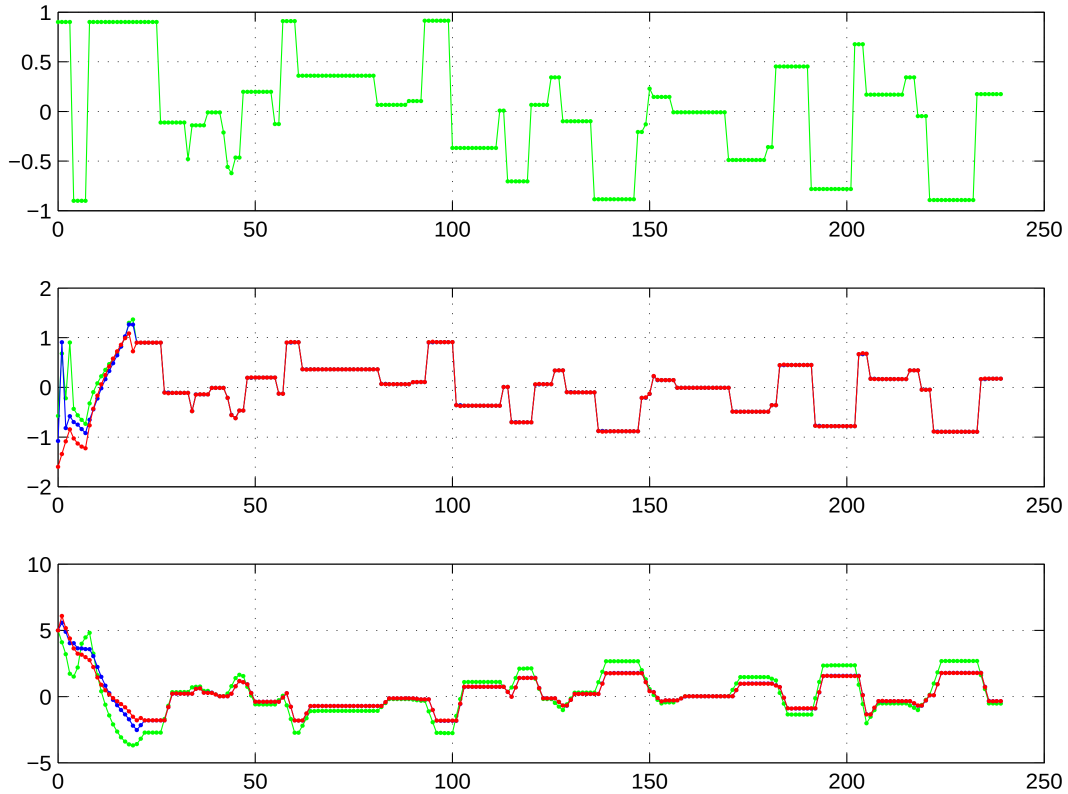


Good linear feedback

Top: time-dependent demand varying in $[-1, 1]$
Middle: replenishment orders $u_1(t)$, $u_2(t)$, $u_3(t)$
Bottom: inventory levels (green: #1, blue: #2, red: #3)

But: At the very beginning, we still have unpleasant jumps in the inventory levels and replenishment orders.

♡ **To improve the behaviour of the process in the beginning, we can use purified-output-based affine control aimed at minimizing the initial jumps and converging to the above feedback control. This is what we get:**



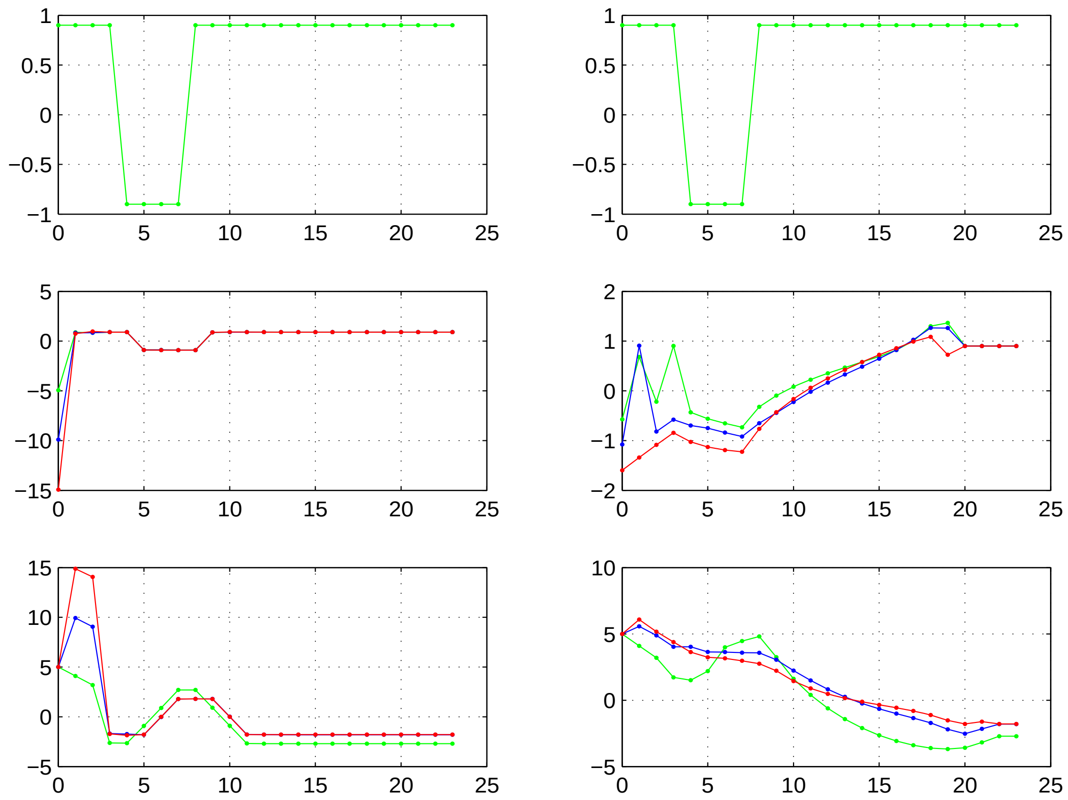
Combined p.o.b./feedback control

Top: time-dependent demand varying in $[-1, 1]$

Middle: replenishment orders $u_1(t)$, $u_2(t)$, $u_3(t)$

Bottom: inventory levels (green: #1, blue: #2, red: #3)

♡ **This is what we gain in the beginning, while loosing nothing in the long run:**



Pure feedback control (left)

vs.

combined p.o.b/feedback control (right)

Top: time-dependent demand varying in $[-1, 1]$

Middle: replenishment orders $u_1(t)$, $u_2(t)$, $u_3(t)$

Bottom: inventory levels (green: #1, blue: #2, red: #3)

Handling Infinite-Horizon Design Specifications

♣ **Let the open loop system be** *time-invariant***:**

$$\begin{aligned} x_0 &= z \\ x_{t+1} &= Ax_t + Bu_t + Rd_t\ ,\ t = 0, 1, ... \\ y_t &= Cx_t + Dd_t \end{aligned}$$

and *stable***: the spectral radius of** A **is** < 1**.**

♠ **Let us use** *nearly time-invariant affine POB control:*

$$u_t = h_t + \sum_{s=0}^{k-1} H_s^t v_{t-s}$$
$$\textbf{for } t \geq N_* : h_t = 0 \ \&\ H_\tau^t = H_\tau$$

N_**: stabilization time***.**

$$\begin{array}{rcl} x_{t+1} & = & Ax_t + Bu_t + Rd_t \\ y_t & = & Cx_t + Dd_t \end{array}, \; t = 0, 1, \ldots$$

$$u_t = h_t + \sum_{s=0}^{k-1} H_s^t v_{t-s}$$

$$\text{for } t \geq N_* : h_t = 0 \;\&\; H_\tau^t = H_\tau$$

♠ **Setting**

$$\delta_t = x_t - \widehat{x}_t, \;\; H = [H_0, \ldots, H_{k-1}],$$

for $t \geq N_*$, evolution of the closed loop system will be given by the *time-invariant LDS*

$$\underbrace{\begin{bmatrix} x_{t+1} \\ \delta_{t+1} \\ \delta_t \\ \vdots \\ \delta_{t-k+2} \end{bmatrix}}_{\omega_{t+1}} = \underbrace{\left[\begin{array}{c|cccc} A & BH_0C & BH_1C & \ldots & BH_{k-1}C \\ \hline & A & & & \\ & & A & & \\ & & & \ddots & \\ & & & & A \end{array}\right]}_{A_+[H]} \omega_t$$

$$+ \underbrace{\left[\begin{array}{c|cccc} R & BH_0D & BH_1D & \ldots & BH_{k-1}D \\ \hline & R & & & \\ & & R & & \\ & & & \ddots & \\ & & & & R \end{array}\right]}_{R_+[H]} \underbrace{\begin{bmatrix} d_t \\ d_t \\ d_{t-1} \\ \vdots \\ d_{t-k+1} \end{bmatrix}}_{d^t}$$

$$u_t = \sum_{\nu=0}^{k-1} H_\nu [C\delta_{t-\nu} + Dd_{t-\nu}].$$

♠ **Facts:**

- *The matrix $A_+[H]$ is affine in H and is stable for all H*
- *The resolvent*

$$\mathcal{R}_H(s) := (sI - A_+[H])^{-1} \qquad [s \in \mathbf{C}]$$

has all its singularities in spectrum of A and is affine in H.

$$\boxed{\begin{array}{rcl} x_{t+1} &=& Ax_t + Bu_t + Rd_t \\ y_t &=& Cx_t + Dd_t \end{array}, \; t = 0,1,\ldots}$$

$$\boxed{\begin{array}{c} u_t = h_t + \sum_{s=0}^{k-1} H_s^t v_{t-s} \\ \text{for } t \geq N_* : h_t = 0 \ \&\ H_\tau^t = H_\tau \end{array}}$$

♣ **Starting with time N_*, the closed loop system can be "embedded" into the time-invariant LDS**

$$\boxed{\begin{array}{rcl} \omega_{t+1} &=& A_+[H]\omega_t + R_+[H]d^t \\ u_t &=& \sum_{\nu=0}^{k-1} H_\nu [C\delta_{t-\nu} + Dd_{t-\nu}] \end{array}}$$

♠ **Let $s \in \mathbf{C}$ differ from 0 and from all eigenvalues of A, and let the external disturbance be**

$$d_t = s^t f, \; t = 0,1,\ldots$$

♠ **Facts:**

• *As $t \to \infty$, the state-control trajectory $[x_t; u_t]$ approaches the trajectory*

$$s^t \mathcal{H}_{x,u}(s) f \qquad .$$

The transfer matrix $\mathcal{H}_{x,u}(s)$ is easy to compute: setting $\mathcal{R}_A(s) = (sI - A)^{-1}$, one has

$$\mathcal{H}_{xu}(s) = \left[\begin{array}{c} \overbrace{\mathcal{R}_A(s)\left[R + \sum_{\nu=0}^{k-1} s^{-\nu} B H_\nu \left[D + C\mathcal{R}_A(s) R\right]\right]}^{\mathcal{H}_x(s)} \\ \hline \underbrace{\left[\sum_{\nu=0}^{k-1} s^{-\nu} H_\nu\right]\left[D + C\mathcal{R}_A(s)R\right]}_{\mathcal{H}_u(s)} \end{array}\right]$$

The transfer matrix has all its singularities in the spectrum of A and is affine in the steady-state parameters $H = [H_0, \ldots, H_{k-1}]$ of the POB control law we use.

$$\boxed{\begin{array}{c}\left.\begin{array}{rcl} x_{t+1} &=& Ax_t + Bu_t + Rd_t \\ y_t &=& Cx_t + Dd_t \end{array}\right., t = 0, 1, \dots \\ \hline u_t = h_t + \sum_{s=0}^{k-1} H_s^t v_{t-s} \\ \text{for } t \geq N_* : h_t = 0 \ \&\ H_\tau^t = H_\tau \end{array}}$$

♠ **The transfer matrix of the closed loop system has all its singularities in the spectrum of (stable)** A **and** *is affine in the parameters* $H = [H_0, ..., H_{k-1}]$ *of the POB affine control law we use.*

$\Rightarrow$ *Let the design specifications on* asymptotic behaviour *of the closed loop system be given by* convex constraints *on the transfer matrix. Then design of the POB affine control law meeting these specifications reduces to Convex Programming.*

• *After the above design specifications are met, we can further adjust the "transitional behaviour" of the closed loop system by modifying the resulting POB on finite horizon* $0 \leq t \leq N_*$*, which again reduces to Convex Programming.*

Example: Discrete time H_∞ control.

♣ **Discrete time H_∞ design specifications impose constraints on the transfer matrix along the unit circumference $s = \exp\{\imath\omega\}$, $0 \leq \omega \leq 2\pi$, that is, on the steady state response of the closed loop system to harmonic oscillations in the role of disturbances.**

Let the only nonzero entry in the disturbances be the $\jmath$-th one, and let it be a harmonic oscillation of unit amplitude and frequency ω. Induced steady-state behavior of $\imath$-th state is the harmonic oscillation of the same frequency with the amplitude $|(\mathcal{H}_x(\exp\{\imath\omega\}))_{ij}|$ and the phase shifted by $\arg(\mathcal{H}_x(\exp\{\imath\omega\})_{ij})$.
$\Rightarrow$ *The state-to-input responses* $(\mathcal{H}_x(\exp\{\imath\omega\}))_{ij}$ *explain the steady-state behavior of states when input is comprised of harmonic oscillations, and similarly for control-to-input responses* $(\mathcal{H}_u(\exp\{\imath\omega\}))_{ij}$.

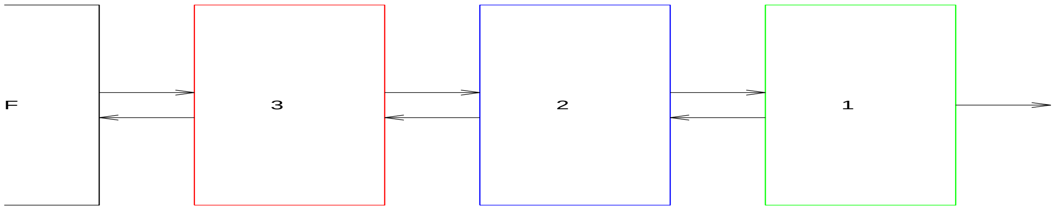
♠ **A wide spectrum of H_∞ design specifications are given by the systems of constraints of the form**

$$\forall (s = \exp\{\imath\omega\} : \omega \in \Delta_i \subset [0, 2\pi]) :$$
$$\|G_i(s) - L_i(s)\mathcal{H}_{xy}(s)R_i(s)\| \leq \tau_i,\ 1 \leq i \leq m$$

- $G_i(s), L_i(s), R_i(s)$: rational in s matrix-valued functions with no singularities on $|s| = 1$
- Δ_i: given segments • $\|\cdot\|$: standard matrix norm.

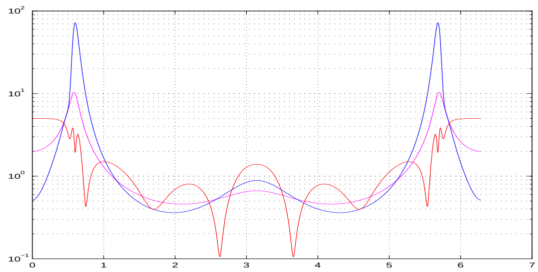
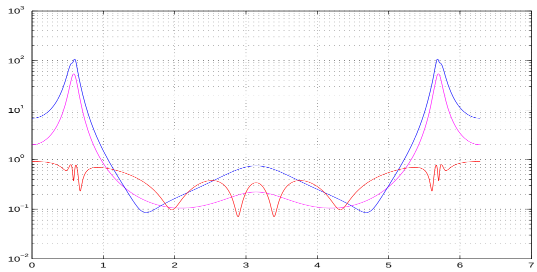
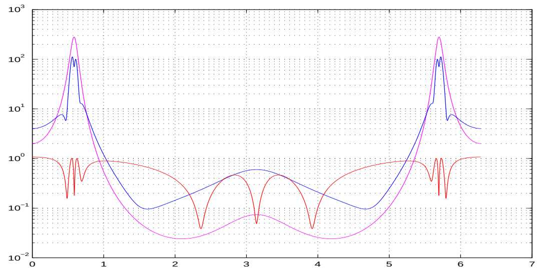
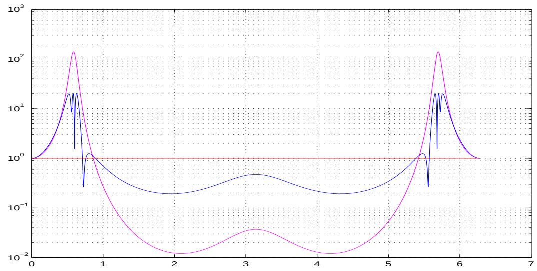
♡ *These specifications reduce to explicit convex constraints on the parameters H of the affine POB control law we intend to use and thus are tractable*

Control of 3-Level Serial Inventory (continued)



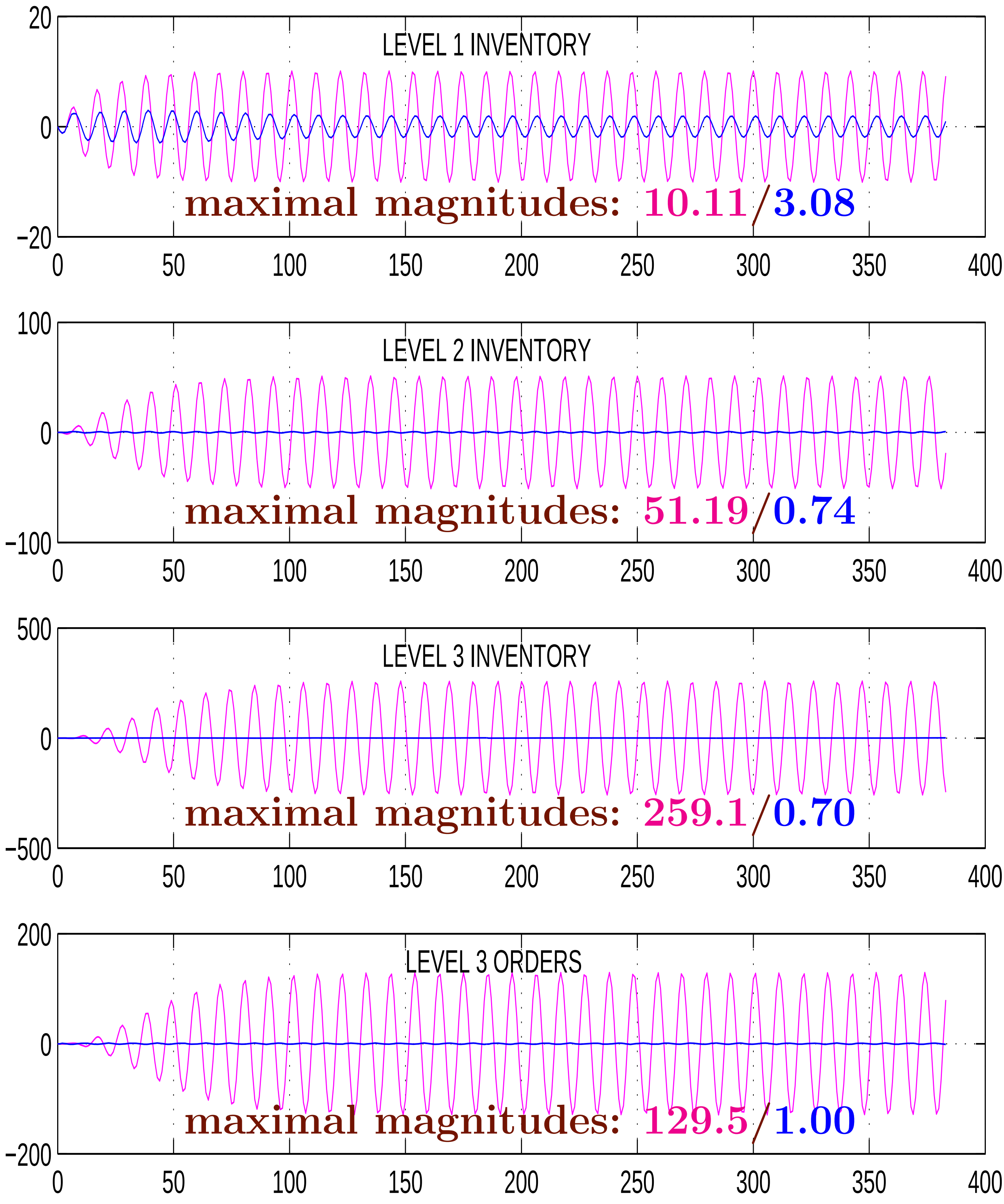
3−LEVEL SERIAL INVENTORY

♣ **Minimizing the maximal, over frequencies, magnitude of the frequency response of level 3 orders to the demand, we get results as follows:**

Inventory, level 1	Inventory, level 2
	
max: 10.39/71.95/5.00	max: 54.00/107.13/0.92
Inventory, level 3	**Orders, level 3**
	
max: 280.53/111.77/1.07	max: 140.29/20.38/1.00

Magnitudes of frequency responses

- Magenta: naive linear feedback
- Blue: optimal POB affine control with $k = 1$
- Red: optimal POB affine control with $k = 6$



Inventories and Orders vs. Time

Disturbance: $\cos(0.6t)$ **starting with** $t = 0$

Magenta: naive feedback

Blue: optimal affine POB control with $k = 6$

We have dramatically reduced and inverted the bullwhip effect!