Optimization under uncertainty: robust optimization

Imre Pólik, PhD

McMaster University
School of Computational Engineering and Science

March 6, 2008

Outline

1. Motivation
   - The role of uncertainty
2. Theoretical background
   - Robust linear optimization
   - Robust least squares optimization
3. Applications
   - Truss topology design
   - Antenna array design

Why uncertainty?

Design on computer,
measure with a ruler,
cut with an axe!

- Uncertain data
  - measurement errors
  - noise
  - unpredictability
  - vibration
- Implementation errors
  - technological limits
  - imperfect materials

Robust linear optimization

Robust linear programming

the parameters in optimization problems are often uncertain, *e.g.*, in an LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m, \end{array}$$

there can be uncertainty in c, a_i, b_i

two common approaches to handling uncertainty (in a_i, for simplicity)

- deterministic model: constraints must hold for all $a_i \in \mathcal{E}_i$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i \text{ for all } a_i \in \mathcal{E}_i, \quad i = 1, \ldots, m, \end{array}$$

- stochastic model: a_i is random variable; constraints must hold with probability η

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{prob}(a_i^T x \leq b_i) \geq \eta, \quad i = 1, \ldots, m \end{array}$$

deterministic approach via SOCP

- choose an ellipsoid as $\mathcal{E}_i$:

$$\mathcal{E}_i = \{\bar{a}_i + P_i u \mid \|u\|_2 \leq 1\} \qquad (\bar{a}_i \in \mathbf{R}^n, \quad P_i \in \mathbf{R}^{n \times n})$$

  center is $\bar{a}_i$, semi-axes determined by singular values/vectors of P_i

- robust LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i \quad \forall a_i \in \mathcal{E}_i, \quad i = 1, \ldots, m \end{array}$$

  is equivalent to the SOCP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i^T x + \|P_i^T x\|_2 \leq b_i, \quad i = 1, \ldots, m \end{array}$$

  (follows from $\sup_{\|u\|_2 \leq 1} (\bar{a}_i + P_i u)^T x = \bar{a}_i^T x + \|P_i^T x\|_2$)

stochastic approach via SOCP

- assume a_i is Gaussian with mean $\bar{a}_i$, covariance Σ_i ($a_i \sim \mathcal{N}(\bar{a}_i, \Sigma_i)$)
- $a_i^T x$ is Gaussian r.v. with mean $\bar{a}_i^T x$, variance $x^T \Sigma_i x$; hence

$$\mathbf{prob}(a_i^T x \leq b_i) = \Phi\left(\frac{b_i - \bar{a}_i^T x}{\|\Sigma_i^{1/2} x\|_2}\right)$$

where $\Phi(x) = (1/\sqrt{2\pi}) \int_{-\infty}^{x} e^{-t^2/2}\, dt$ is CDF of $\mathcal{N}(0,1)$

- robust LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{prob}(a_i^T x \leq b_i) \geq \eta, \quad i = 1, \ldots, m, \end{array}$$

with $\eta \geq 1/2$, is equivalent to the SOCP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i^T x + \Phi^{-1}(\eta) \|\Sigma_i^{1/2} x\|_2 \leq b_i, \quad i = 1, \ldots, m \end{array}$$

Robust least squares optimization

Robust approximation

minimize $\|Ax - b\|$ with uncertain A

two approaches:

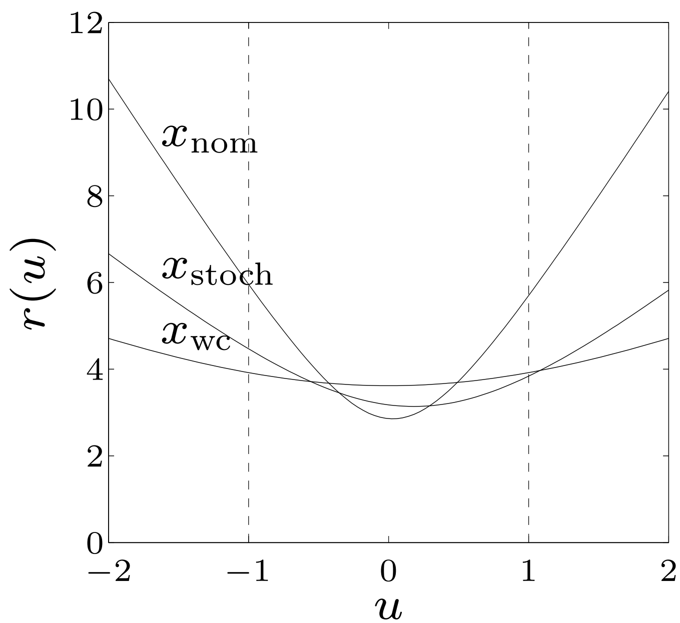
- **stochastic**: assume A is random, minimize $\mathbf{E}\,\|Ax - b\|$
- **worst-case:** set $\mathcal{A}$ of possible values of A, minimize $\sup_{A \in \mathcal{A}} \|Ax - b\|$

tractable only in special cases (certain norms $\|\cdot\|$, distributions, sets $\mathcal{A}$)

example: $A(u) = A_0 + uA_1$

- x_{nom} minimizes $\|A_0 x - b\|_2^2$
- x_{stoch} minimizes $\mathbf{E}\,\|A(u)x - b\|_2^2$ with u uniform on $[-1, 1]$
- x_{wc} minimizes $\sup_{-1 \le u \le 1} \|A(u)x - b\|_2^2$

figure shows $r(u) = \|A(u)x - b\|_2$



stochastic robust LS with $A = \bar{A} + U$, U random, $\mathbf{E}\, U = 0$, $\mathbf{E}\, U^T U = P$

$$\text{minimize} \quad \mathbf{E}\, \|(\bar{A} + U)x - b\|_2^2$$

- explicit expression for objective:

$$\begin{aligned} \mathbf{E}\, \|Ax - b\|_2^2 &= \mathbf{E}\, \|\bar{A}x - b + Ux\|_2^2 \\ &= \|\bar{A}x - b\|_2^2 + \mathbf{E}\, x^T U^T U x \\ &= \|\bar{A}x - b\|_2^2 + x^T P x \end{aligned}$$

- hence, robust LS problem is equivalent to LS problem

$$\text{minimize} \quad \|\bar{A}x - b\|_2^2 + \|P^{1/2}x\|_2^2$$

- for $P = \delta I$, get Tikhonov regularized problem

$$\text{minimize} \quad \|\bar{A}x - b\|_2^2 + \delta\|x\|_2^2$$

worst-case robust LS with $\mathcal{A} = \{\bar{A} + u_1 A_1 + \cdots + u_p A_p \mid \|u\|_2 \leq 1\}$

$$\text{minimize} \quad \sup_{A \in \mathcal{A}} \|Ax - b\|_2^2 = \sup_{\|u\|_2 \leq 1} \|P(x)u + q(x)\|_2^2$$

where $P(x) = \left[\begin{array}{cccc} A_1 x & A_2 x & \cdots & A_p x \end{array}\right]$, $q(x) = \bar{A}x - b$

- from page 5–14, strong duality holds between the following problems

$$\begin{array}{ll} \text{maximize} & \|Pu + q\|_2^2 \\ \text{subject to} & \|u\|_2^2 \leq 1 \end{array} \qquad \begin{array}{ll} \text{minimize} & t + \lambda \\ \text{subject to} & \left[\begin{array}{ccc} I & P & q \\ P^T & \lambda I & 0 \\ q^T & 0 & t \end{array}\right] \succeq 0 \end{array}$$

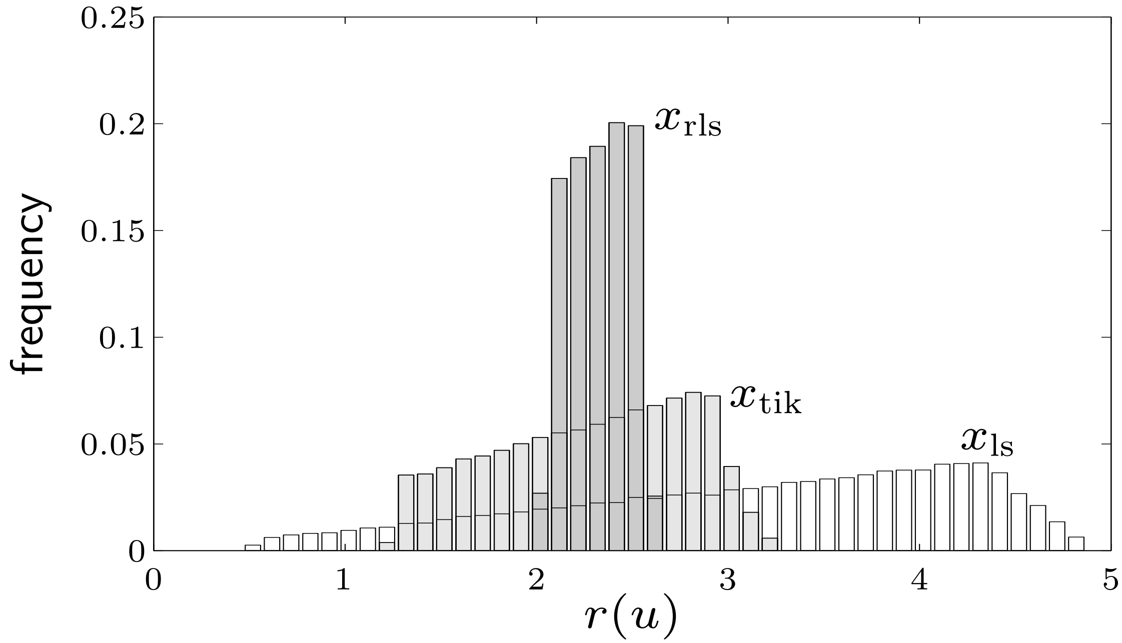
- hence, robust LS problem is equivalent to SDP

$$\begin{array}{ll} \text{minimize} & t + \lambda \\ \text{subject to} & \left[\begin{array}{ccc} I & P(x) & q(x) \\ P(x)^T & \lambda I & 0 \\ q(x)^T & 0 & t \end{array}\right] \succeq 0 \end{array}$$

example: histogram of residuals

$$r(u) = \|(A_0 + u_1 A_1 + u_2 A_2)x - b\|_2$$

with u uniformly distributed on unit disk, for three values of x



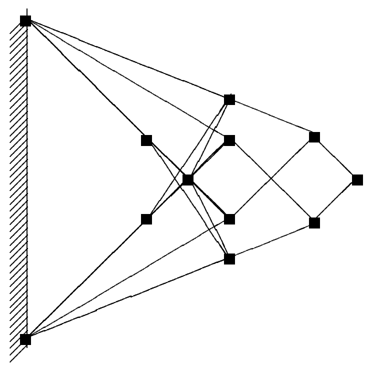
- $x_{\rm ls}$ minimizes $\|A_0 x - b\|_2$
- $x_{\rm tik}$ minimizes $\|A_0 x - b\|_2^2 + \|x\|_2^2$ (Tikhonov solution)
- $x_{\rm wc}$ minimizes $\sup_{\|u\|_2 \leq 1} \|A_0 x - b\|_2^2 + \|x\|_2^2$

Truss topology design

Applications, V: Structural Design

• **Structural design is an engineering area dealing with mechanical constructions like trusses and plates.**

• **Truss is a construction comprised of thin elastic bars linked with each other at nodes from a given 2D/3D nodal set:**



A truss

• **Numerous problems of the type**

"Given

• type of material to be used,

• resource (an upper bound on the amount of material),

and

• set of loading scenarios – external loads of interest,

find a truss/shape capable to withstand best of all the loads in question"

can be casted as semidefinite programs.

A. Building the model

• Mechanical constructions to be considered ("linearly elastic constructions") are characterized as follows.

I. A construction C can be characterized by

I.1) Linear space $\mathbf{V} = \mathbf{R}^m$ of *virtual displacements* **of the construction;**

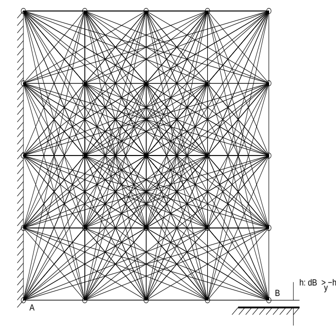
I.2) *Potential energy* **– a nonnegative quadratic form**

$$E_C(v) = \frac{1}{2} v^T A_C v \qquad [v \in \mathbf{V}]$$

indicating what is the energy stored in the construction as a result of displacement v;

I.3) A closed convex subset $\mathcal{V} \subset \mathbf{R}^m$ of *kinematically admissible displacements of the construction.*

Example: Truss.



• The set of virtual displacements of a truss is the direct product, over the nodes, of the spaces of "physical" virtual displacements of the nodes.

For the depicted planar truss, all nodes but the most left ones can move along the plane, and the most left nodes are fixed. In this case,

$$\mathbf{V} = \underbrace{\mathbf{R}^2 \times \ldots \times \mathbf{R}^2}_{25-5=20 \text{ times}} = \mathbf{R}^{40}$$

• The potential energy is

$$E_C(v) = \frac{1}{2} v^T A(t) v \qquad \left[A(t) = \sum_{i=1}^{n} t_i b_i b_i^T \right]$$

(n: # of tentative bars; t_i: volume of bar i; $b_i \in \mathbf{V}$ are given by geometry of the nodal set).

• The set $\mathcal{V}$ of kinematically admissible displacements is cut off $\mathbf{V}$ by rigid external obstacles.

For the depicted truss, there is one obstacle which restricts the y-displacement of node B to be $\geq -h$.

<u>**A construction C:**</u>

- $\mathbf{V} = \mathbf{R}^m$ – **space of virtual displacements;**
- $A_C \succeq 0$ – **matrix of potential energy:** $E_C(v) = \frac{1}{2}v^T A_C v$
- $\mathcal{V} \subset \mathbf{V}$ – **set of kinematically admissible displacements**

II. An <u>external load</u> applied to C can be represented by a vector $f \in \mathbf{R}^m$; the static equilibrium of C loaded by f is given by the

> <u>**Variational Principle:**</u> *A construction C is capable to carry an external load f if and only if the quadratic form*
>
> $$E_C^f(v) = \frac{1}{2}v^T A_C v - f^T v$$
>
> *attains its minimum in $v \in \mathcal{V}$, and displacement yielding (static) equilibrium is a minimizer of $E_C^f(\cdot)$ on $\mathcal{V}$.*
>
> <u>**Compliance:**</u> **The minus minimum value of $E_C^f(v)$ over $v \in \mathcal{V}$:**
>
> $$\mathrm{Compl}_f = \max_{v \in \mathcal{V}} \left[f^T v - \frac{1}{2} v^T A_c v \right].$$
>
> **The less is compliance, the better C withstands the load.**

<u>**A construction C:**</u>

- $\mathbf{V} = \mathbf{R}^m$ – **space of virtual displacements;**
- $A_C \succeq 0$ – **matrix of potential energy:** $E_C(v) = \frac{1}{2}v^T A_C v$
- $\mathcal{V} \subset \mathbf{V}$ – **set of kinematically admissible displacements**

<u>**Compliance of C w.r.t. a load $f \in \mathbf{V}$:**</u>

$$\mathrm{Compl}_f = \max_{v \in \mathcal{V}} \left[f^T v - \frac{1}{2} v^T A_c v \right].$$

III. The stiffness matrix A_C depends on mechanical characteristics $t_1, ..., t_n$ of "elements" $E_1, ..., E_n$ comprising C, and these characteristics t_i are *positive semidefinite symmetric $d \times d$ matrices*. **Specifically,**

$$A_C = \sum_{i=1}^{n} \sum_{s=1}^{S} b_{is} t_i b_{is}^T,$$

where b_{is} are given $m \times d$ matrices.

A natural measure of material resource used by element E_i is the trace $\mathrm{Tr}(t_i)$.

I. A construction C:

- $\mathbf{V} = \mathbf{R}^m$ – **space of virtual displacements;**
- $A_C \succeq 0$ – **matrix of potential energy:** $E_C(v) = \frac{1}{2}v^T A_C v$
- $\mathcal{V} \subset \mathbf{V}$ – **set of kinematically admissible displacements**

II. Compliance of C w.r.t. a load $f \in \mathbf{V}$:

$$\mathrm{Compl}_f = \max_{v \in \mathcal{V}} \left[f^T v - \frac{1}{2} v^T A_c v \right].$$

III. Design variables: $t_i \in \mathbf{S}^d_+$,

$$A_C = \sum_{i=1}^n \sum_{s=1}^S b_{is} t_i b_{is}^T;$$

total material resource: $\Sigma_i \operatorname{Tr}(t_i)$.

• **These assumptions are valid for truss ($t_i \geq 0$ are bar volumes, i.e., $d = 1$).**

• **These assumptions are valid for (finite element approximation of) shape.**

• **Thus, a general Structural Design problem can be modeled as follows. The mechanical data ("ground structure") are**

- **The space $\mathbf{V} = \mathbf{R}^m$ of virtual displacements;**
- **A subset $\mathcal{V} \subset \mathbf{V}$ of kinematically admissible displacements; below, $\mathcal{V} = \{v \mid Rv \leq r\}$;**
- **A collection $\{b_{is} \mid i = 1, ..., n; s = 1, ..., S\}$ of $d \times m$ matrices defining the "energy matrix"**

$$A(t) = \sum_{i=1}^n \sum_{s=1}^S b_{is}^T t_i b_{is},$$

$t = \{t_i \in \mathbf{S}^d_+\}_{i=1}^n$ **being the design variables.**

• **An external load is represented by a vector $f \in \mathbf{V}$, the compliance of a construction w.r.t. f is**

$$\mathrm{Compl}_f(t) = \sup_{v \in \mathcal{V}} \left[f^T v - \frac{1}{2} v^T A(t) v \right],$$

while the material resource is $\Sigma_{i=1}^n \operatorname{Tr}(t_i)$.

• **A typical Optimal Structural Design problem is:**

Given (a) a ground structure, (b) an upper bound w on total material resource along with bounds on the resources of elements, and (c) a set $\mathcal{F}$ of loading scenarios, find a construction with the smallest possible worst case, w.r.t. $f \in \mathcal{F}$, compliance:

$$\boxed{\begin{array}{c} \mathrm{Compl}_{\mathcal{F}}(t) \equiv \sup_{f \in \mathcal{F}} \mathrm{Compl}_f(t) \to \min \\ t_i \succeq 0, \\ \underline{\rho}_i \leq \operatorname{Tr}(t_i) \leq \overline{\rho}_i, \ i = 1, ..., n, \\ \sum_{i=1}^n \operatorname{Tr}(t_i) \leq w \end{array}}$$

• **A typical Structural Design problem is**

$$\boxed{\begin{array}{c}\mathrm{Compl}_{\mathcal{F}}(t) \equiv \sup\limits_{f\in\mathcal{F}} \mathrm{Compl}_f(t) \to \min \\ t_i \succeq 0, \\ \underline{\rho}_i \le \mathrm{Tr}(t_i) \le \overline{\rho}_i,\ i=1,...,n, \\ \sum\limits_{i=1}^{n} \mathrm{Tr}(t_i) \le w\end{array}}$$

where

- $\mathrm{Compl}_f(t) = \sup_{v\in\mathcal{V}} \left[f^T v - \frac{1}{2} v^T A(t) v\right]$
- $A(t) = \Sigma_{i=1}^{n} \Sigma_{s=1}^{S} b_{is}^T t_i b_{is}$
- $\mathcal{V} = \{v \in \mathbf{R}^m \mid Rv \le r\}$
- $t_i \in \mathbf{S}^d$**,** $i = 1, ..., n$**, are design variables**
- $\mathcal{F} \subset \mathbf{R}^m$ **is the set of loading scenarios.**

• **From now on, we assume that**

A) $\mathcal{V}$ **satisfies the Slater condition:**

$$\exists \bar{v} : R\bar{v} < r$$

B) The ground structure does not admit rigid body motions:

$$\sum_{i,s} b_{is} b_{is}^T \succ 0$$

C) $0 \le \underline{\rho}_i < \overline{\rho}_i$ **and**

$$\sum_i \underline{\rho}_i < w$$

Multi-load and Robust Structural Design

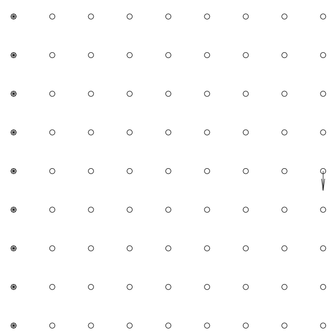
• **In the usual settings of the SD problem,** $\mathcal{F} = \{f_1, ..., f_k\}$ **is a small finite set of "loads of interest" ("multi-load design").**

An alternative is to use a "massive" ellipsoidal set of loads:

$$\mathcal{F} = \{f = Qu \mid u^T u \le 1\} \qquad \left[Q \in \mathbf{M}^{m,k}\right]$$

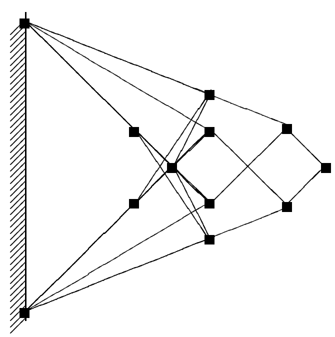
• **The multi-load approach achieves the best possible results, as far as the scenario loads are concerned, but can yield unstable construction which can be crushed by "small occasional load".**

Example: **Assume we are designing a planar truss – a cantilever; the 9×9 nodal structure and the only load of interest f^* are as shown on the picture:**

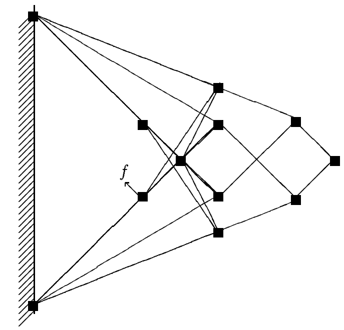


9×9 **ground structure and the load of interest**

The optimal single-load design yields a nice truss as follows:



Optimal cantilever (single-load design)
the compliance is 1.000



Optimal cantilever (single-load design)
the compliance is 1.000

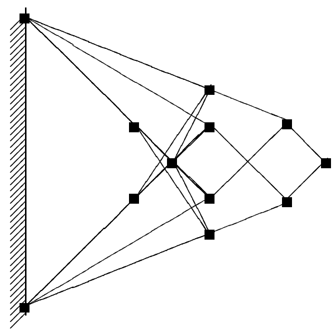
- **It turns out that when the "load of interest" f^* is replaced with the small ($\| f \| = 0.005 \| f^* \|$) "badly placed" occasional load f, the compliance jumps from 1.000 to 8.4 (!)**

- **In order to improve the stability of the design, let us replace the single load of interest f^* by the ellipsoid containing this load and all loads f of magnitude $\| f \|$ not exceeding 10% of the magnitude of f^*:**

$$\mathcal{F} = \{f = u_1 f^* + u_2 f_2 + f_3 u_3 + ... + u_{20} f_{20} \mid u^T u \leq 1\},$$

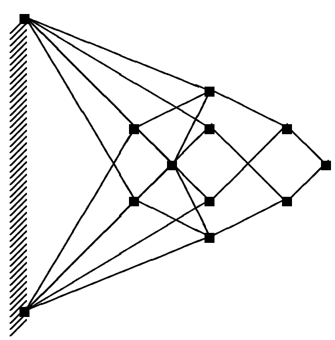
where $f_2, ..., f_{20}$ are of the norm $0.1 \| f^* \|$ and $f^*, f_2, ..., f_{20}$ is an orthogonal basis in $\mathbf{V} = \mathbf{R}^{20}$.

• **Passing from the single-load to the robust design, we modify the result as follows:**



Optimal cantilever (single-load design)

$\Downarrow$



"Robust" cantilever

Compliances	**Design**	
	Single-load	**Robust**
Compliance w.r.t. f^*	**1.000**	**1.0024**
$\max$ **compliance w.r.t. loads** f: $\| f \| \leq 0.1 \| f^* \|$	**32000**	**1.03**

Antenna array design

Synthesis of arrays of antennae

• The diagram of an antenna. Consider a (monochromatic) antenna placed at the origin. The components of the electro-magnetic field generated by the antenna at a point $r\delta$ (δ is a unit direction) are

$$\begin{aligned} E &= a(\delta)r^{-1}\cos(\phi(\delta)+t\omega-2\pi r/\lambda)+o(r^{-1}) \\ H &= a(\delta)r^{-1}\sin(\phi(\delta)+t\omega-2\pi r/\lambda)+o(r^{-1}) \end{aligned}$$

• t: time • ω: frequency λ: wavelength

$$\Updownarrow$$

$$E+iH = Z(\delta)r^{-1}\exp\{i[\omega t-2\pi r/\lambda]\}+o(r^{-1}),$$
$$Z(\delta)=a(\delta)\exp\{i\phi(\delta)\}.$$

The function $Z(\delta)$ is called the diagram of the antenna. The quantity $a^2(\delta)=|Z(\delta)|^2$ is proportional to the directional density of the energy sent by the antenna.

• The diagram of an array of antennae – a complex antenna comprised of N elements – is just the sum of diagrams of the elements.

• Synthesis of Array of Antennae: Given a target diagram $Z_*(\cdot)$ along with N "building blocks" – antenna elements with diagrams $Z_1(\cdot),...,Z_N(\cdot)$ – find "weights" $z_j \in \mathbf{C}$ such that the function

$$\sum_{j=1}^{N} z_j Z_j(\cdot)$$

is as close as possible to the target diagram $Z_*(\cdot)$.
Remark. Physically, multiplication of a diagram $Z_j(\cdot)$ by a complex weight z_j means that the corresponding standard "bulding block" is preceded by appropriate amplification and delay devices.

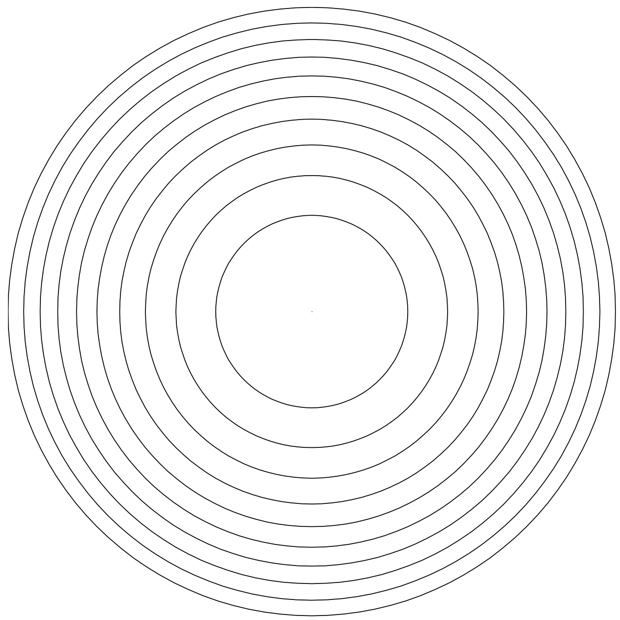
• Choosing a fine grid Δ of directions δ, we may pose the Antenna Synthesis problem as a discrete Tschebyshev problem with complex-valued data and design variables:

$$\max_{\delta\in\Delta}\left|Z_*(\delta)-\sum_{j=1}^{N} z_j Z_j(\delta)\right| \to \min$$

with subsequent approximation of the Tschebyshev problem by an LP program.

Antenna synthesis: Example

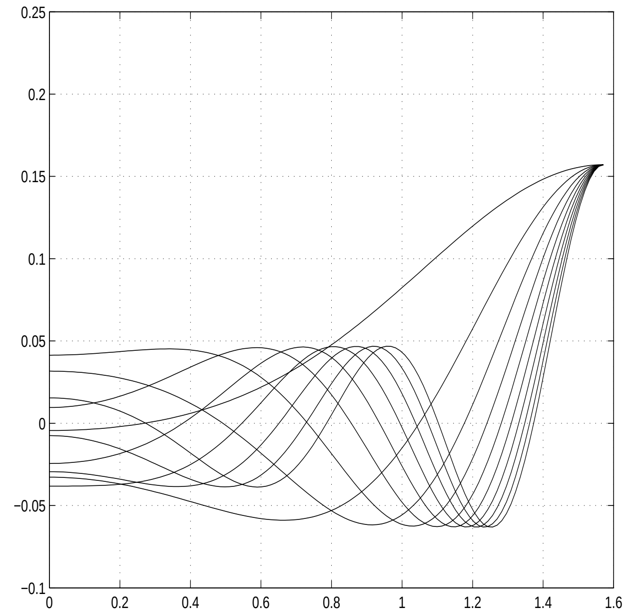
• **Let a planar antenna be comprised of a central circle and 9 concentric rings of the same area placed in the XY-plane ("Earth's surface"):**



The radius of the antenna is 1m

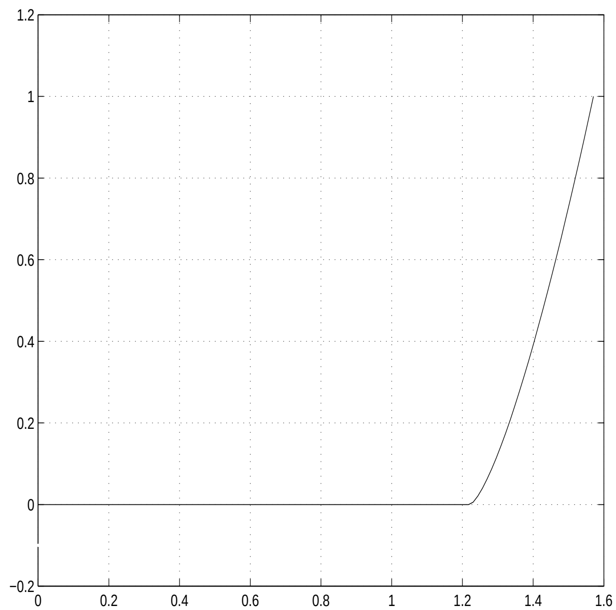
• **The diagram of a ring $\{a \le r \le b\}$ in the XY-plane is real-valued and depends on direction's altitude θ only:**

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



Diagrams of 10 rings as functions of altitude $\theta \in [0, \pi/2]$, $\lambda =$0.5m

• **Assume the target diagram to be real-valued function of the altitude "concentrated" in the angle $\frac{\pi}{2} - \frac{\pi}{12} \leq \theta \leq \frac{\pi}{2}$:**



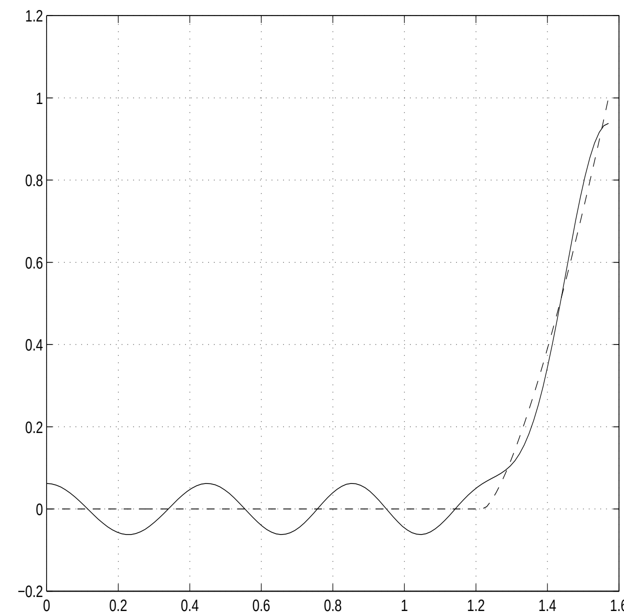
The target diagram

• **With 120-point discretization of altitudes, the Antenna Synthesis problem becomes an LP program with 11 variables and 240 linear constraints:**

$$\tau \to \min \mid -t \leq Z_*(\theta) - \Sigma_{j=1}^{10} x_j Z_j(\theta) \leq t \quad \forall \theta \in \Theta$$

$$\left[\begin{array}{c} Z_j(\theta) = \frac{1}{2} \int\limits_{r_{j-1}}^{r_j} \left[\int\limits_0^{2\pi} \rho \cos(4\pi\rho\cos(\theta)\cos(\phi)d\phi\right] d\rho \\ r_j = \sqrt{0.1j},\ j = 0, 1, ..., 10 \\ \Theta = \left\{\frac{\pi l}{240}\right\}_{l=1}^{120} \end{array}\right]$$

• **The resulting diagram approximates the target within absolute inaccuracy 0.0621:**



The target diagram (dashed) and the synthesied diagram (solid)

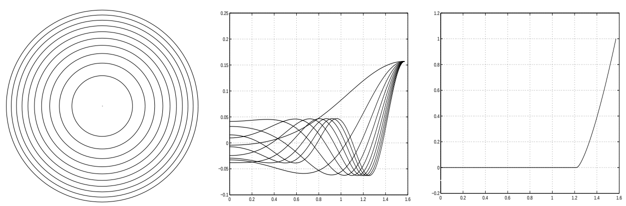
• **The optimal weights (rounded to 5 digits) are**

element #	1	2	3	4	5	6	7	8	9	10
weight	1624.4	-14700	55383	-107247	95468	19221	-138620	144870	-69303	13311

Example: Antenna Synthesis revisited

- **Recall the Antenna Synthesis example from Lecture 1:**

Given 10 planar antenna elements and a target diagram:



Elements   **Diagrams** $Z_j(\cdot)$ **Target** $Z_*(\cdot)$

approximate best of all the target diagram by a linear combination of the diagrams of the elements:

$$\min_x \left[\max_{\theta\in\Theta}\left|Z_*(\theta)-\sum_{j=1}^{10} x_j Z_j(\theta)\right|\right]$$

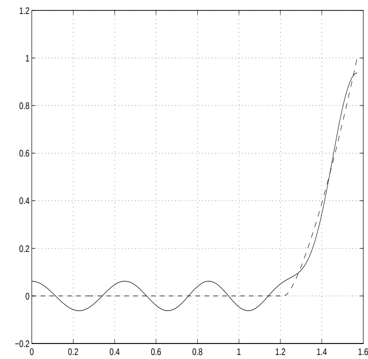
$[\Theta$: **120-point grid on** $[0,\pi/2]]$

$$\Updownarrow$$

$$\boxed{t \to \min \mid \; -t \le Z_*(\theta) - \sum_{j=1}^{10} x_j Z_j(\theta) \le t, \;\; \theta\in\Theta}$$

$$\boxed{t \to \min \mid \; -t \le Z_*(\theta) - \sum_{j=1}^{10} x_j Z_j(\theta) \le t, \;\; \theta\in\Theta}$$

$$\Downarrow$$



Target (dashed) and synthezied (solid) diagrams

$\max_\theta |Z_*(\theta)| = 1;\; t_{\rm opt} = 0.0625$

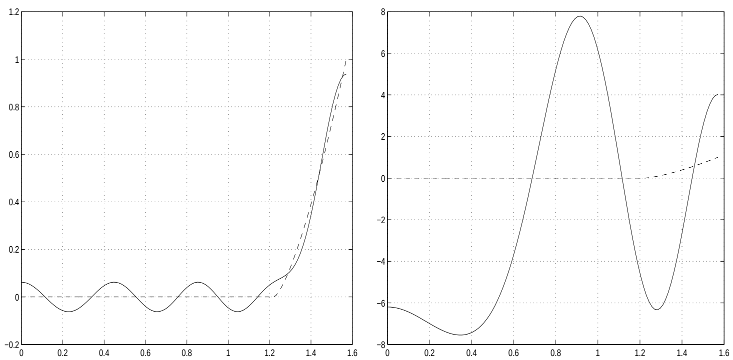
- **The optimal weights** x_j^*, $j=1,...,10$, **are characteristics of physical devices. In reality, they somehow drift around their computed values.**

What happens when the weights are affected by small (just 0.1%) random perturbations:

$$x_j = (1+\epsilon_j)x_j^*$$
$$\left[\{\epsilon_j \sim \textbf{Uniform}[-0.001, 0.001]\}_{j=1}^{10}\right] \qquad ?$$

• **The results of 0.1% "implementation errors" are disastrous:**

"Dream and reality"



"Nominal" diagram **Actual diagram**
[dashed: the target diagram]

• **The target diagram is of the uniform norm 1, and its uniform distance from the nominal diagram is** ≈ 0.06.

• **The realization of "actual diagram" shown on the picture is at the uniform distance 7.8 from the target diagram!**

$$t \to \min \mid -t \le Z_*(\theta) - \sum_{j=1}^{10} x_j Z_j(\theta) \le t,\ \theta \in \Theta$$

$$\Updownarrow$$

$$\boxed{t \to \min \mid \ Ax + ta + b \ge 0 \qquad \textbf{(LP)}}$$

• **The influence of "implementation errors"**

$$x_j \mapsto (1+\epsilon_j)x_j$$

is <u>**as if**</u> **there were no implementation errors, but the part** A **of the constraint matrix was uncertain and known "up to multiplication by a diagonal matrix with diagonal entries from** $[0.999, 1.001]$**":**

$$\mathcal{U}_{\rm ini} = \{A = A^{\rm nom}\,{\rm Diag}(1+\epsilon_1, ..., 1+\epsilon_{10}) \mid |\epsilon_j| \le 0.001\} \quad \textbf{(U)}$$

• **To improve stability of our design, we could replace the uncertain LP program (LP), (U) with its robust counterpart.**

However, to work directly with $\mathcal{U}_{\rm ini}$ **would be "too conservative" – we would ignore the fact that the implementation errors are random and independent, so that the probability for all of them to take simultaneously the "most unfavourable" values is negligibly small.**

Let us try to define the uncertainty set in a smarter way.

• **Consider a linear constraint**

$$\sum_{j=1}^{n} a_j x_j + b \geq 0 \qquad \textbf{(L)}$$

and let x_j be randomly perturbed:

$$x_j \mapsto (1+\epsilon_j)x_j$$

ϵ_j being independent unbiased random variables:

$$\mathcal{E}\{\epsilon_j\} = 0; \quad \mathcal{E}\{\epsilon_j^2\} = \sigma_j^2.$$

What is a "reliable version" of (L)?

• **With randomly perturbed x_j, the left hand side in (L) becomes a random variable:**

$$\zeta = \Sigma_{j=1}^n a_j(1+\epsilon_j)x_j + b$$

$$\left[\begin{array}{rcl} \mathrm{Mean}\{\zeta\} & \equiv & \mathcal{E}\{\zeta\} = \Sigma_{j=1}^n a_j x_j + b, \\ \mathrm{StD}\{\zeta\} & \equiv & \left(\mathcal{E}\left\{(\zeta - \mathrm{Mean}\{\zeta\})^2\right\}\right)^{1/2} = \sqrt{\Sigma_{j=1}^n \sigma_j^2 a_j^2 x_j^2}. \end{array}\right]$$

• **Let us choose a "safety parameter" κ and ignore all events where**

$$\zeta < \mathrm{Mean}\{\zeta\} - \kappa \mathrm{StD}\{\zeta\},$$

taking full responsibility for all remaining events.

With this "common sense" approach, a "reliable" version of the inequality $\zeta \geq 0$ – i.e., of (L) – becomes the inequality

$$\underbrace{\sum_{j=1}^{n} a_j x_j + b - \kappa \sqrt{\sum_{j=1}^{n} \sigma_j^2 a_j^2 x_j^2} \geq 0}_{\textbf{c.q.i.}} \qquad (\mathbf{L}_{\mathrm{rel}})$$

$$\boxed{\begin{array}{c} \Sigma_{j=1}^n a_j(1+\epsilon_j)x_j + b \geq 0 \qquad \textbf{(L)} \\ \mathcal{E}\{\epsilon_j\} = 0; \quad \mathcal{E}\{\epsilon_j^2\} = \sigma_j^2 \end{array}}$$

$$\Downarrow$$

$$\boxed{\Sigma_{j=1}^n a_j x_j + b - \kappa\sqrt{\Sigma_{j=1}^n \sigma_j^2 a_j^2 x_j^2} \geq 0 \qquad (\mathbf{L}_{\mathrm{rel}})}$$

• **Note: $(\mathbf{L}_{\mathrm{rel}})$ is exactly the robust counterpart of (L) associated with the ellipsoidal uncertainty set**

$$\mathcal{U}_\kappa = \left\{a' = a + \kappa\,\mathrm{Diag}(\sigma_1 a_1, ..., \sigma_n a_n)u \mid u^T u \leq 1\right\} \qquad \textbf{(Ell)}$$

Thus, ignoring "rare events" is equivalent to replacing the actual set $\mathcal{U}_{\mathrm{act}}$ of values of the perturbed coefficient vector

$$a' = ((1+\epsilon_1)a_1, ..., (1+\epsilon_n)a_n)^T$$

with ellipsoid (Ell).

• **Assume that ϵ_j run through $[-\epsilon, \epsilon]$. Then $\mathcal{U}_{\mathrm{act}}$ is just the box with edges $2\epsilon|a_j|$ around the "nominal" vector of coefficients a.**

It is easily seen that in the case in question

$$\mathrm{Prob}\left\{\zeta < \mathrm{Mean}\{\zeta\} - \kappa \mathrm{StD}\{\zeta\}\right\} \leq \exp\left\{-\frac{\kappa^2}{2}\right\}$$

The probability of the "rare event" we are ignoring when replacing $\mathcal{U}_{\mathrm{act}}$ with $\mathcal{U}_{5.26}$ is $< 10^{-6}$. Note that for n large, the ellipsoid $\mathcal{U}_{5.26}$ is a "negligible part" of the box $\mathcal{U}_{\mathrm{act}}$!

• **Let us apply the outlined methodology to our Antenna example:**

$$t \to \min \mid -t \leq Z_*(\theta) - \sum_{j=1}^{10} x_j Z_j(\theta) \leq t \;\forall \theta \in \Theta \qquad \textbf{(LP)}$$

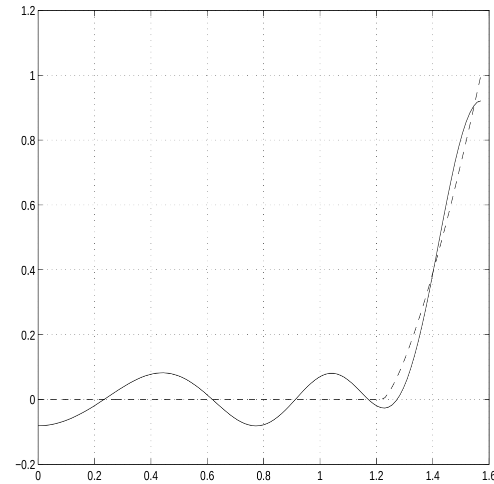
$$\Downarrow$$

$$\boxed{\begin{array}{rcl} & t \to \min & \\ Z_*(\theta) - \Sigma_{j=1}^{10} x_j Z_j(\theta) + \kappa\epsilon\sqrt{\Sigma_{j=1}^{10} x_j^2 Z_j^2(\theta)} & \leq & t \\ Z_*(\theta) - \Sigma_{j=1}^{10} x_j Z_j(\theta) - \kappa\epsilon\sqrt{\Sigma_{j=1}^{10} x_j^2 Z_j^2(\theta)} & \geq & -t \\ & \forall \theta \in \Theta & \end{array}} \qquad \textbf{(RC)}$$

$$[\epsilon = 0.001]$$

• **The results of "Robust Antenna Design"** $(\kappa = 1)$ **are as follows:**

Dream and reality



A typical "robust" diagram

• **The diagram shown on the picture is at uniform distance 0.0822 from the target (just by 30% larger than the "nominal optimal value" is 0.0622 given by "nominal design" which ignores the implementation errors)**

• **As a compensation, robust design is incomparably more stable than the nominal one: in a sample of 40 realizations of "robust diagrams", the uniform distance to the target varies from 0.0814 to 0.0830.**

- **When implementation errors become 10 times larger (1% instead of 0.1%), the "robust design" remains nearly as good as in the case of 0.1%-perturbations: now in a sample of 40 realizations of "robust diagrams", the uniform distance to the target varies from from 0.0834 to 0.116.**

$$t \to \min \mid -t \le Z_*(\theta) - \sum_{j=1}^{10} x_j Z_j(\theta) \le t \ \forall \theta \in \Theta \qquad \textbf{(LP)}$$

$$\Downarrow$$

$$\boxed{\begin{array}{rcl} & t \to \min & \\ Z_*(\theta) - \sum_{j=1}^{10} x_j Z_j(\theta) + \epsilon\sqrt{\sum_{j=1}^{10} x_j^2 Z_j^2(\theta)} & \le & t \\ Z_*(\theta) - \sum_{j=1}^{10} x_j Z_j(\theta) - \epsilon\sqrt{\sum_{j=1}^{10} x_j^2 Z_j^2(\theta)} & \ge & -t \\ & \forall \theta \in \Theta & \end{array}} \qquad \textbf{(RC)}$$

- **Why "nominal design" is so unstable?**

The basic diagrams $Z_j(\cdot)$ are "nearly linearly dependent". As a result, the nominal problem is "ill-posed" – it possesses a huge domain comprised of "nearly optimal" solutions. Indeed, look what are the optimal solutions in (LP) with added box constraints $|x_j| \le L$ on the variables:

L	1	10	10^2	10^3	10^4	10^5	10^6	10^7
Opt_Val	0.09449	0.07994	0.07358	0.06955	0.06588	0.06272	0.06215	0.06215

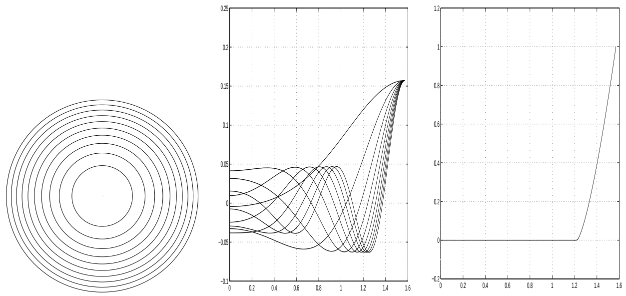
The "exactly optimal" solution to the nominal problem (LP) is very large, and even small <u>relative</u> implementation errors may destroy the design completely.

In the robust counterpart, magnitudes of candidate solutions are penalized, so that (RC) implements a smart trade-off between the optimality and the magnitude (i.e., the stability) of the solution.

j	1	2	3	4	5	6	7	8	9	10
$x_j^{\rm nom}$	1.6e3	-1.4e4	5.5e4	-1.1e5	9.5e4	1.9e4	-1.3e5	1.4e6	-6.9e4	1.3e4
$x_j^{\rm rob}$	-0.30	5.0	-3.4	-5.1	6.9	5.5	5.3	-7.5	-8.9	13

• **Consider the Least Squares Antenna Synthesis with the same data as in Lectures 1,3:**

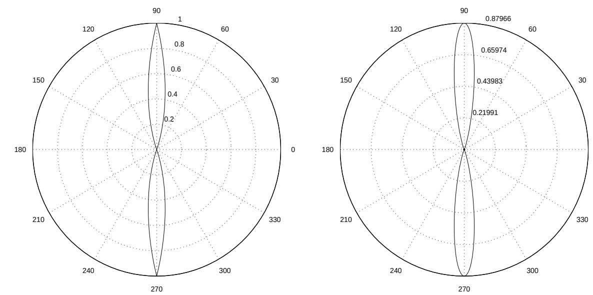
Given 10 planar antenna elements and a target diagram:



Elements **Diagrams** $Z_j(\cdot)$ **Target** $Z_*(\cdot)$

find weights x_j, $j = 1, ..., 10$ **minimizing the relative approximation error**

$$\frac{\sqrt{\int\limits_0^{\frac{\pi}{2}} |Z_*(\theta) - \sum\limits_{j=1}^{10} x_j Z_j(\theta)|^2 d\theta}}{\sqrt{\int\limits_0^{\frac{\pi}{2}} |Z_*(\theta)|^2 d\theta}}$$

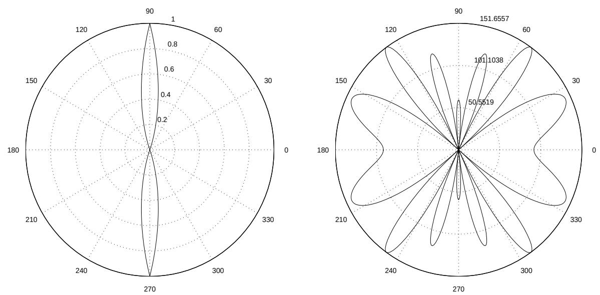


Least Squares design; relative $\|\cdot\|_2$**-error 0.072**

$\left[\begin{array}{c}\textbf{polar coordinates in the XZ-plane} \\ \textbf{left: the target; right: the design}\end{array}\right]$

• <u>Question:</u> **What happens if the "nominal" weights given by the Least Squares design are affected by 0.5% random perturbations?**

<u>Answer:</u> **A disaster. You can observe, as a (random) realization of the antenna diagram, a picture as follows:**



Nominal Least Squares design

$\left[\begin{array}{c}\textbf{left: the target; right: a sample diagram}\\ \textbf{“implementation errors” 0.5\%}\end{array}\right]$

• **Our Antenna Synthesis problem is just a particular Least Squares problem**

$$\tau \to \min \mid \| Ax - b \|_2 \leq \tau \qquad \textbf{(LS)}$$

with $(N+1) \times N$ **matrix** A **(in our example,** $N = 10$**).**

• **The effect of the "implementation errors"**

$$x_i \mapsto (1+\epsilon_i)x_i, \ \ \epsilon_i \in [-\epsilon, \epsilon]$$

is exactly the same as if there were no implementation errors at all, but the matrix A **in (LS) was uncertain and varying in the uncertainty set**

$$\mathcal{U}_\epsilon = \{A = A_* \,\mathrm{Diag}(1+\epsilon_1, ..., 1+\epsilon_N) \mid |\epsilon_i| \leq \epsilon\},$$

A_* **being the matrix of the "nominal" Least Squares problem.**

$$\tau \to \min \mid \parallel Ax - b \parallel_2 \leq \tau$$
$$A \in \mathcal{U}_\epsilon = \{A = A_* \,\mathrm{Diag}(1+\epsilon_1, ..., 1+\epsilon_N) \mid \ |\epsilon_i| \leq \epsilon\}$$

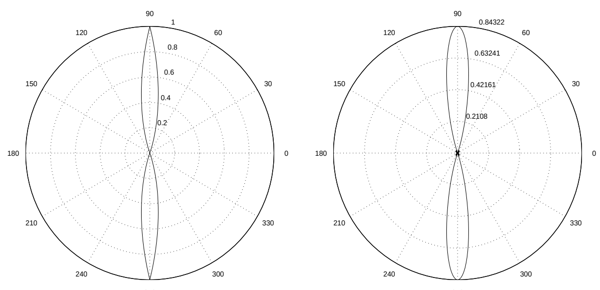
• Let us embed the uncertainty set $\mathcal{U}_\epsilon$ into the ellispoid

$$\mathcal{U} = \{A = A_* \,\mathrm{Diag}(1+\epsilon_1, ..., 1+\epsilon_N) \mid \ \sum_{i=1}^{N} \epsilon_i^2 \leq N\epsilon^2\}$$

(exploiting the fact that the implementation errors are independent, we could find a "smarter" way to define the uncertainty ellipsoid) and solve the resulting Robust Least Squares problem:

$$\tau \to \min \mid \parallel Ax - b \parallel_2 \leq \tau \quad \forall A \in \mathcal{U}.$$

As a result, we improve dramatically the stability of the design, loosing almost nothing in the optimality:



Robust Least Squares design

$$\left[\begin{array}{c}\textbf{left: the target; right: a sample diagram} \\ \textbf{“implementation errors” 0.5\%}\end{array}\right]$$

The relative approximation error all the time are less than 0.094.