ROBUST OPTIMIZATION THE NEED, THE CHALLENGE, THE SUCCESS

Aharon Ben-Tal

MINERVA Optimization Center
Technion – Israel Institute of Technology
Visiting Professor - CWI Amsterdam

DATA UNCERTAINTY IN OPTIMIZATION

♣ **Consider a generic optimization problem of the form**

$$\min_x \{f(x;\zeta) : F(x;\zeta) \in \mathbf{K}\}$$

• $x \in \mathbf{R}^n$: **decision vector** • $\zeta \in \mathbf{R}^M$: **data** • $\mathbf{K} \subset \mathbf{R}^m$: **closed convex set**

♠ **More often than not the data ζ is** *uncertain* **– not known exactly when problem is solved.**

Sources of data uncertainty:

- **part of the data is measured/estimated** ⇒ *estimation errors*
- **part of the data (e.g., future demands/prices) does not exist when problem is solved** ⇒ *prediction errors*
- **some components of a solution cannot be implemented exactly as computed** ⇒ *implementation errors* **which in many models can be mimicked by appropriate data uncertainty**

Example

Effect of implementation errors

Antenna Design (continued)

Dream: no implementation errors
Sidelobe level 0.018

Reality: $x_j \mapsto (1+\epsilon_j)x_j$
$[\epsilon_j \sim \textbf{Uniform}[-0.001, 0.001]]$
100-diagram sample: Sidelobe level $\in [97, 536]$

Reality: $x_j \mapsto (1+\epsilon_j)x_j$
$[\epsilon_j \sim \textbf{Uniform}[-0.02, 0.02]]$
100-diagram sample: Sidelobe level $\in [2469, 11552]$

Example

Effect of data inaccuracy

Data Uncertainty in Optimization

♣ **Consider a real-world LP program `PILOT4` from the `NETLIB` library (1,000 variables, 410 constraints). The constraint # 372 is:**

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

♠ **Most of the coefficients are “ugly reals” (like -15.79081 or -84.644257). It is highly unlikely that the corresponding real-life parameters are known to high accuracy, so that the ugly coefficients can be thought of as uncertain – not known exactly.**

The only exception is the coefficient `1` at x_{880} – it perhaps reflects the structure of the problem and might be exact.

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

(?) What happens with the constraint, evaluated at the nominal solution $x^{\rm n}$ as reported by `CPLEX`, when the accuracy in the uncertain data is 0.1%:

$$|a_i^{\rm true} - a_i^{\rm n}| \leq 0.001|a_i^{\rm n}| \qquad (*)$$

- **In the worst case, the constraint can be violated by as much as 450%:**

$$\min_{a^{\rm true}} \left\{[a^{\rm true}]^T x^{\rm n} \middle|\, a^{\rm true} \textbf{ satisfies } (*)\right\} - b < -128.2 \approx 4.5|b|.$$

- **Assuming "random uncertainty":**

$$a_i^{\text{true}} = a_i^{\text{n}} + \epsilon_i |a_i^{\text{n}}|, \quad \epsilon_i \sim \textbf{Uniform}[-0.001, 0.001]$$

and running 1,000 simulations, we come to the results as follows:

$\textbf{Prob}\{V > 0\}$	$\textbf{Prob}\{V > 150\%\}$	$\textbf{Mean}(V)$
0.50	**0.18**	**125%**

$$V = \max\left[\frac{b - (a^{\text{true}})^T x^{\text{n}}}{|b|}, 0\right]$$

⇒ The nominal solution is highly "unreliable" – small perturbations of (clearly uncertain!) data entries can make the solution heavily infeasible...

♣ Among 90 `NETLIB` LP problems,

- **In 19 problems 0.01%-perturbations of "clearly uncertain" data result in more than 5%-violations of (some of) the constraints as evaluated at the nominal optimal solution;**
- **In 13 of these 19 problems, 0.01%-perturbations of "clearly uncertain" data result in more than 50%-violations of the constraints!**

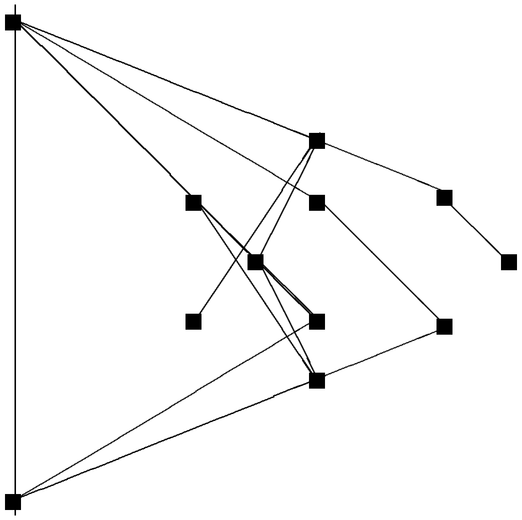
Example

Effect of uncertain predictions

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:

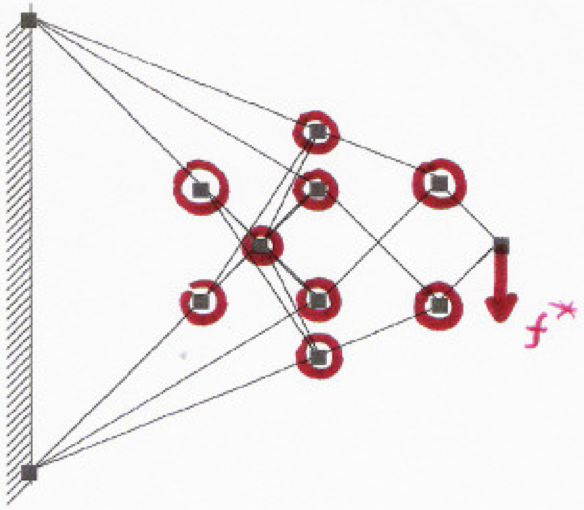
9×9 ground structure and the load of interest

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



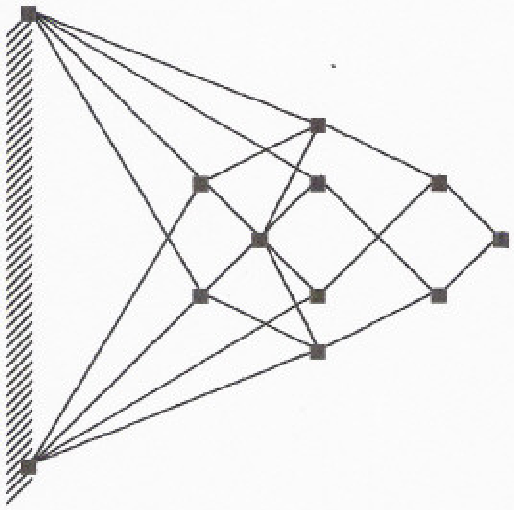
Optimal cantilever (single-load design)
the compliance is 1.000

- **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

♠ **With traditional modelling methodology,**

- **"large" data uncertainty is modelled in a stochastic fashion and then processed via Stochastic Programming techniques**

> **Fact:** *In many cases, it is difficult to specify reliably the distribution of uncertain data and/or to process the resulting Stochastic Programming program.*

♠ **The ultimate goal of** *Robust Optimization* **is to take into account data uncertainty already at the modelling stage in order to "immunize" solutions against uncertainty.**

- **In contrast to Stochastic Programming, Robust Optimization does not assume stochastic nature of the uncertain data (although can utilize, to some extent, this nature, if any).**

“NON-ADJUSTABLE” ROBUST OPTIMIZATION:

Robust Counterpart of Uncertain Problem

$$\boxed{\min_x \{f(x,\zeta) : F(x,\zeta) \in \mathbf{K}\}} \qquad (\mathbf{U})$$

♣ **The initial (“Non-Adjustable”) Robust Optimization paradigm (Soyster ’73, B-T&N ’97–, El Ghaoui et al. ’97–, Bertsimas&Sim ’03–,...) is based on the following tacitly accepted assumptions:**

A.1. All decision variables in (U) represent “here and now” decisions which should get specific numerical values as a result of solving the problem and *before* the actual data “reveals itself”.

A.2. The uncertain data are “unknown but bounded”: one can specify an appropriate (typically, bounded) *uncertainty set* $\mathcal{U} \subset \mathbf{R}^M$ of possible values of the data. The decision maker is fully responsible for consequences of the decisions to be made when, and only when, the actual data is within this set.

A.3. The constraints in (U) are “hard” – we cannot tolerate violations of constraints, even small ones, when the data is in $\mathcal{U}$.

$$\boxed{\begin{array}{c}\min_x \{f(x,\zeta) : F(x,\zeta) \in \mathbf{K}\}\\ \zeta \in \mathcal{U}\end{array}} \qquad (\mathbf{U})$$

♠ **Conclusions:**

• **The only meaningful candidate solutions are the** *robust* **ones – those which remain feasible whatever be a realization of the data from the uncertainty set:**

$$x \text{ robust feasible } \Leftrightarrow\; F(x,\zeta) \in \mathbf{K}\;\; \forall \zeta \in \mathcal{U}$$

• **"Robust optimal" solution to be used is a robust solution with the smallest possible** *guaranteed* **value of the objective, that is, the optimal solution of the optimization problem**

$$\min_{x,t} \{t : f(x,\zeta) \le t,\; F(x,\zeta) \in \mathbf{K}\;\; \forall \zeta \in \mathcal{U}\} \qquad (\mathbf{RC})$$

called the *Robust Counterpart* **of** (**U**).

Optimization Problems with Uncertain Data

♣ **A generic optimization problem is of the form**

$$\min_x \{ f(x;\zeta) \mid F(x;\zeta) \leq 0 \} \qquad (\mathbf{P}_\zeta)$$

- x **is the design vector**
- f, F **are specified by the description of the problem**
- ζ **is a finite-dimensional vector specifying the data.**

Example 1. Linear Programming:

$$\min_x \{c^T x \mid Ax \leq b\} \qquad [\zeta = (c, A, b)]$$

Example 2. Convex Quadratic Programming:

$$\min_x \{c^T x \mid x_i^T A_i^T A_i x_i - 2b_i^T x_i + c_i \leq 0,\ i = 1, ..., m\} \qquad [\zeta = (c, \{A_i, b_i, c_i\}_{i=1}^m)]$$

Example 3. Conic Quadratic Programming:

$$\min_x \{c^T x \mid \| A_i x - b_i \|_2 \leq c_i^T x - d_i,\ i = 1, ..., m\} \qquad [\zeta = (c, \{A_i, b_i, c_i, d_i\}_{i=1}^m)]$$

Example 4. Semidefinite Programming:

$$\min_x \left\{c^T x \mid A_0 + \sum_{j=1}^{\dim x} x_i A_i \succeq 0\right\} \qquad [\zeta = (c, A_0, ..., A_{\dim x})]$$

Semi-Infinite Conic Programs

♣ Conic Program:

$$\min_x \left\{c^T x : Ax - b \in \mathbf{K}\right\} \qquad \textbf{(C)}$$

- (c, A, b) – **problem's data**
- **closed pointed convex cone K,** $\mathrm{int}\,\mathbf{K} \neq \emptyset$**, in a Euclidean space – problem's structure**

Examples:

- **Linear Programming: $\mathbf{K} = \mathbf{R}^n_+$**
- **Conic Quadratic Programming: K is a direct product of Lorentz cones**

  $$\mathbf{L}^k = \{y \in \mathbf{R}^k : y_k \geq \sqrt{y_1^2 + ... + y_{k-1}^2}\}$$

- **Semidefinite Programming: $\mathbf{K} = \mathbf{S}^n_+$ is the cone of positive semidefinite matrices in the space $\mathbf{S}^n$ of $n \times n$ symmetric matrices**

♣ Semi-Infinite Conic Program:

$$\min_x \left\{c^T x : Ax - b \in \mathbf{K}\, \forall [A, b] \in \mathcal{U}\right\}$$

where $\mathcal{U}$ is a given "uncertainty set" (assumed to be convex and compact).

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

♣ **The main mathematical question associated with semi-infinite problem (S) is:**

> **(?) When and how (S) can be reformulated as a "computationally tractable" optimization problem?**

The answer to (?) clearly depends on the interplay between the geometries of the cone K and of the uncertainty set $\mathcal{U}$.

Robust Linear Programming

Focus on a *single* uncertainty-affected linear inequality—a family

$$\left\{a^T x \leq b\right\}_{[a;b]\in\mathcal{U}}$$

of linear inequalities with the data varying in the uncertainty set

$$\mathcal{U} = \left\{[a;b] = \left[a^0;b^0\right] + \sum_{\ell=1}^{L} \zeta_\ell \left[a^\ell;b^\ell\right];\ \zeta \in \mathcal{Z}\right\}$$

and on "tractable representation" of the RC

$$a^T x \leq b\,\forall\left([a;b] = \left[a^0;b^0\right] + \sum_{\ell=1}^{L} \zeta_\ell \left[a^\ell;b^\ell\right];\ \zeta \in \mathcal{Z}\right) \tag{2}$$

of this uncertain inequality

Theorem *Let the perturbation set* Z *be given by*

$$Z = \{\zeta \in \mathbb{R}^L \mid \exists\, u \in \mathbb{R}^K : P\zeta + Qu + p \in K\}$$

where K *is a closed convex pointed cone in* $\mathbb{R}^N$ *which is either polyhedral, or is such that*

$$\exists\, \bar{\zeta}, \bar{u} : P\bar{\zeta} + Q\bar{u} + p \in int\ K\,.$$

Consider the robust counterpart of a linear inequality:

$$a^T x \le b \quad \forall \begin{pmatrix} a \\ b \end{pmatrix} \in \mathcal{U}$$

where

$$\mathcal{U} = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a^o \\ b^o \end{pmatrix} + \sum_{\ell=1}^{L} \zeta_\ell \begin{pmatrix} a^\ell \\ b^\ell \end{pmatrix} \,\middle|\, \zeta \in Z \right\}\,.$$

Then a vector $x \in \mathbb{R}^n$ is robust feasible *if and only if* $\exists\, y \in \mathbb{R}^L$, which together with x satisfies the following linear/conic inequalities:

$$
\begin{aligned}
& p^T y + (a^o)^T x \le b_0 \\
& Q^T y = 0 \\
& (P^T y)_\ell + (a^\ell)^T x = b^\ell, \quad \ell = 1, 2, \ldots, L \\
& y \in K_*
\end{aligned}
$$

where $K_* = \{y \mid y^T v \ge 0, \forall\, v \in K\}$ is the dual cone of K.

EXAMPLE

$$(1) \qquad a^T x \le b \qquad \binom{a}{b} \in \mathcal{U}$$

$$\mathcal{U} = \left\{ \binom{a}{b} \middle| \binom{a_0}{b_0} + \sum_{i=1}^{L} \zeta_\ell \binom{a_\ell}{b_\ell} \middle| \zeta \in Z \right\}$$

Ball-Box uncertainty

$$\swarrow$$

$$Z = \{\zeta \in \mathbb{R}^L \mid \|\zeta\|_\infty \le 1, \ \|\zeta\|_2 \le \rho\}$$

RESULT:

x solves the semi-infinite linear system

IFF (x, z, w) solves the SOCQ system

$$\left\{ \begin{array}{l} \displaystyle\sum_{\ell=1}^{L} |z_\ell| + \Omega \sqrt{\sum_{\ell=1}^{L} w_\ell^2} \le b_0 - a_0^T x \\ z_\ell + w_\ell = b_\ell - a_\ell^T x \qquad \ell = 1 \dots L \end{array} \right.$$

Illustration: Single-Period Portfolio Selection

There are 200 assets. Asset #200 (“money in the bank”) has yearly return $r_{200} = 1.05$ and zero variability. The yearly returns r_ℓ, $\ell = 1, \dots, 199$ of the remaining assets are independent random variables taking values in the segments $[\mu_\ell - \sigma_\ell, \mu_\ell + \sigma_\ell]$ with expected values μ_ℓ; here

$$\mu_\ell = 1.05 + 0.3\frac{(200-\ell)}{199}, \ \sigma_\ell = 0.05 + 0.6\frac{(200-\ell)}{199}, \ \ell = 1, \dots, 199\,.$$

The goal is to distribute $1 between the assets in order to maximize the return of the resulting portfolio, the required risk level being $\varepsilon = 0.5\%$.

We want to solve the uncertain LO problem

$$\max_{y,t} \left\{ t : \sum_{t=1}^{199} r_\ell y_\ell + r_{200} y_{200} - t \geq 0, \sum_{\ell=0}^{200} y_\ell = 1, \ y_\ell \geq 0 \, \forall \, \ell \right\},$$

where y_ℓ is the capital to be invested into asset $\#\ell$.

The uncertain data are the returns r_ℓ, $\ell = 1, \ldots, 199$; their natural parameterization is

$$r_\ell = \mu_\ell + \sigma_\ell \zeta_\ell \,,$$

where ζ_ℓ, $\ell = 1, \ldots, 199$, are independent random perturbations with zero mean varying in the segments $[-1, 1]$. Setting $x = [y; -t] \in \mathbb{R}^{201}$, the problem becomes

$$\left\{\begin{array}{lll} \text{minimize} & x_{201} & \\ \text{subject to} & & \\ (a) & \left[a^0 + \sum_{\ell=1}^{199} \zeta_\ell a^\ell\right]^T x - \left[b^0 + \sum_{\ell=1}^{199} \zeta_\ell b^\ell\right] \leq 0 & \\ (b) & \sum_{j=1}^{200} x_\ell = 1 & \\ (c) & x_\ell \geq 0, \; \ell = 1, \ldots, 200 & \end{array}\right. \tag{4}$$

where

$$\begin{array}{rcl} a^0 & = & [-\mu_1; -\mu_2; \ldots; -\mu_{199}; -r_{200}; -1]; a^\ell = \sigma_\ell \cdot [0_{\ell-1,1}; 1; 0_{201-\ell,1}], \; \ell = 1, \ldots, 199; \\ b^\ell & = & 0, \; \ell = 0, 1, \ldots, 199 \,. \end{array}$$

The only uncertain constraint in the problem is the linear inequality (a). We consider 3 perturbation sets along with the associated robust counterparts of problem (4).

1. *Box RC* which ignores the information on the stochastic nature of the perturbations affecting the uncertain inequality and uses the only fact that these perturbations vary in $[-1, 1]$. The underlying perturbation set $\mathcal{Z}$ for (a) is $\{\zeta : \|\zeta\|_\infty \leq 1\}$;
2. *Ball-Box* with the safety parameter $\Omega = \sqrt{2\ln(1/\varepsilon)} = 3.255$, which ensures that the optimal solution of the associated RC (a CQ prob.) satisfies (a) with probability at least $1 - \varepsilon = 0.995$. The underlying perturbation set $\mathcal{Z}$ for (a) is $\{\zeta : \|\zeta\|_\infty \leq 1\}, \ \|\zeta\|_2 \leq 3.255\}$;
3. *Budgeted uncertainties* with the uncertainty budget $\gamma = \sqrt{2\ln(1/\varepsilon)}\sqrt{199} = 45.921$, which results in the same probabilistic guarantees as for the Ball-Box RC. The underlying perturbation set $\mathcal{Z}$ for (a) is $\{\zeta : \|\zeta\|_\infty \leq 1\}, \ \|\zeta\|_1 \leq 45.921\}$;

Results

Box RC. The associated RC is the LP

$$\max_{y,t}\left\{ t : \begin{array}{l} \sum_{\ell=1}^{199}(\mu_\ell - \sigma_\ell)y_\ell + 1.05 y_{200} \geq t \\ \sum_{\ell=1}^{200} y_\ell = 1,\ y \geq 0 \end{array} \right\};$$

as it should be expected, this is nothing but the instance of our uncertain problem corresponding to the worst possible values $r_\ell = \mu_\ell - \sigma_\ell,\ \ \ell = 1, \ldots, 199$, of the uncertain returns. Since these values are less than the guaranteed return for money, the robust optimal solution prescribes to keep our initial capital in the bank with guaranteed yearly return 1.05.

Ball-Box RC. The associated RC is the conic quadratic problem

$$\max_{y,z,w,t}\left\{t: \begin{array}{l} \sum_{\ell=1}^{199}(\mu_\ell y_\ell + 1.05y_{200} - \sum_{\ell=1}^{199}|z_\ell| - 3.255\sqrt{\sum_{\ell=1}^{199} w_\ell^2} \geq t \\ z_\ell + w_\ell = y_\ell, \quad \ell = 1,\dots,199, \ \sum_{\ell=1}^{200} y_\ell = 1, \ y \geq 0 \end{array}\right\}.$$

The robust optimal value is 1.1200, meaning 12.0% profit with risk as low as $\varepsilon = 0.5\%$.

Semi-Infinite Convex Quadratic Programming

♣ **Tractability of a semi-infinite convex quadratically constrained program**

$$\min_x \left\{ c^T x : \underbrace{x^T A_i^T A_i x + 2b_i^T x + c_i \leq 0}_{\Updownarrow \atop \left[\begin{array}{c} 2A_i x \\ 1 + c_i + 2b_i^T x \\ 1 - c_i - 2b_i^T x \end{array}\right] \in \mathbf{L}}, \forall \left(i = 1, ..., m, \{A_i, b_i, c_i\}_{i=1}^m \in \mathcal{U} \right) \right\}$$

reduces to tractability of a single semi-infinite convex quadratic constraint

$$x^T A^T A x + 2b^T x + c \leq 0 \quad \forall (A, b, c) \in \mathcal{V} \qquad (*)$$

♣ **Fact:** $(*)$ **can be NP-hard already for a simple polyhedral set** $\mathcal{V}$**.**
⇒ **We need to look for tight tractable approximations of** $(*)$**.**

- **What is "an approximation" ?**
- **How to measure the quality of an approximation?**

The RC of a Quadratically Constrained Problem

$$\min_{x\in\mathbb{R}^n} \left\{\gamma^T x : x^T A^T A x \leq 2b^T x + c, \quad \forall (A,b,c)\in\mathcal{U}\right\}, \tag{UQ}$$

with single ellipsoid uncertainty, namely

$$\mathcal{U} = \left\{(A,b,c) = (A^0,b^0,c^0) + \sum_{\ell=1}^{L} u_\ell (A^\ell, b^\ell, c^\ell) : \|u\|_2 \leq 1\right\}. \tag{1}$$

Theorem 1 *A robust counterpart of (UQ) with the uncertainty set $\mathcal{U}$ given by (1) is equivalent to the SDP problem*

$$\min_{(x,\lambda)\in\mathbb{R}^n\times\mathbb{R}} \gamma^T x$$

$$s.t. \quad \left(\begin{array}{c|ccc|c}
c^0 + 2x^T b^0 - \lambda & \frac{1}{2}c^1 + x^T b^1 & \cdots & \frac{1}{2}c^L + x^T b^L & (A^0 x)^T \\
\hline
\frac{1}{2}c^1 + x^T b^1 & \lambda & & & (A^1 x)^T \\
\vdots & & \ddots & & \vdots \\
\frac{1}{2}c^L + x^T b^L & & & \lambda & (A^L x)^T \\
\hline
A^0 x & A^1 x & \cdots & A^L x & I_m
\end{array}\right) \succeq 0, \quad (RUQ)$$

Approximations of Semi-Infinite Conic Constraints

$$Ax+b\in\mathbf{K}\quad\forall[A,b]\in\mathcal{U}\tag{SC}$$

♣ **Assume (as it is usually is the case in applications) that the uncertainty set $\mathcal{U}$ is given as**

$$\mathcal{U}=\left\{[A^0,b^0]+\rho\sum_{\ell=1}^{L}\zeta_\ell[A^\ell,b^\ell]:\zeta\in\mathcal{Z}\right\}$$

- $[A^0,b^0]$: **nominal data**
- $\mathcal{Z}$: **set of data perturbations** ζ **"of magnitude** ≤ 1**"**
- $\rho\geq 0$: **level of perturbations**

♠ **With this assumption, (SC) becomes a member of a family of semi-infinite conic constraints**

$$\textbf{(SC)}$$
$$\Updownarrow$$
$$A^0x+b^0+\rho\sum_\ell(A^\ell x+b^\ell)\in\mathbf{K}\quad\forall\zeta\in\mathcal{Z}\tag{SC[ρ]}$$

$$A^0x + b^0 + \rho\sum_\ell(A^\ell x + b^\ell) \in \mathbf{K} \quad \forall \zeta \in \mathcal{Z} \qquad \textbf{(SC}[\rho]\textbf{)}$$

Definition: **A Linear Matrix Inequality**

$$\begin{array}{c} A_\rho(x,u) + B_\rho \succeq 0 \\ [A_\rho(x,y) \in \mathbf{S}^n \text{ is linear in } (x,u)] \end{array} \qquad \textbf{(LMI}[\rho]\textbf{)}$$

is called an approximation of (SC$[\rho]$) tight within factor ϑ, if

(i) whenever x can be extended to a feasible solution (x,u) of (LMI$[\rho]$), x is feasible for (SC$[\rho]$), and

(ii) whenever x can<u>not</u> be extended to a feasible solution of (LMI$[\rho]$), x is <u>not</u> feasible for (SC$[\vartheta\rho]$).

Approximating Quadratic Constraints
Basic Fact

♣ **Consider a nonconvex problem of quadratic maximization**

$$\mathbf{OPT}(\rho) = \max\left\{x^TAx + 2b^Tx : x^TQ_jx \leq \rho^2,\ j = 1, ..., J\right\} \qquad [Q_j \succeq 0, \sum_j Q_j \succ 0]$$

along with the standard semidefinite relaxation of this problem:

$$\mathbf{SDP}(\rho) = \max_{\lambda,\mu}\left\{\mu + \rho^2\sum_j\lambda_j : \left[\begin{array}{c|c}\mu & b^T \\ \hline b & \sum_j\lambda_jQ_j - A\end{array}\right] \succeq 0,\ \lambda \geq 0\right\}.$$

Then $\mathbf{SDP}(\rho) \geq \mathbf{OPT}(\rho)$ (trivial) and

- **$\mathbf{SDP}(\rho) = \mathbf{OPT}(\rho)$ when $J = 1$ (trivial),**
- **$\mathbf{SDP}(\rho) \leq \frac{\pi}{2}\cdot\mathbf{OPT}(\rho)$ when $A \succeq 0$, $b = 0$ and the matrices $Q_1, ..., Q_J$ commute with each other (Nesterov '97),**
- **$\mathbf{SDP}(\rho) \leq \mathbf{OPT}(\vartheta\rho) \leq \vartheta^2\mathbf{OPT}(\rho)$ with $\vartheta \leq O(1)\sqrt{\ln(J)}$ (Ben-Tal, Nem., Roos '02, Nem. 03).**

Approximating Semi-Infinite Convex Quadratic Constraint with Ellipsoidal Uncertainty

Theorem. **[Ben-Tal, Nem., Roos '02] Consider semi-infinite convex quadratic inequality**

$$x^TA^TAx+2b^Tx+c\le 0 \quad \forall \left((A,b,c)=(A_*,b_*,c_*)+\rho\sum_{\ell=1}^{L}\zeta_\ell(A^\ell,b^\ell,c^\ell):\zeta\in\mathcal{Z}\right) \qquad \textbf{(SC}[\rho]\textbf{)}$$

where the perturbation set $\mathcal{Z}$ is an intersection of ellipsoids centered at the origin:

$$\mathcal{Z}=\{\zeta:\zeta^TQ_j\zeta\le 1, j=1,...,N\}$$
$$\left[Q_j\succeq 0,\ \sum_j Q_j\succ 0\right]$$

(SC$[\rho]$) admits an explicit ϑ-approximation, where ϑ is as follows:

♠ **In the case $N=1$ of simple ellipsoidal uncertainty, $\vartheta=1$;**

♠ **In the case of box uncertainty ($N=\dim\zeta$, $\zeta^TQ_j\zeta=\zeta_j^2$), $\vartheta=\frac{\pi}{2}$;**

♠ **In the general case, $\vartheta\le O(1)\sqrt{\ln N}$. For example,**

$$LN\le 100,000,000\Rightarrow\vartheta\le 6.36.$$

Semi-Infinite Semidefinite Programming

♣ **A semi-infinite LMI**

$$\mathcal{A}[x] \equiv \sum_{\ell} x_\ell A_\ell + B \succeq 0 \quad \forall \mathcal{A}[\cdot] \in \mathcal{U}$$

can be NP-hard already in the simplest cases where $\mathcal{U}$ is a box or an ellipsoid.

♣ **The strongest generic result on tight tractable approximations of Semi-Infinite LMIs deals with the case of structured norm-bounded perturbations:**

$$\mathcal{U} = \left\{ \mathcal{A}[x] = \mathcal{A}_*[x] + \sum_{\ell=1}^{L} \left[L_\ell^T[x] \Delta_\ell R_\ell + R_\ell^T \Delta_\ell^T L_\ell[x] \right] : \begin{array}{l} \Delta_\ell \in \mathbf{R}^{d_\ell \times d_\ell} \\ \|\Delta_\ell\| \leq \rho \\ \Delta_\ell = \delta_\ell I_{d_\ell},\ \ell \in I^{\rm s} \end{array} \right\}$$

- $A_*[x]$: **symmetric $m \times m$ matrix affinely depending on x**
- $L_\ell[x]$: **$d_\ell \times m$ matrix affinely depending on x**

$$\mathcal{A}_*[x] + \sum_{\ell=1}^{L} \left[L_\ell^T[x]\Delta_\ell R_\ell + R_\ell^T \Delta_\ell^T L_\ell[x]\right] \succeq 0 \;\forall \left(\{\Delta_\ell\}_{\ell=1}^L : \begin{array}{l} \Delta_\ell \in \mathbf{R}^{d_\ell \times d_\ell} \\ \|\Delta_\ell\| \leq \rho \\ \Delta_\ell = \delta_\ell I_{d_\ell},\ \ell \in I^{\rm s} \end{array} \right) \quad \textbf{(SLMI)}$$

♣ **Theorem:** **[Ben-Tal, Nem., Roos '02] (SLMI) admits a ϑ-tight computationally tractable approximation, where ϑ depends solely on the maximum of sizes d_ℓ of scalar perturbation blocks**

$$\mu = \max\left\{d_\ell : \ell \in I^{\rm s}\right\}$$

specifically,

$$\vartheta \leq \frac{\pi\sqrt{\mu}}{2}.$$

If there are no scalar perturbation blocks, or all scalar perturbation blocks are of size 1, then

$$\vartheta = \frac{\pi}{2}.$$

In the case of a single perturbation block, $\vartheta = 1$.

♣ When processing well-structured convex programs (e.g., Linear/Conic Quadratic/Semidefinite ones) with *uncertain* and *stochastic* data, the entity of primary interest is a *chance constrained conic inequality*

$$\text{Prob}\{A_\zeta x + b_\zeta \in \mathbf{K}\} \geq 1 - \epsilon$$

- x: decision vector
- ζ: random perturbation
- A_ζ, b_ζ: affine in ζ
- $\mathbf{K}$: simple cone ($\mathbb{R}_+/\mathbb{R}^n_+$/Lorentz/Semidefinite)
- $\epsilon \ll 1$: a given tolerance

♠ Chance constraints (primarily with $\mathbf{K} = \mathbb{R}_+$, $\mathbf{K} = \mathbb{R}^n_+$) were introduced by A. Charnes, W. Cooper, G. Symonds (1958) and intensively studied by many authors (T. Badics, D. Dentcheva, A. Dupačová, L. Miller, A. Prékopa, A. Ruszczynski, B. Vizvari, H. Wagner,...)

$$p(w) \equiv \text{Prob}\left\{\zeta^w \equiv w_0 + \sum_{\ell=1}^d \zeta_\ell w_\ell \notin \mathbf{K}\right\} \leq \epsilon \qquad (C)$$

♣ In general, (C) can be difficult to process:

- The feasible set X of (C) can be nonconvex, which makes it problematic to optimize under the constraint
- Even when convex, X can be "computationally intractable"

  Let $\mathbf{K} = \mathbb{R}_+$ and $\zeta \sim \text{Uniform}([0,1]^d)$. In this case, X is convex (Lagoa et al, 2005); however, *unless P=NP, there is no algorithm capable to compute $p(w)$ within accuracy δ in time polynomial in the size of the (rational) data w and in* $\ln(1/\delta)$ (L. Khachiyan, 1989).

♣ When (C) is difficult to process "as it is", one can look for a *safe tractable approximation of (C) – a computationally tractable convex set W_ϵ such that $W_\epsilon \subset X \equiv \{w : p(w) \leq \epsilon\}$.*

Probabilistic guaranties via RO

$$(1) \quad f_0(y) + \sum_{l=1}^{d} \xi_l f_l(y) \le 0$$

Assumption $\xi_1, \xi_2, \ldots, \xi_d$ independent rv's

$\xi_l \sim \mathrm{P}_l \in \mathcal{P}_l$ *compact all prob. dist. in $\mathcal{P}_l$ has common support = [–1,1].

Def. A vector **y** satisfying, for a given $0 < \xi < 1$:

$$(2) \qquad \Pr\{ f_0(y) + \Sigma\, \xi_l f_l(y) \le 0 \} \ge 1 - \varepsilon$$

provides a safe approximation of (1).

Challenge Find uncertainty set for ξ, U s.t. the RC of (1):

$$(3) \qquad f_0(y) + \Sigma\, \xi_l f_l(y) \le 0 \;\forall\; \xi \in U$$

is a safe approximation of (1), i.e. satisfies (2).

Theorem

(4) where
$$
\begin{aligned}
U &= B \cap (M + E)\\
B &= \left\{ \xi \in \Re^d \,\middle|\, \|\xi\|_\infty \le 1 \right\}\\
M &= \left\{ \xi \,\middle|\, \mu_l^- \le \xi_l \le \mu_l^+, l = 1, \ldots d \right\}\\
E &= \left\{ \xi \,\middle|\, \sum \xi_l^2 / \sigma_l^2 \le 2 \log(1/\varepsilon) \right\}
\end{aligned}
$$

μ_l^-, μ_l^+ and σ_l are such that

$$A_l(z) \le \max\left(\mu_l^- z, \mu_l^+ z\right) + \frac{\sigma_l^2}{2} z_l^2, \forall l = 1, \ldots d$$

where $A_l(z) = \max_{\mathrm{P}_l \in \mathcal{P}_l} \log\left(\int \exp(zs) d\mathrm{P}_l(s)\right).$

Also E can be replaced by a scaled

$$E = \left\{ \xi \,\middle|\, \sum |\xi_l| / \sigma_l \le \sqrt{2d \log(1/\varepsilon).} \right\}$$

Values of $\mu_l^{\pm}$, σ_l are explicitly known for various families $\mathcal{P}_l$, e.g.

$$\left.\begin{array}{c}\mathrm{supp}(\mathcal{P})\subset[-1,1]\\ \text{P unimodal and symmetric}\end{array}\right\}\begin{array}{c}\mu_l^{\pm}=0\\ \sigma_l=\sqrt{1/6}\end{array}$$

$$\left.\begin{array}{c}\mathrm{supp}(\mathcal{P})\subset[-1,1]\\ \text{P unimodal}\end{array}\right\}\begin{array}{c}\mu_l^{-}=-1/2,\mu_l^{+}=1/2\\ \sigma_l=\sqrt{1/24}\end{array}$$

Moreover, for the LP case ($f_l(y)$ affine) the RC of (3), with U as in the Theorem, is conic quadratic or LP.

(1) $a(\zeta)^T x \le b(\zeta)$

$$a(\zeta) = a^0 + \sum_{\ell=1}^{L} \zeta_\ell a^\ell, \quad b(\zeta) = b^0 + \sum_{\ell=1}^{L} \zeta_\ell b^\ell$$

$$\zeta_1, \ldots, \zeta_L \ \text{ i.i.d.}, \ E(\zeta_\ell) = 0\,, \quad |\zeta_\ell| \le 1$$

$$(\mathrm{CC})_\varepsilon \ \boxed{\mathrm{Prob}_\zeta \left(a(\zeta)^T x \le b(\zeta)\right) \ge 1 - \varepsilon}$$

Let $\mathcal{U}_\Omega = \left\{\zeta \in I\!R^L \mid \|\zeta\|_2 \le \Omega\right\}$.

Consider the RC of (1) w.r.t. $\mathcal{U}_\Omega$:

$$a(\zeta)^T x \le b(\zeta) \quad \forall\, \zeta \in \mathcal{U}_\Omega$$

which we already know is equivalent to

$$(\mathrm{RC})_\Omega \ \boxed{(a^0)^T x + \Omega \sqrt{\textstyle\sum_{\ell=1}^{L} \left((a^\ell)^T x - b_\ell\right)^2} \le b^0}$$

Theorem 1 *If x solves $(\mathrm{RC})_\Omega$ with $\Omega \ge \sqrt{2\log(1/\varepsilon)}$, then x solves $(\mathrm{CC})_\varepsilon$*

$$\mathrm{OR:} \quad \left\{ \begin{array}{l} x \ \textit{solves}\ (\mathrm{RC})_\Omega \quad \textit{then}\ x\ \textit{solves} \\ (\mathrm{CC})_\varepsilon \quad \textit{with} \quad \varepsilon < e^{-\Omega^2/2} \end{array} \right.$$

e.g., $\Omega = 7.44$, $1 - \varepsilon = 1 - 10^{-12}$.

Example (one of many): Range and mean a priori information on ζ_ℓ. Consider a chance constraint

$$\begin{array}{c}\mathrm{Prob}\{w_0(x)+\sum_{\ell=1}^d \zeta_\ell w_\ell(x)>0\}\leq\epsilon\\ \bullet\ \zeta_\ell\in[-1,1]\text{: independent }\bullet\ \mathbf{E}\{\zeta_\ell\}\in[\mu_\ell^-,\mu_\ell^+]\end{array}\qquad (C)$$

The Bernstein approximation of (C) is

$$\forall \zeta\in\mathcal{Z}=\{\zeta:\sum_{\ell=1}^d \phi_\ell(\zeta_\ell)\leq 2\ln(1/\epsilon)\}:w_0(x)+\sum_{\ell=1}^d \zeta_\ell w_\ell(x)\leq 0,$$

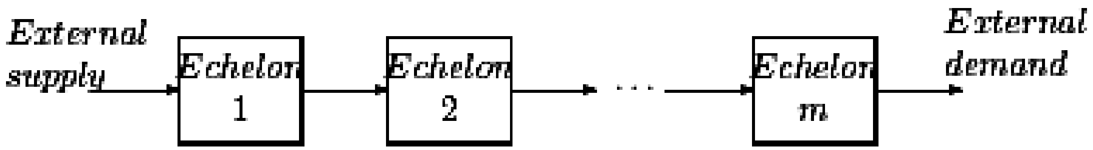
$$\phi_\ell(s)=\begin{cases}(1+s)\ln\left(\frac{1+s}{1+\mu_\ell^-}\right)+(1-s)\ln\left(\frac{1-s}{1-\mu_\ell^-}\right) & ,-1\leq s\leq\mu_\ell^-\\ 0 & ,\mu_\ell^-\leq s\leq\mu_\ell^+\\ (1+s)\ln\left(\frac{1+s}{1+\mu_\ell^+}\right)+(1-s)\ln\left(\frac{1-s}{1-\mu_\ell^+}\right) & ,\mu_\ell^+\leq s\leq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

Robust Optimization of Uncertainty Affected Linear Dynamic Problems

- **In dynamical decision-making, some of the decisions x_j should be made when the actual data becomes partially known and thus can depend on the corresponding portions of the data**

  Example: **In the Inventory problem with uncertain demand, replenishment orders v_t of day t usually can depend on the actual demands at days $1, ..., t-1$.**

The Supply chain Problem



- x_t^j=amount echelon j orders from j-1 at the beginning of period t
- Y_t^j =inventory level in echelon j at the end of period t
- z^j =initial inventory level at echelon j
- d_t =external demand at period t

- $I(j)$= Information delay, $M(j)$= manufacturing delay, $L(j)$= Lead time
- $T^L(j) = I(j) + M(j\text{-}1) + L(j)$ the delay between the time an order is placed and received in echelon j
- $T^M(j) = I(j+1) + M(j)$ the delay between the time an order is placed and shipped from echelon j

The Supply chain Problem

- Main objective : minimizing cost
- Sub objective:
  stabilizing the system
- Problem Characteristics:
  - Finite horizon
  - Multi echelon
  - Delays
  - Backlogging
  - Demand must be satisfied and is uncertain
- Eliminating the equalities recursively yields a LP with only inequalities

$$
\begin{array}{l}
\min_{x,y} \sum_{j,t} \left[c_t^j x_t^j + w_t^j \right] \\
s.t. \\
\quad y_t^j = y_{t-1}^j + x_{t-T^L(j)}^j - x_{t-T^M(j)}^{j+1} \qquad \forall j \in \{1,..,m-1\} \\
\quad y_t^m = y_{t-1}^m + x_{t-T^L(j)}^m - d_{t-T^M(m)} \\
\quad \left. \begin{array}{l}
w_t^j \geq h_t^j y_t^j \\
w_t^j \geq -p_t^j y_t^j \\
y_t^j \geq \underline{a}^j \\
y_t^j \leq \overline{a}^j \\
x_t^j \leq b^j \\
x_t^j \geq 0 \\
w_t^j \geq 0 \\
y_0^j = z^j
\end{array} \right\} \forall j \in \{1,..,m\}
\end{array}
$$

open loop dynamics: $$\begin{cases} x_{t+1} &= A_t x_t + B_t u_t + R_t d_t \\ y_t &= C_t x_t \\ x_0 &= z \end{cases}$$

control law: $$u_t = \xi_{t0} + \sum_{\tau=0}^{t} \Xi_{t\tau} y_\tau$$

$$\Downarrow$$

$$\boxed{w := (u_0, ..., u_T, x_0, ..., x_{T+1}) = W(\xi; d, z)}$$

$$\Downarrow$$

$$\boxed{\min_\xi \{f(W(\xi; d, z)) : D_i W(\xi; d, z) - b_i \in \mathcal{Q}_i,\ i = 1, ..., m\} \qquad \textbf{(U)}}$$

Note: **Due to presence of uncertain input trajectory d and possible uncertainty in the initial state, (U) is an uncertain problem.**

Difficulty: **While linearity of the dynamics and the control law make $W(\xi; d, z)$ linear in (d, z), the dependence of $W(\cdot, \cdot)$ on the parameters $\xi = \{\xi_{t0}, \Xi_{t\tau}\}_{0 \leq \tau \leq t \leq T}$ of the control law is highly nonlinear**

$\Rightarrow$ (U) is *not* a bi-affine problem, which makes inapplicable the theory we have developed. In fact, (U) seems to be intractable already when there is no uncertainty in d, z!

Dynamic Control Problems

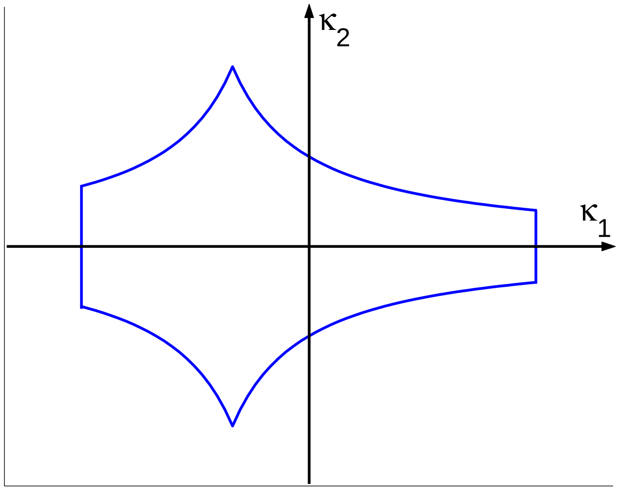
- Example:

$$\mathcal{X}_{t+1} = \mathcal{X}_t + \boldsymbol{u}_t + \boldsymbol{d}_t$$
$$\mathcal{Y}_t = \mathcal{X}_t$$
$$\mathcal{X}_0 = \mathbf{0}$$
$$\boldsymbol{u}_t = \kappa_t \mathcal{Y}_t$$

- Control wanted: when $|d_t| \leq 1$ then $|u_t| \leq 3$ for t=0,1,2
- Since: $\boldsymbol{u}_2 = \kappa_2 \mathcal{X}_2 = \kappa_2(\mathcal{X}_1 + \kappa_1 \boldsymbol{d}_0 + \boldsymbol{d}_1) = \kappa_2\left[(1+\kappa_1)\boldsymbol{d}_0 + \boldsymbol{d}_1\right]$

  is not bi-affine

  the control coefficients have a highly non convex domain.



Remedy: suitable re-parameterization of affine control laws.

♣ Consider a closed loop system along with its *model*:

closed loop system:	**model:**
$x_{t+1} = A_t x_t + B_t u_t + R_t d_t$	$\widehat{x}_{t+1} = A_t \widehat{x}_t + B_t u_t$
$y_t = C_t x_t$	$\widehat{y}_t = C_t \widehat{x}_t$
$x_0 = z$	$\widehat{x}_0 = 0$
$u_t = U_t(y_0, \ldots, y_t)$	

♠ **Observation:** We can run the model in an on-line fashion, so that at time t, before the decision on u_t should be made, we have in our disposal *purified outputs*

$$v_t = y_t - \widehat{y}_t.$$

♠ **Fact I [Equivalence]:** Every transformation $(d, z) \mapsto w$ which can be obtained from an affine control law based on outputs:

$$u_t = \xi_{t0} + \sum_{\tau=0}^{t} \Xi_{t\tau} y_\tau \qquad (*)$$

can be obtained from an affine control law based on purified outputs:

$$u_t = \eta_{t0} + \sum_{\tau=0}^{t} H_{t\tau} v_\tau \qquad (**)$$

and vice versa.

system:

$$\begin{array}{rcl} x_{t+1} &=& A_t x_t + B_t u_t + R_t d_t \\ y_t &=& C_t x_t \\ x_0 &=& z \end{array}$$

model:

$$\begin{array}{rcl} \widehat{x}_{t+1} &=& A_t \widehat{x}_t + B_t u_t \\ \widehat{y}_t &=& C_t \widehat{x}_t \\ \widehat{x}_0 &=& 0 \end{array}$$

control law:

$$\begin{array}{rcl} v_t &=& y_t - \widehat{y}_t \\ u_t &=& \eta_{t0} + \sum\limits_{\tau=0}^{t} H_{t\tau} v_\tau \qquad (**) \end{array} \qquad \textbf{(S)}$$

♠ **Fact II [bi-affinity]: The state-control trajectory $w = W(\eta; d, z)$ of (S) is affine in (d, z) when the parameters $\eta = \{\eta_{t0}, H_{t\tau}\}_{0\leq\tau\leq t\leq T}$ of the control law $(**)$ are fixed,** *and is affine in η when (d, z) is fixed.*

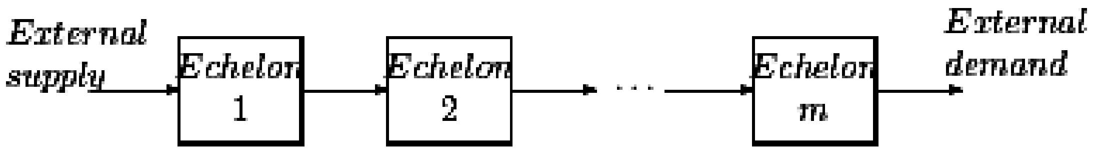
♠ **Corollary:** *With parameterization $(**)$ of affine control laws, the problem*

> *Find an affine control law $(*)$ which ensures that the resulting state-control trajectory w satisfies the system of convex inclusions*
>
> $$D_i w - b_i \in \mathcal{Q}_i,\ \ i = 1, ..., m$$
>
> *and minimizes, under this restriction, a given linear objective $f(w)$.*

becomes an uncertain bi-affine optimization problem and as such can be processed via the CRC approach.

In particular, in the case when $\mathcal{Q}_i$ are one-dimensional, the CRC of the problem is computationally tractable, provided that the normal range $\mathcal{U}$ of (d, z) and the associated cone $\mathcal{L}$ are so. If $\mathcal{U}$, $\mathcal{L}$ and the norms used to measure distances are polyhedral, CRC is just an explicit LP program.

The Supply chain Problem



- x_t^j=amount echelon j orders from j-1 at the beginning of period t
- Y_t^j =inventory level in echelon *j* at the end of period *t*
- z^j =initial inventory level at echelon *j*
- d_t =external demand at period *t*

- *I(j)*= Information delay, *M(j)*= manufacturing delay, *L(j)*= Lead time
- $T^L(j) = I(j) + M(j\text{-}1) + L(j)$ the delay between the time an order is placed and received in echelon *j*
- $T^M(j) = I(j+1) + M(j)$ the delay between the time an order is placed and shipped from echelon j

The Supply chain Problem

- Main objective : minimizing cost
- Sub objective:
  stabilizing the system
- Problem Characteristics:
  - Finite horizon
  - Multi echelon
  - Delays
  - Backlogging
  - Demand must be satisfied and is uncertain
- Eliminating the equalities recursively yields a LP with only inequalities

$$\min_{x,y} \sum_{j,t} \left[c_t^j x_t^j + w_t^j \right]$$

$$s.t.$$

$$y_t^j = y_{t-1}^j + x_{t-T^L(j)}^j - x_{t-T^M(j)}^{j+1} \qquad \forall j \in \{1,..,m-1\}$$

$$y_t^m = y_{t-1}^m + x_{t-T^L(j)}^m - d_{t-T^M(m)}$$

$$\left.\begin{array}{l} w_t^j \geq h_t^j y_t^j \\ w_t^j \geq -p_t^j y_t^j \\ y_t^j \geq \underline{a}^j \\ y_t^j \leq \overline{a}^j \\ x_t^j \leq b^j \\ x_t^j \geq 0 \\ w_t^j \geq 0 \\ y_0^j = z^j \end{array}\right\} \forall j \in \{1,..,m\}$$

The Bullwhip effect

- This problem has a well known phenomenon associated with it called “Bullwhip effect”
- The Bullwhip effect is described as “amplification of oscillation from down stream demands to upstream echelons”
- Such amplification can occur both in orders and inventory levels.
- Large variations in these measures are disruptive to the system and generates high cost.
- One of the aims of good control is to reduce the Bullwhip effect.

LV control

- Let's assume we take the control suggested by Love [Love,1979] using target inventory:

$$\boxed{x_t^j = x_{t-1}^{j+1} + {}^{1}\!/_{2}\left(\Gamma^j - y_{t-1}^j\right)}$$

- Resulting in the LP problem :

$$
\boxed{
\begin{array}{l}
\min\limits_{x,\tilde{x},y,w,\pi} \sum\limits_{j,t} \left[c_t^j x_t^j + w_t^j\right] \\
s.t. \\
\quad y_t^j = y_{t-1}^j + x_{t-T^L(j)}^j - x_{t-T^M(j)}^{j+1} \qquad \forall j \in \{1,..,m-1\} \\
\quad y_t^m = y_{t-1}^m + x_{t-T^L(j)}^m - d_{t-T^M(m)} \\
\quad \left.\begin{array}{l}
w_t^j \ge h_t^j y_t^j \\
w_t^j \ge -p_t^j y_t^j \\
\tilde{x}_t^j = x_{t-1}^{j+1} + {}^{1}\!/_{2}\left(\Gamma^j - y_{t-1}^j\right) \\
\tilde{x}_t^j \le \mathcal{M}\pi_t^j \\
x_t^j \ge 0 \\
x_t^j \ge \tilde{x}_t^j \\
x_t^j \le \mathcal{M}\pi_t^j \\
x_t^j \le \mathcal{M}(1-\pi_t^j) + \tilde{x}_t^j \\
w_t^j \ge 0 \\
y_0^j = z^j \\
\pi_t^j \in \{0,1\}
\end{array}\right\} \forall j \in \{1,..,m\}
\end{array}
}
$$

- We can further improve this method by making the reference inventory a decision variable rather than a constant. – **ILV control**

Example

- [Love, 1979], Oscillating demand:

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
d_t	6	6	6	6	6	6	6	6	7	8	9	10	9	8	7	6	5	4	5	6

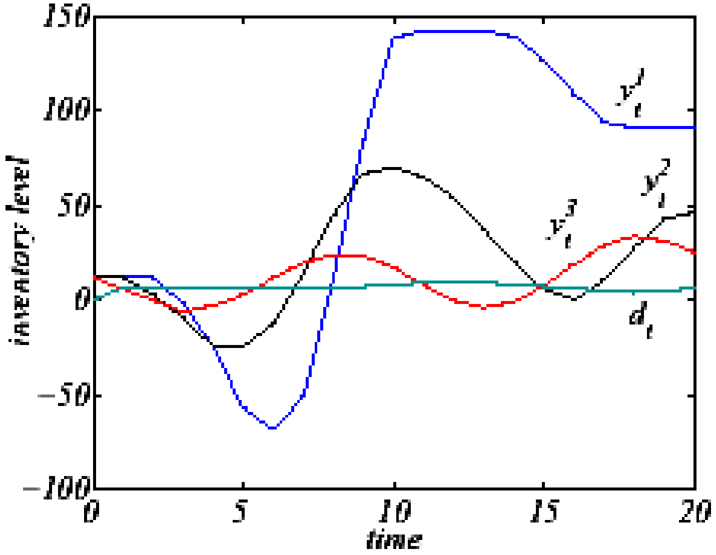
- Horizon: n=20
- Echelons: m=3
- Cost: c=2, p=3,h=1
- Initial inventory: z=12
- Lead time: L=2
- No other delays

The importance of good control

- The Bullwhip effect –

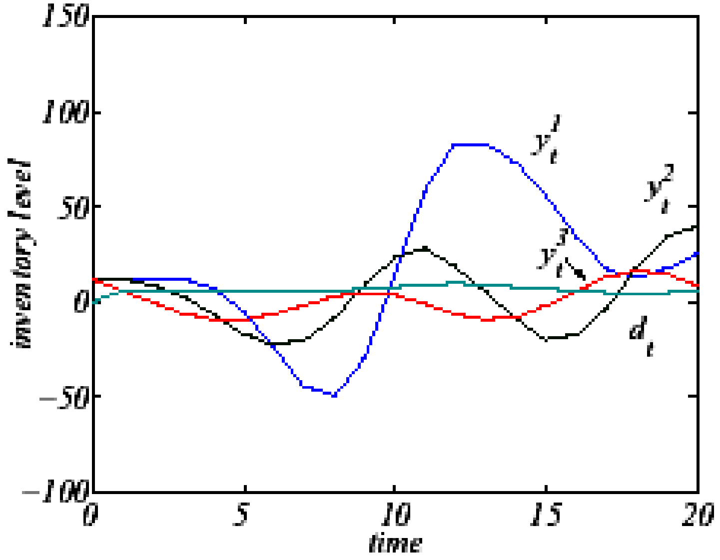
LV

OFV: 4795



ILV

OFV: 3140



- These are deterministic methods which do not work well with varying demand → a more robust method is needed

Purified output-based AARC control

The supply chain problem as a control problem

- The dynamics of the supply chain problem is given by:

$$y_t^j = y_{t-1}^j + x_{t-T^L(j)}^j - x_{t-T^M(j)}^{j+1} \quad \forall j \in \{1,..,m-1\}$$

$$y_t^m = y_{t-1}^m + x_{t-T^L(j)}^m - d_{t-T^M(m)}$$

with initial state: $y_0^j = z^j$

whose form matches the classical dynamic control problem

The purified outputs corresponding to the dynamic system (1) – (3) are here

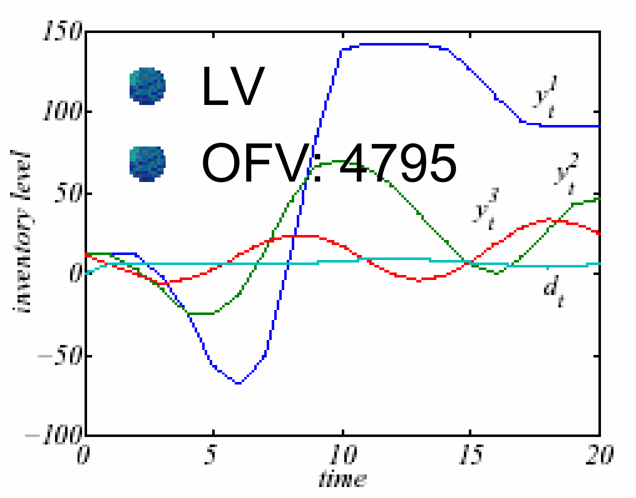
$$v_t^j = \begin{cases} z_0^m - \sum_{t=1}^{t-T^M(m)} d_z & \text{if } j = m \\ \\ z_0^j & j < m \end{cases}$$

The affine control law is here

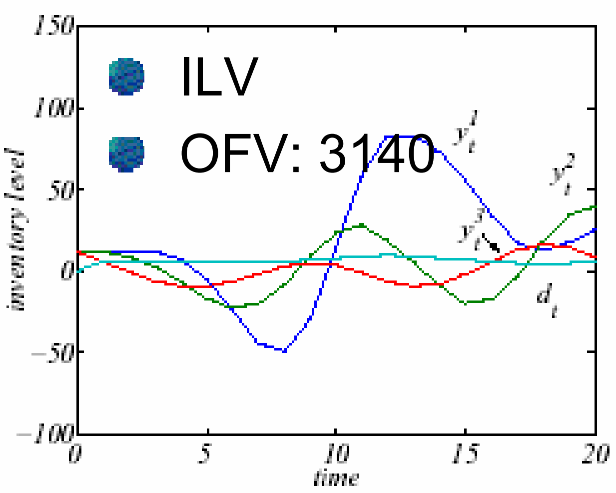
$$x_t^j = \eta_0^{x,t,j} + \sum_{l=1}^{m} \sum_{\tau=1}^{T} \eta_{\tau l}^{x,t,j} \, v_\tau^l$$

where $\eta_{\tau l}^{x,t,j} = 0 \qquad \forall \tau \geq t$ (non anticipativity)

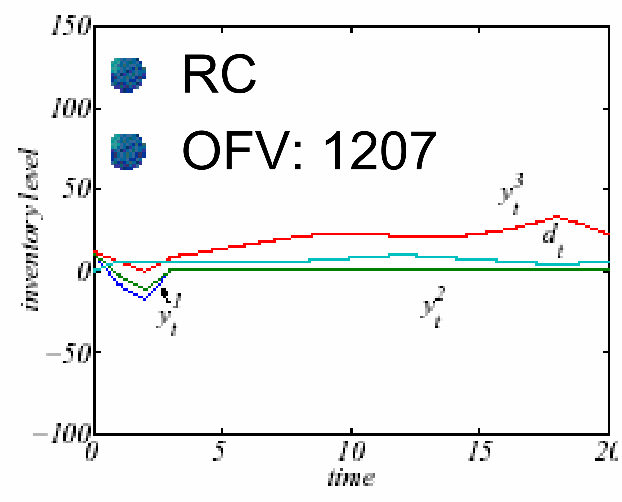
Inventory Behavior – “amplification of oscillation”



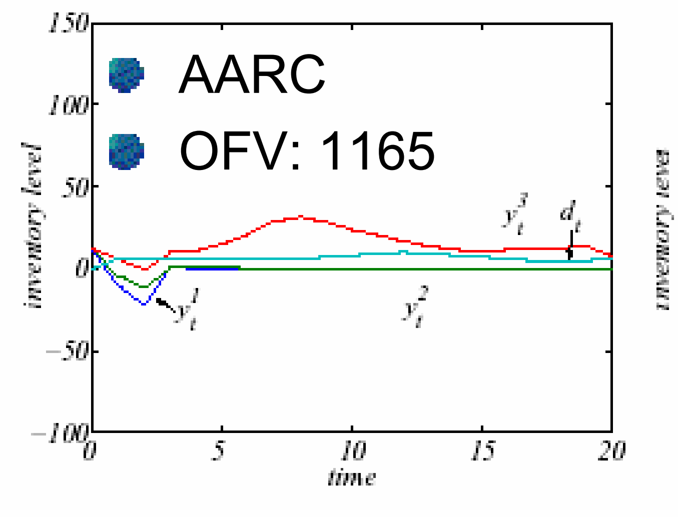
- LV
- OFV: 4795



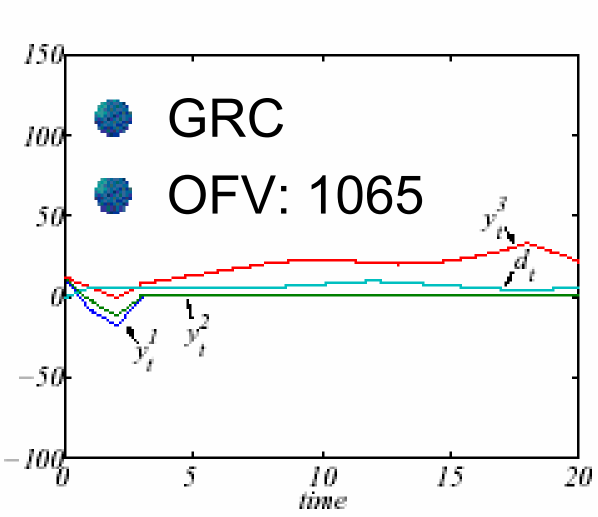
- ILV
- OFV: 3140



- RC
- OFV: 1207



- AARC
- OFV: 1165

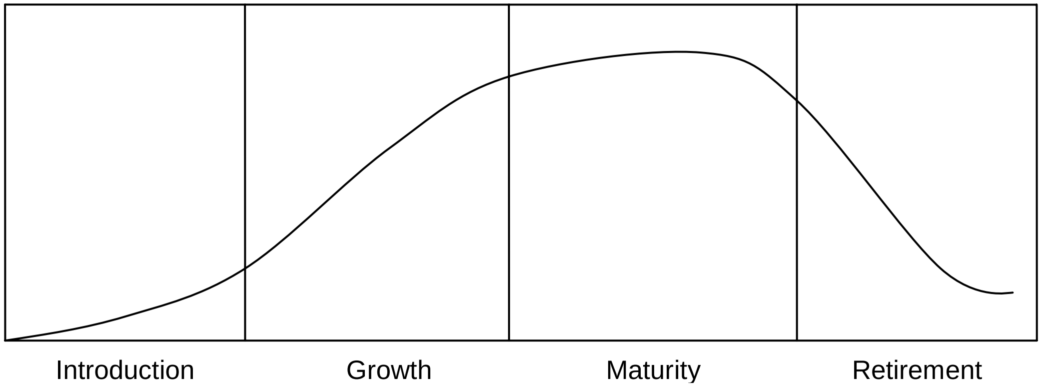


- GRC
- OFV: 1065

The Price(?) of Robustness

Problem definition

- Single-product finite horizon T production planning problem with discrete periods and uncertain demand
- Expected (nominal) demand follows a typical life-cycle pattern:



Problem definition

- In each period $t = 1, \ldots, T$:
  - holding (shortage) costs incur for each unit surplus (shortage)
  - Income is realized through sales
- In period T:
  - salvage value incur for surplus units
  - Unsatisfied demand is lost
- In each period $t = 1, \ldots, T-1$:
  - Unsatisfied demand is backlogged

Mathematical model (1)

T	planning horizon		
q_t	production quantity in units	c_t	production cost per unit
d_t	demand in units	m_t	selling price per unit
$\bar{d}_t$	nominal demand in units	s_T	salvage value per unit
I_t	inventory level	h_t	holding cost per unit
k	initial inventory	p_t	shortage cost per unit

$$q^t = (q_1, q_2, \dots, q_t)^T$$

$$d^t = (d_1, d_2, \dots, d_t)^T$$

- denote $I_t = I_t\left(q^t, d^t\right) \triangleq \sum_{\tau=1}^{t} \left(q_\tau - d_\tau\right) + k$

- The model:

Sales Salvage

$$\underset{q \geq 0}{Max} \left\{ \begin{array}{l} F(q,d) = \sum_{t=1}^{T} \left[m_t \min\left(d_t, I_{t-1}\left(q^{t-1}, d^{t-1}\right) + q_t\right) \right] + s_T \max\left(I_T\left(q^T, d^T\right), 0\right) - \\ -\sum_{t=1}^{T} \left[c_t q_t + h_t \max\left(I_t\left(q^t, d^t\right), 0\right) + p_t \max\left(-I_t\left(q^t, d^t\right), 0\right) \right] \end{array} \right\} \quad (1)$$

Production Holding Shortage

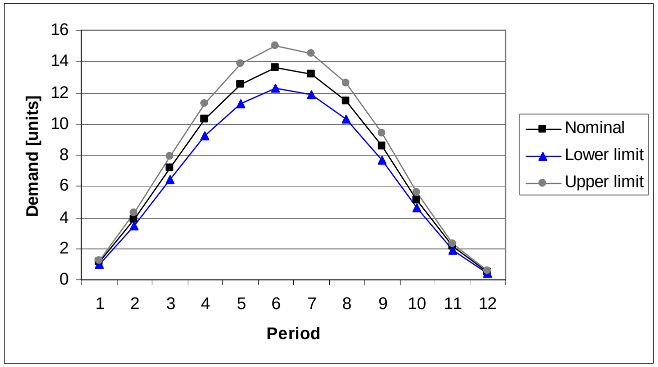
LP model (2)

- The piecewise linear model (1) can be written as the following LP model:

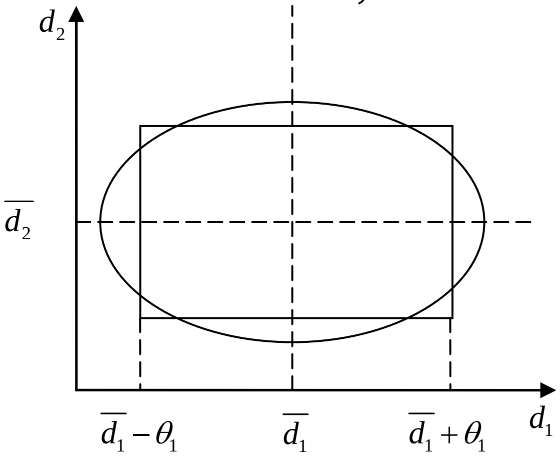
$$
\begin{aligned}
&\underset{q,I,y,g}{Max}\ F \\
&s.t. \\
&\qquad \left.\begin{aligned}
&\sum_{t=1}^{T}\left[g_t - c_t q_t - y_t\right] \geq F \\
&y_t \geq \overline{h}_t I_t\left(q^t, d^t\right) \\
&y_t \geq -p_t I_t\left(q^t, d^t\right) \\
&g_t \leq m_t d_t \\
&g_t \leq m_t\left(I_{t-1}\left(q^{t-1}, d^{t-1}\right) + q_t\right) \\
&q_t, g_t, y_t \geq 0
\end{aligned}\right\} \quad t = 1, \ldots, T \qquad (2)
\end{aligned}
$$

Uncertainty set

- $U_{box}=\left\{d^T \mid \; \left|d_t-\overline{d_t}\right|\le\theta_t,\quad \forall t=1,\dots,T\right\}$ where θ_t is the uncertainty level



- $U_{ell}=\left\{d\in\mathbb{R}^T \mid \; d_t=\overline{d_t}+\theta_t\zeta_t\quad t=1,\ldots,T\quad \text{where } \|\zeta\|_2\le\rho\right\}$ where ρ is the ellipsoid radius



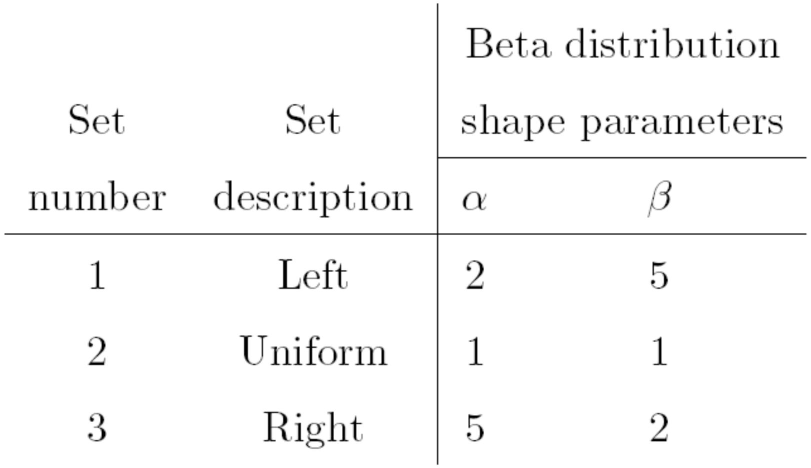
Simulations

- For several ellipsoid radius a set of L demand vectors was generated
- Each vector consisting of $T = 12$ entries
- The entries of the demand vectors were generated from a uniform distribution supported by the uncertainty set for the demand.
- denote the average "actual" profit of method A by $AP(q_A) = \frac{1}{L}\sum_{l=1}^{L} F(q_A, d^l)$ where $F(q,d)$ is the objective function (total profit) given in (1)

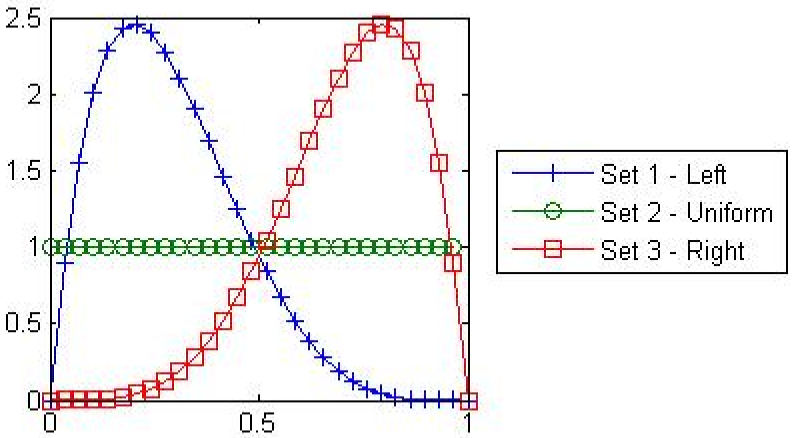
Simulations

- For several uncertainty levels three sets of L=100 demand vectors were generated
- Each vector consists of $T=12$ entries
- The demand vectors entries of each set were generated from a Beta distribution with specific

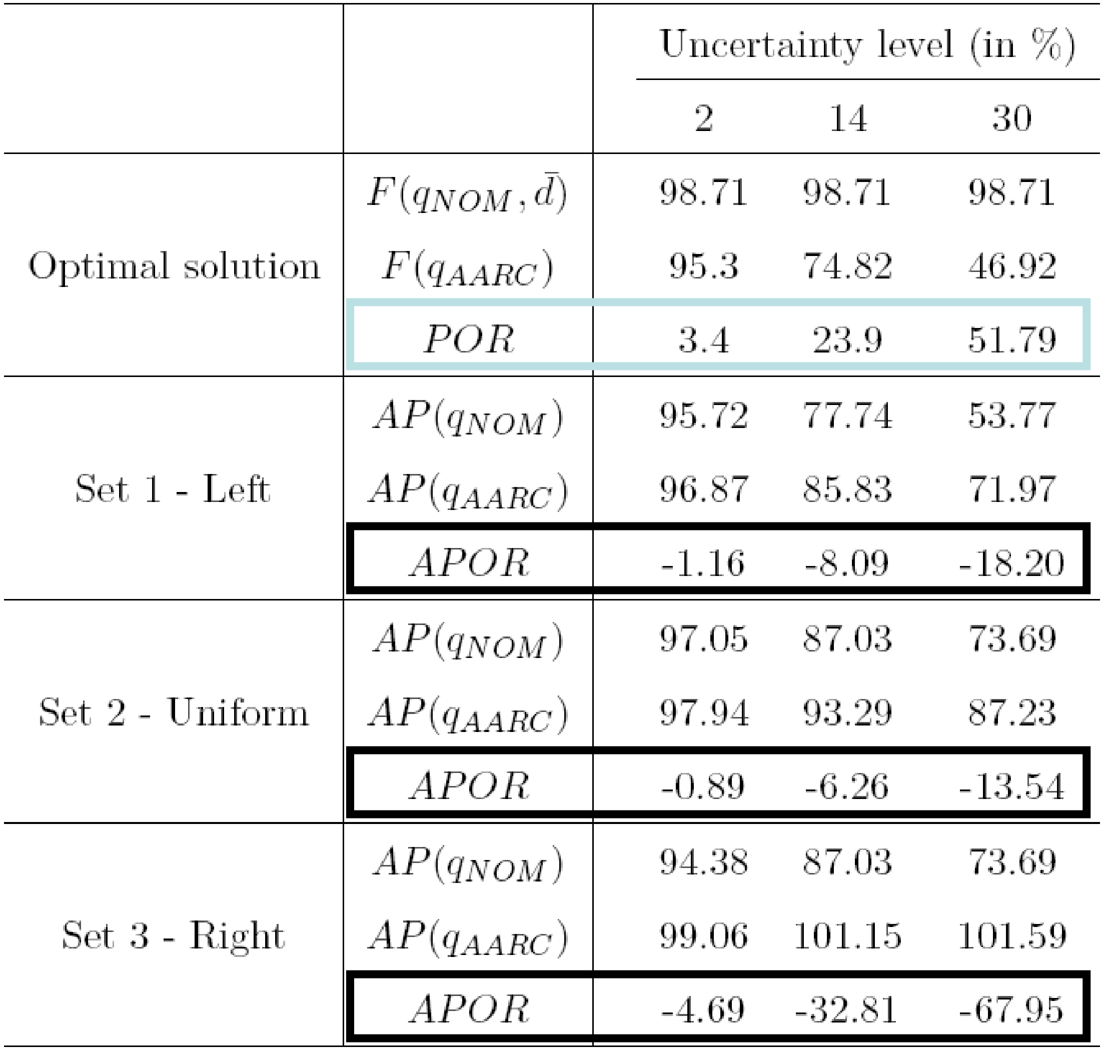
pported by the uncertainty



Set number	Set description	Beta distribution shape parameters α	β
1	Left	2	5
2	Uniform	1	1
3	Right	5	2



The actual price of robustness (APOR)



		Uncertainty level (in %)		
		2	14	30
Optimal solution	$F(q_{NOM}, \bar{d})$	98.71	98.71	98.71
	$F(q_{AARC})$	95.3	74.82	46.92
	POR	3.4	23.9	51.79
Set 1 - Left	$AP(q_{NOM})$	95.72	77.74	53.77
	$AP(q_{AARC})$	96.87	85.83	71.97
	$APOR$	-1.16	-8.09	-18.20
Set 2 - Uniform	$AP(q_{NOM})$	97.05	87.03	73.69
	$AP(q_{AARC})$	97.94	93.29	87.23
	$APOR$	-0.89	-6.26	-13.54
Set 3 - Right	$AP(q_{NOM})$	94.38	87.03	73.69
	$AP(q_{AARC})$	99.06	101.15	101.59
	$APOR$	-4.69	-32.81	-67.95

Simulations results

- Denote the average "actual" profit of solution method A by $AP(q_A) = \frac{1}{L}\sum_{l=1}^{L} F(q_A, d^l)$ where $F(q,d)$ is the objective function (total profit) given in (1)