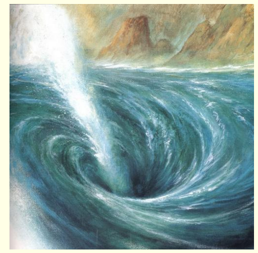
How to avoid choosing between Scylla and Charybdis

in Stochastic Programming

Claudia Sagastizábal

Eletrobras Cepel

mailto:sagastiz@impa.br

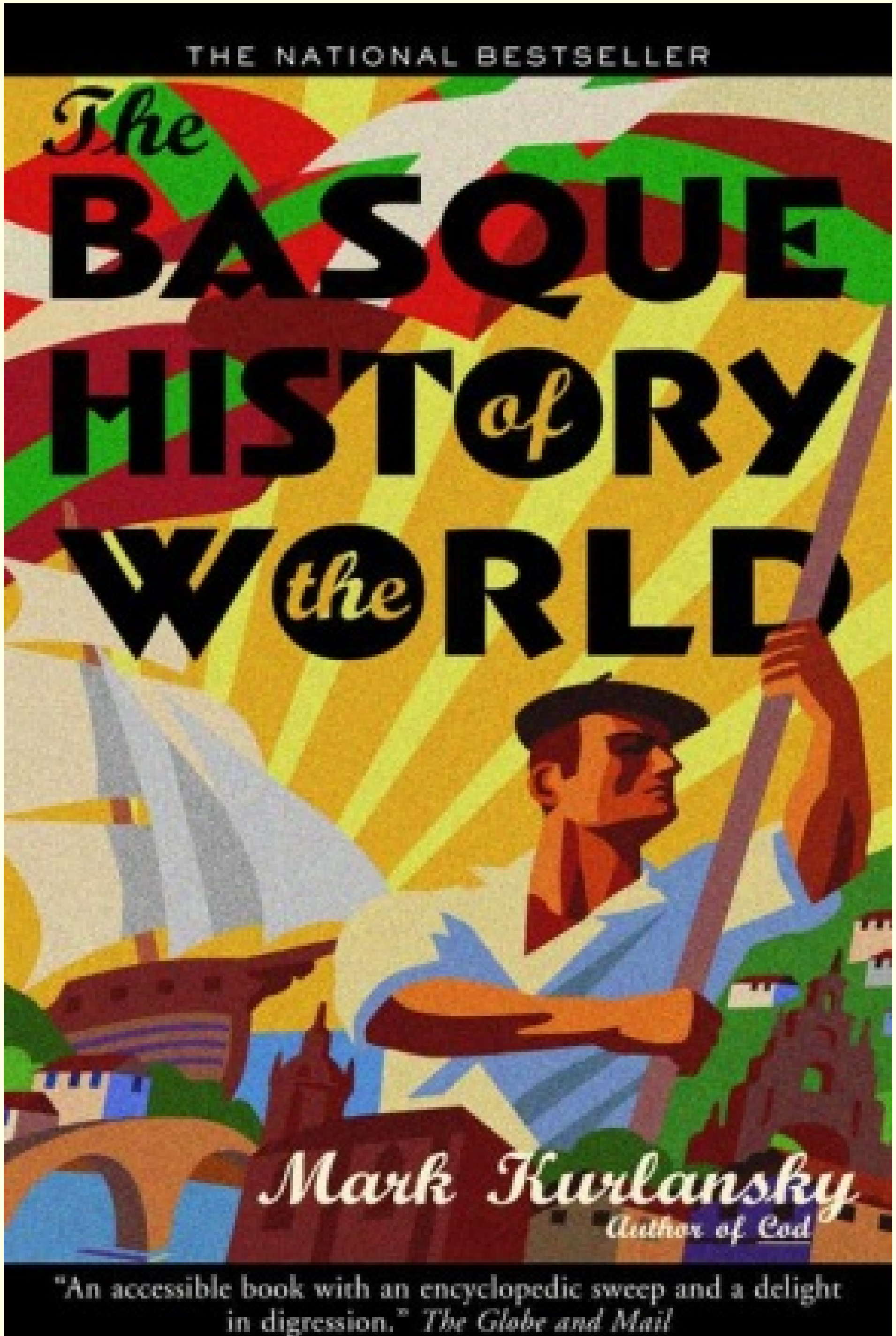
http://www.impa.br/~sagastiz

JBHU-Fest, Bayonne 2010

Before getting started

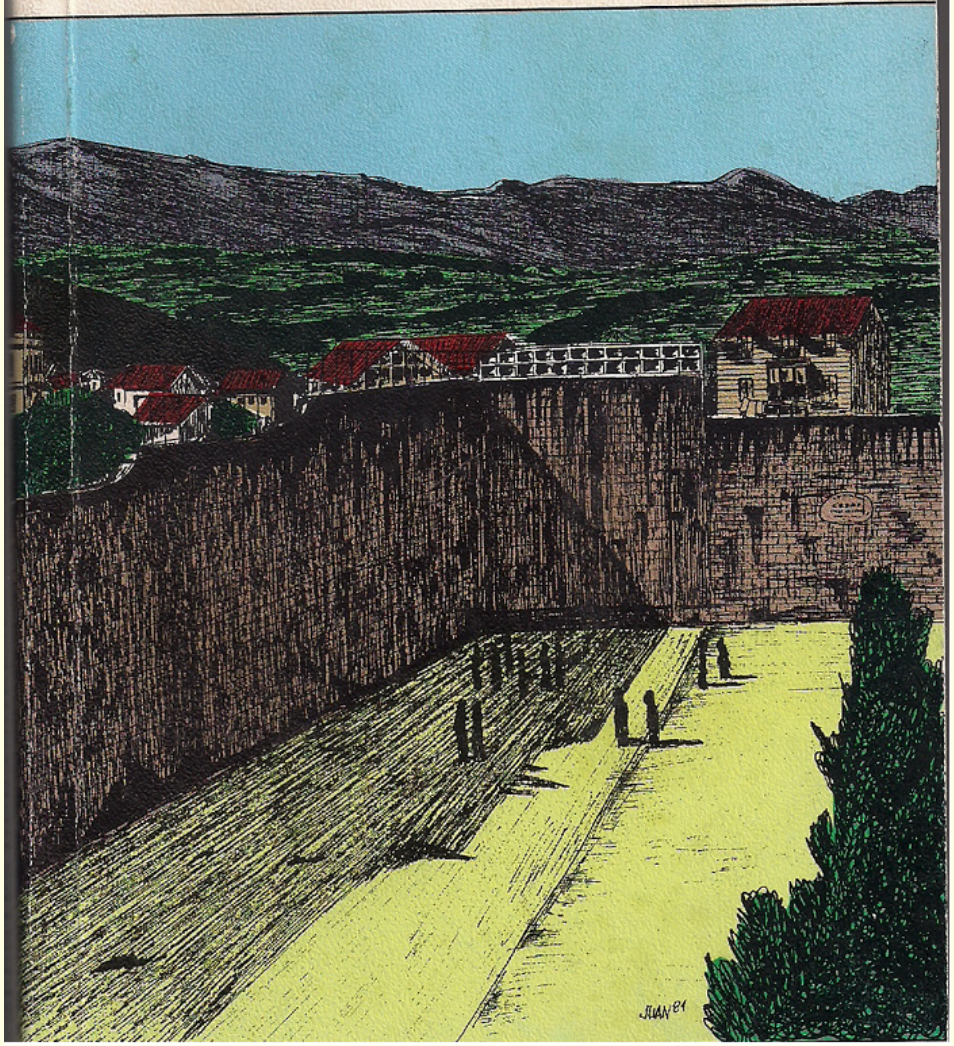
. . .

some Basque history


THE NATIONAL BESTSELLER
The BASQUE HISTORY of the WORLD
Mark Kurlansky
Author of Cod
"An accessible book with an encyclopedic sweep and a delight in digression." The Globe and Mail

Francisco Javier de Sagastizabal

APUNTES DURANGUESES SOBRE EL NUEVO JUEGO DE PELOTA

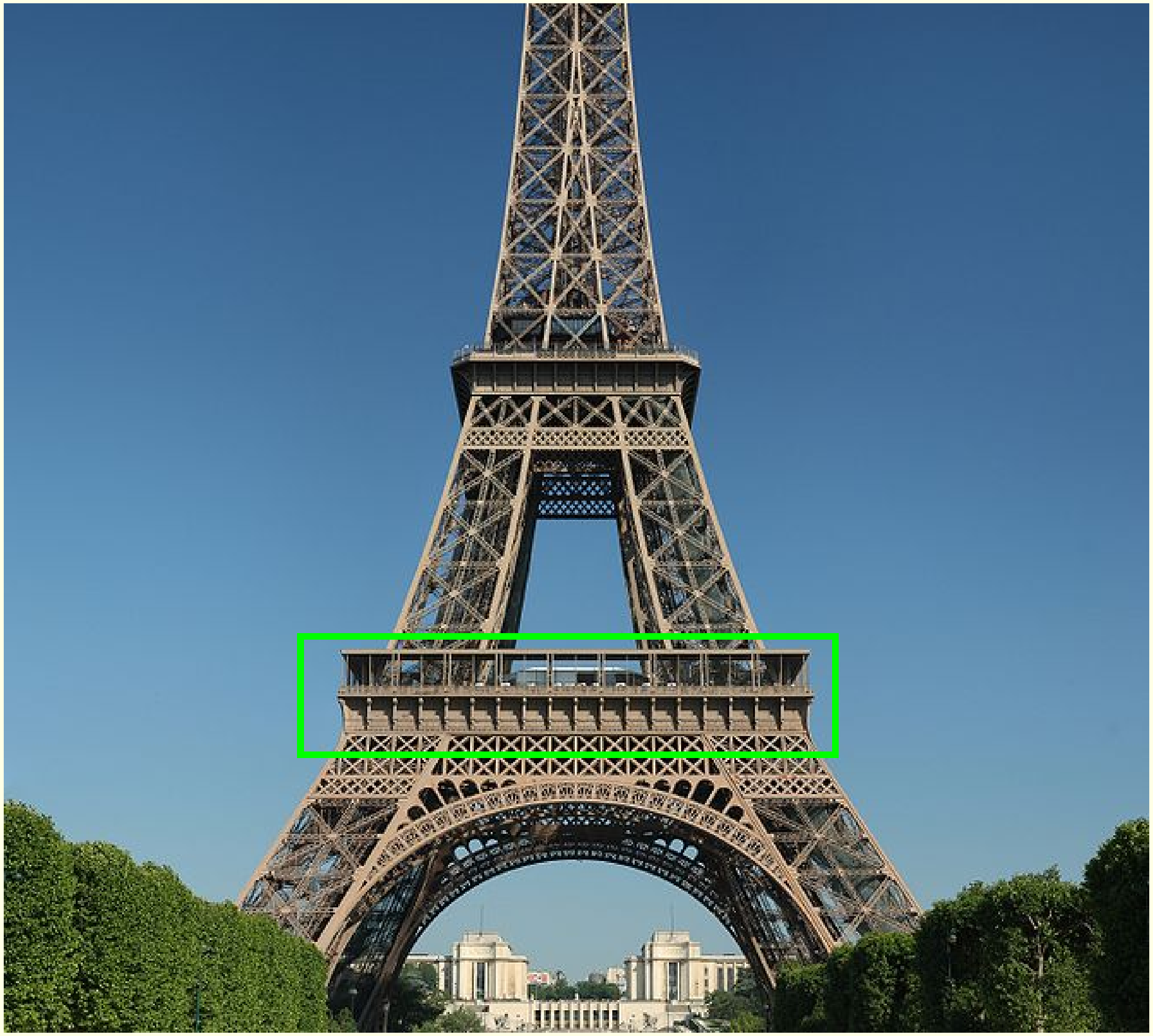


Now

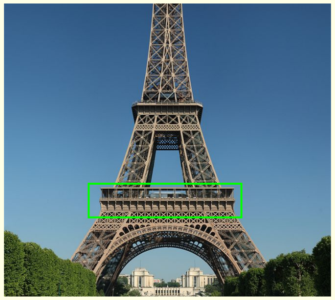
. . .

some French history

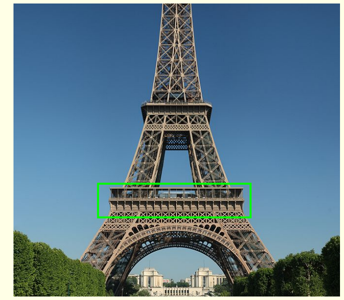
The 72 names on the Eiffel Tower



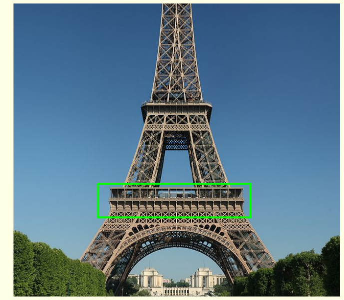
The 72 names on the Eiffel Tower



The 72 names on the Eiffel Tower



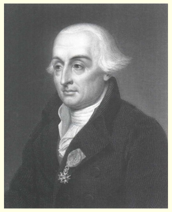
The 72 names on the Eiffel Tower



Lagrange, Laplace and Gauss

"The great masters of modern analysis are Lagrange, Laplace, and Gauss, who were contemporaries. It is interesting to note the marked contrast in their styles. Lagrange is perfect both in form and matter, he is careful to explain his procedure, and though his arguments are general they are easy to follow. Laplace on the other hand explains nothing, is indifferent to style, and, if satisfied that his results are correct, is content to leave them either with no proof or with a faulty one. Gauss is as exact and elegant as Lagrange, but even more difficult to follow than Laplace, for he removes every trace of the analysis by which he reached his results, and studies to give a proof which while rigorous shall be as concise and synthetical as possible." W.W.R. Ball, History of Mathematics (3rd Ed., 1901), p. 68

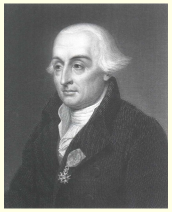
Joseph-Louis Lagrange (1736-1813)



I regard as quite useless the reading of large treatises of pure analysis: too large a number of methods pass at once before the eyes. It is in the works of application that one must study them; one judges their utility there and appraises the manner of making use of them.

Je considère comme complètement inutile la lecture de gros traités d'analyse pure: un trop grand nombre de méthodes passent en même temps devant les yeux. C'est dans les travaux d'application qu'on doit les étudier; c'est lá qu'on juge leurs capacités et qu'on apprend la manière de les utiliser.

Joseph-Louis Lagrange (1736-1813)



I regard as quite useless the reading of large treatises of pure analysis: too large a number of methods pass at once before the eyes. It is in the works of application that one must study them; one judges their utility there and appraises the manner of making use of them.

Before we take to sea we walk on land. Before we create we must understand.

Now

. . .

modern times

Back to business: general context

Real-life optimization problems

– complex in nature

modeling issues

– evolve in time

dynamic relations $\Rightarrow$ coupling decisions

– large scale

solvability issues

Trade-off between accuracy and speed

Back to business: general context

Real-life optimization problems

– complex in nature

modeling issues, **today's talk: high accuracy required**

– evolve in time

dynamic relations $\Rightarrow$ coupling decisions

– large scale

solvability issues

Trade-off between accuracy and speed

Our motivation: Energy problems

For a large hydrothermal power system

Optimal management of the resources given

– Limited availability of hydro-power

– High hydrological uncertainty

Optimal management of an asset of unknown value:

the water

How to price (the lack of) water?

The problem of pricing water

Decision today		Consequences tomorrow	
• Minimize immediate cost, by emptying reservoirs		It rains	Right decision
		Drought	Rationing (deficit)
or			
• Keep water, by thermal generation		It rains	Too much water
		Drought	Right decision

The problem of pricing water

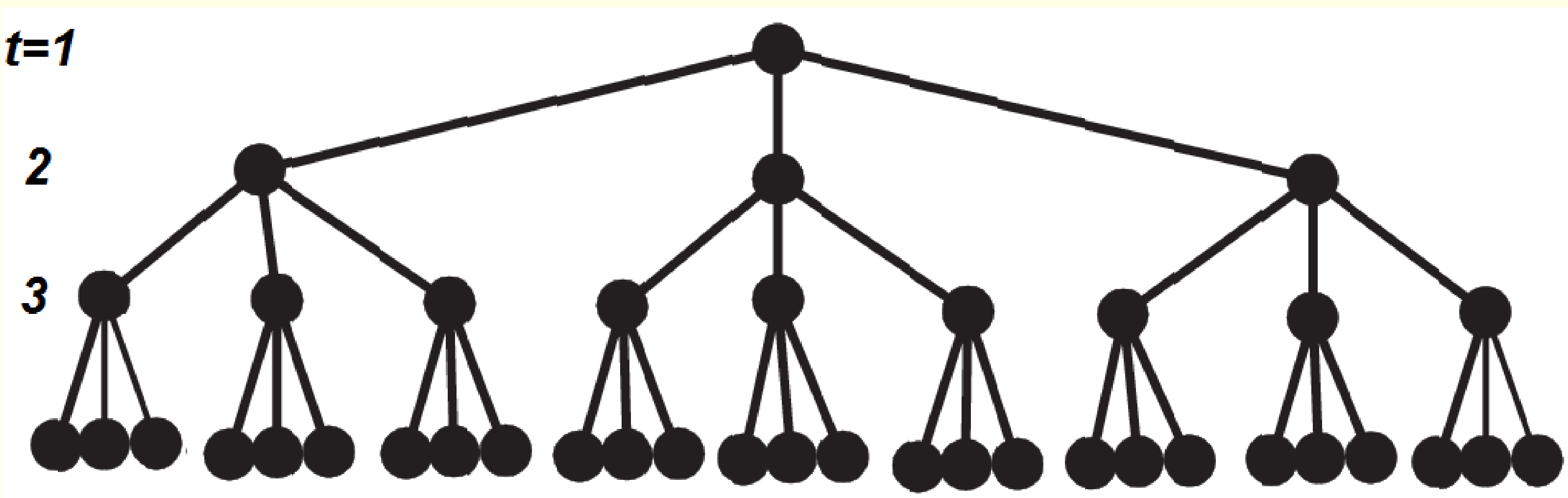
- Water price is given by the value function of a stochastic linear program
- Depends on reservoirs volumes
- Depends on how stochastic process is represented
- Depends on how stochastic program is solved

Used by electrical agents in Brazil (> 400) for financial strategies, Government actions, etc

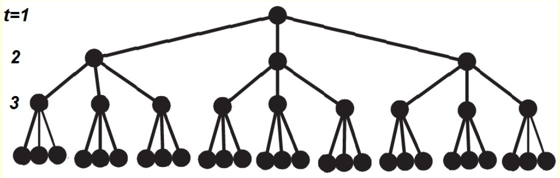
A typical problem: ($\mathrm{SLP_T}$)

$$
v(x_o) = \begin{cases} \min\limits_{(x_t,y_t)} & \sum_{t=1}^{T} c_t^\top y_t \\ x_t & = A x_{t-1} + B y_t + C_t \boldsymbol{\xi_t} + d_t \quad \text{(WB)} \\ E x_t & + F y_t \geq G_t \boldsymbol{\xi_t} + h_t \quad \text{(DEM)} \\ (x_t, y_t) & \in \mathcal{X}_t \times \mathcal{Y}_t \quad \text{(BND)} \end{cases}
$$

If 3 hydrological conditions: $\{normal, wet, dry\}$:



3^3

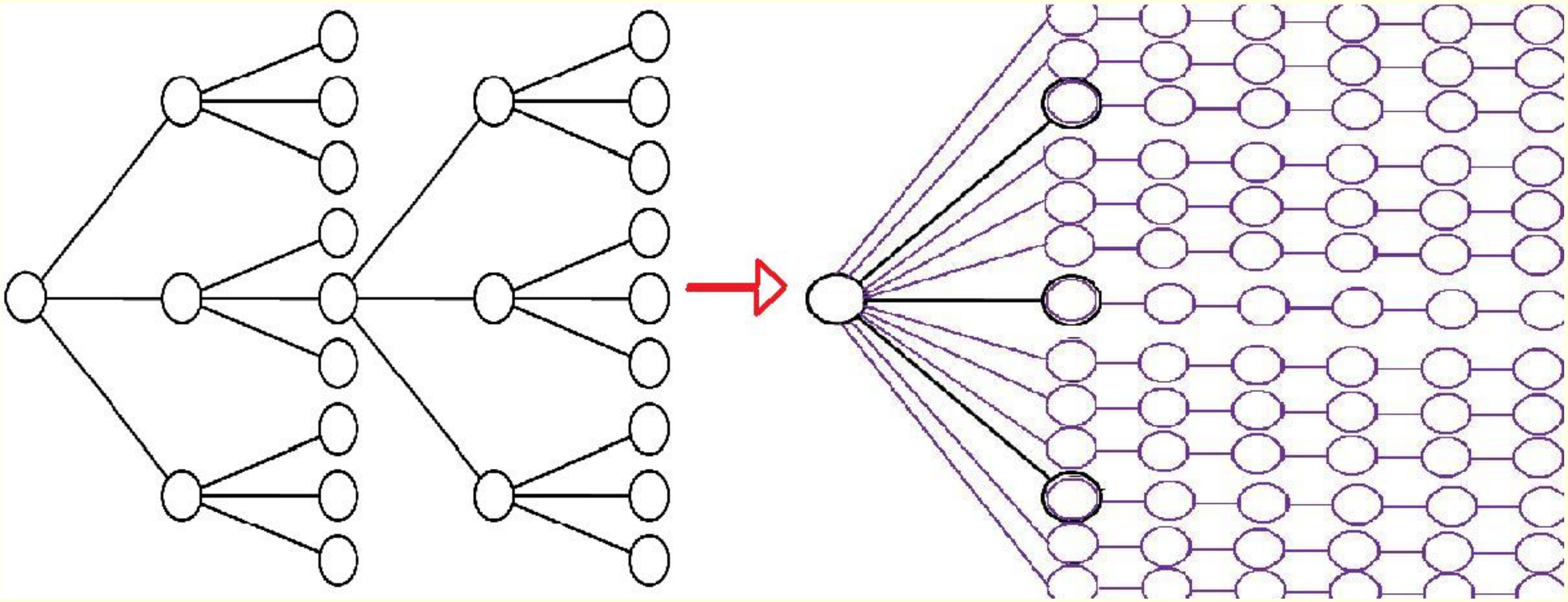
Brazil's **SLP$_\mathbf{T}$** sizes: 

– $T = 120$ time steps

– 20 hydrological conditions for each month and each subsystem (20 different $\xi_{\mathbf{t}}^{\mathbf{i}}$)

Tree has $\mathbf{20^{119}}$ scenarios in $\mathbb{R}^4$!

Proposal for 2-stage problems



$$\mathbf{SLP_2}\begin{cases}\displaystyle\min_{(x,y)\geq 0} c^\top x + q^\top y \\ Ax = b \\ Tx + Wy = h\end{cases}$$

Proposal for 2-stage problems

(joint work with W. de Oliveira, S. Scheinberg)

$$\mathbf{SLP_2}\left\{\begin{array}{l}\min\limits_{(x,y)\geq 0} c^\top x + q^\top y\\ Ax = b\\ Tx + Wy = h\end{array}\right.$$

Inexact bundle method

$$\approx$$

L-shaped method $+$ $\begin{array}{l}\text{regularization}\\ \text{inexactness}\end{array}$

Handling large $\mathbf{SLP_2}$

$$\left.\begin{array}{l} \min\limits_{(x,y)\geq 0} c^\top x + \mathbf{q}^\top y \\ Ax = b \\ \mathbf{T}x + Wy = \mathbf{h} \end{array}\right\} \equiv \left\{\begin{array}{l} \min\limits_{x\geq 0} c^\top x + \mathbb{E}[Q(x;\xi)] \\ Ax = b \\ \text{for } Q(x;\xi) = \left\{\begin{array}{l} \min\limits_{y\geq 0} q^\top y \\ Wy = h - Tx \end{array}\right. \end{array}\right.$$

Assumptions: fixed recourse W, uncertainty $\xi = (q, T, h)$ with finite variance, relatively complete recourse, nonempty $X \Longrightarrow$ the expected recourse function

$$\mathcal{Q}(x) = \mathbb{E}[Q(x;\xi)]$$

is well defined, proper, lsc, and convex.

Handling large SLP$_2$

$$\left.\begin{array}{l}\min\limits_{(x,y)\geq 0} c^\top x + q^\top y \\ Ax = b \\ Tx + Wy = h\end{array}\right\} \equiv \left\{\begin{array}{l}\min\limits_{\mathbf{x}\geq \mathbf{0}} \mathbf{c}^\top \mathbf{x} + \mathbb{E}[\mathbf{Q}(\mathbf{x};\xi)] \\ \mathbf{A}\mathbf{x} = \mathbf{b} \\ \text{for } Q(x;\xi) = \left\{\begin{array}{l}\min\limits_{y\geq 0} q^\top y \\ Wy = h - Tx\end{array}\right.\end{array}\right.$$

First-stage problem $\equiv \min$ **over a polyhedron**

the nonsmooth function $f(x) = c^\top x + \mathcal{Q}(x)$

The first-stage objective function $f(x) = c^\top x + \mathcal{Q}(x)$

SAA:

$$
\begin{aligned}
\mathcal{Q}(x) = \mathbb{E}[Q(x;\xi)] &= \int_\xi Q(x;\xi)\,dp(\xi) \\
&\approx \sum_{i=1}^N p_i Q(x;\xi_i) \\
&= \sum_{i=1}^N p_i Q_i(x) \qquad \text{for } \xi_i = (q_i, h_i, T_i) \\
&= \sum_{i=1}^N p_i \left\{ \begin{array}{l} \min\limits_{y \geq 0} q_i^\top y \\ Wy = h_i - T_i x \end{array} \right. \\
&= \sum_{i=1}^N p_i \left\{ \begin{array}{l} \max\limits_{u} u^\top (h_i - T_i x) \\ W^\top u \leq q_i \end{array} \right.
\end{aligned}
$$

A dual solution u_i satisfies $-T_i^\top u_i \in \partial Q_i(x)$

The first-stage objective function $f(x) = c^\top x + \mathcal{Q}(x)$

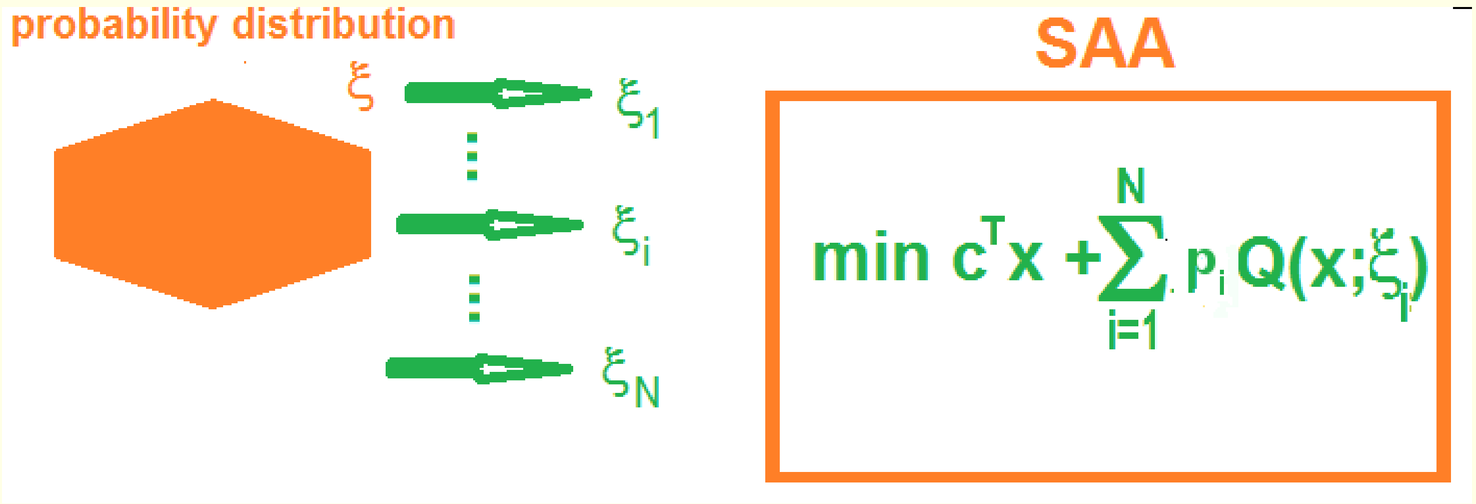
$$
\begin{aligned}
\mathcal{Q}(x) = \mathbb{E}[Q(x;\xi)] \quad &= \quad \int_\xi Q(x;\xi)dp(\xi) \\
&\approx \quad \sum_{i=1}^{N} p_i Q(x;\xi_i) \\
&= \quad \sum_{i=1}^{N} p_i Q_i(x) \\
&= \quad \sum_{i=1}^{N} p_i \begin{cases} \min\limits_{y \geq 0} q_i^\top y \\ Wy = h_i - T_i x \end{cases} \\
&= \quad \sum_{i=1}^{N} p_i \begin{cases} \max\limits_{u} u^\top (h_i - T_i x) \\ W^\top u \leq q_i \end{cases}
\end{aligned}
$$

SAA:

want N to be large

A dual solution u_i satisfies $-T_i^\top u_i \in \partial Q_i(x)$

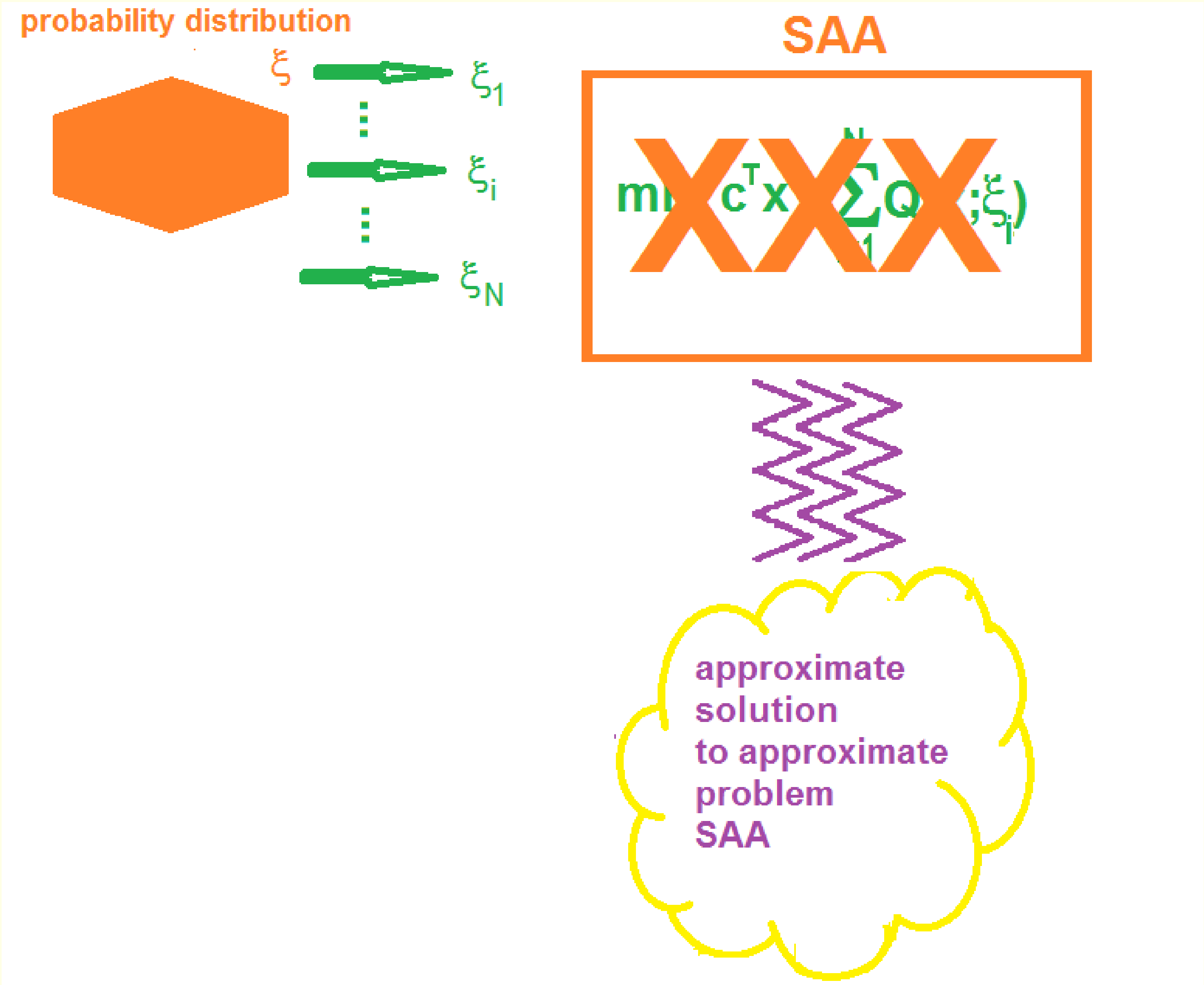
Consequences of large N



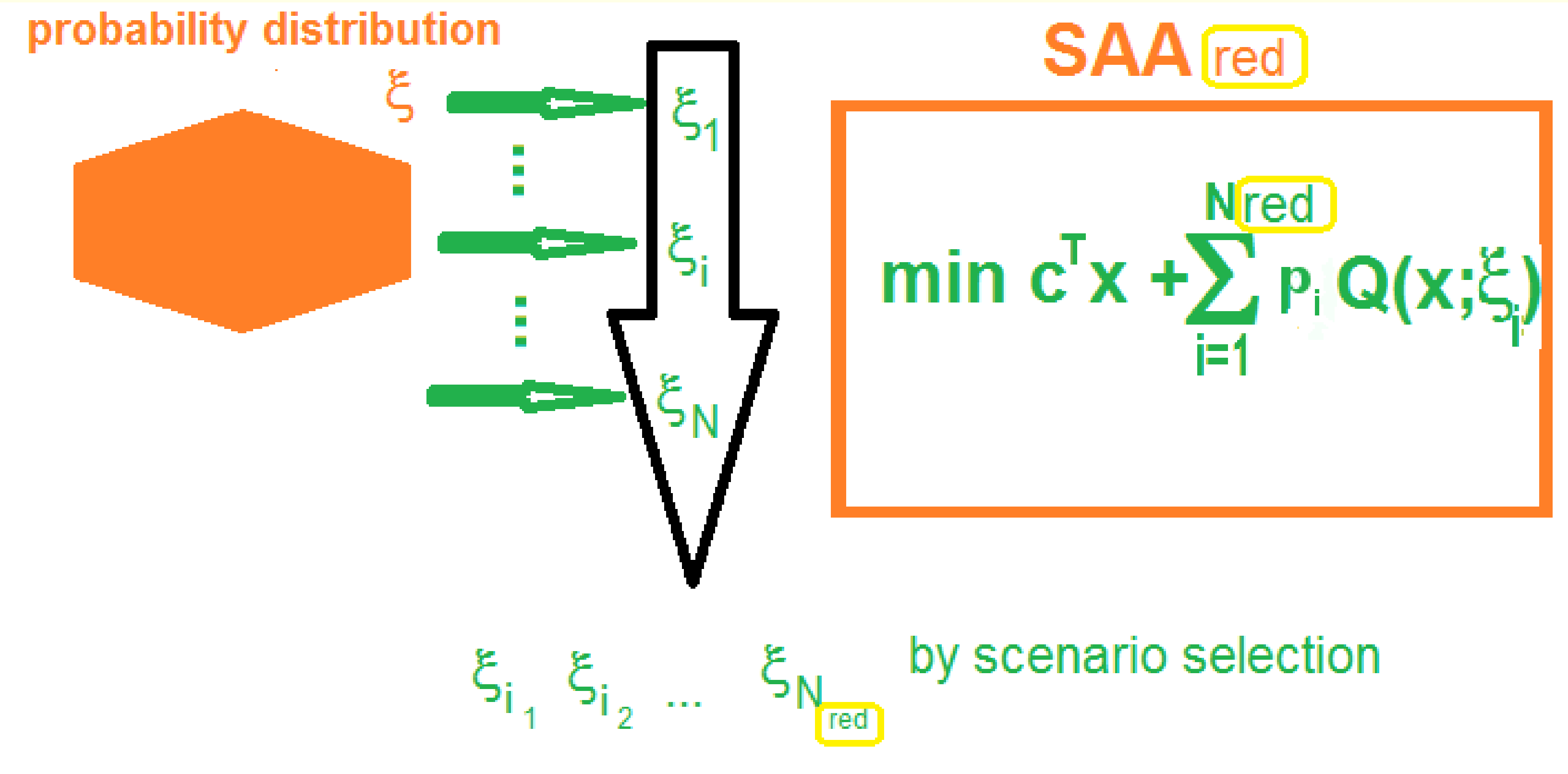
when N is large, LP is not solvable

(not even decomposing, as in L-Shaped$\equiv$ cutting-planes)

Consequences of large N



Static approach

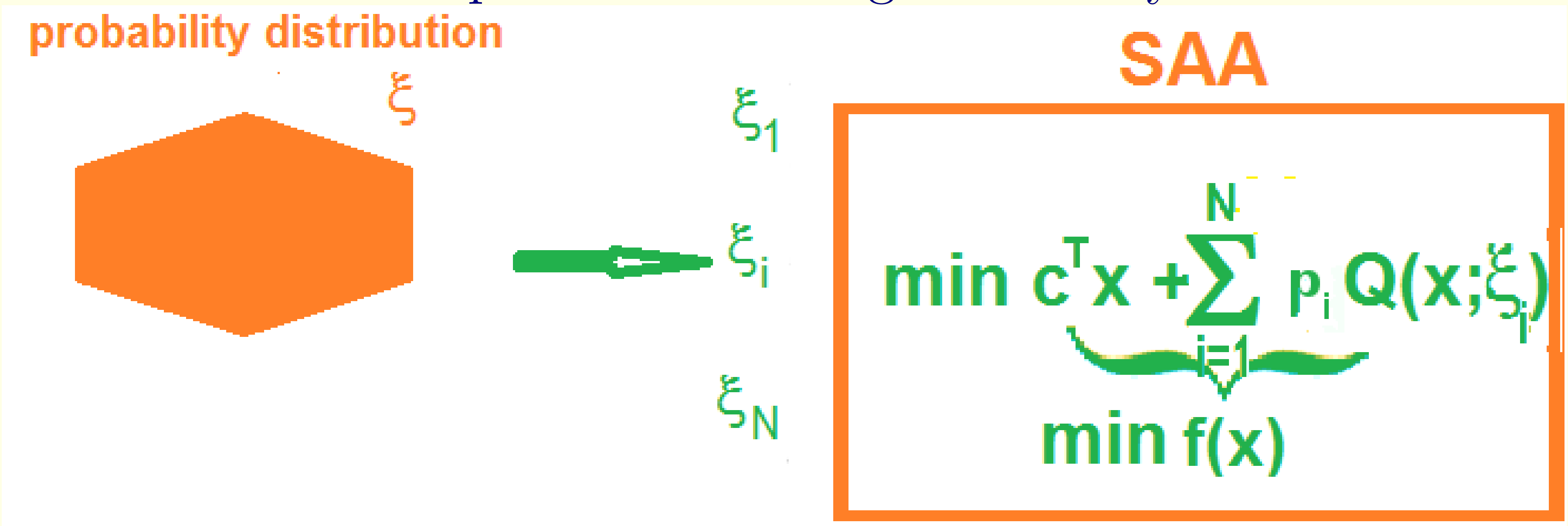


Use scenario selection methods [a] to define a **reduced** SAA problem, solvable by some decomposition method

[a] Eichhorn, Heitsch, Hochreiter, Küchler, Pflüg, Römisch, et al

Our approach

Recall that first-stage problem is nonsmooth and apply a NSO method capable of handling inaccuracy:



There is an oracle giving

$$\left[\begin{array}{lll} \text{a function estimate} & f_x \in & [f(x) - \epsilon_f, f(x) + \epsilon_g] \\ \text{a subgradient estimate} & g_x \in & \partial_{\epsilon_f + \epsilon_g} f(x)\,, \end{array}\right.$$

for errors $\epsilon_f, \epsilon_g \geq 0$ unknown, but bounded.

Our approach

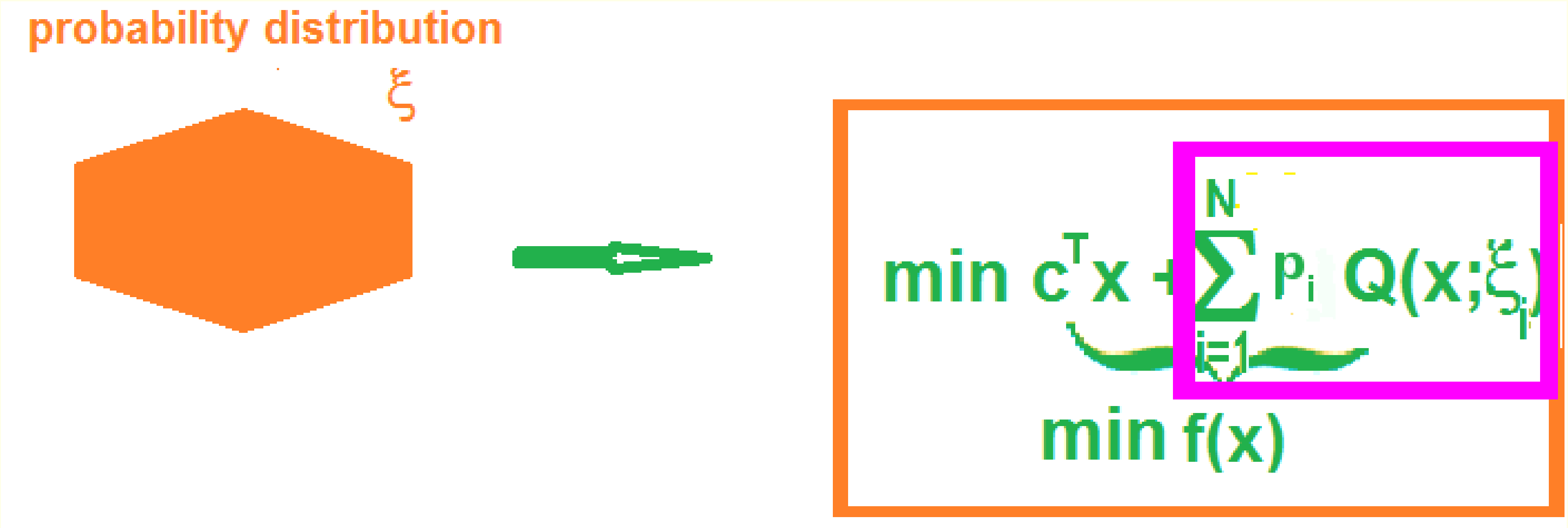
Recall that first-stage problem is nonsmooth and apply a NSO method capable of handling inaccuracy:

There is an oracle giving

- a function estimate $f_x \in [f(x) - \epsilon_f, f(x) + \epsilon_g]$
- $\mathbf{f_x}$ is the sum of $c^\top x$ and a few $\mathbf{Q_i(x) = Q(x;\xi_i)}$

Our approach is dynamic

Recall that first-stage problem is nonsmooth and apply a NSO method capable of handling inaccuracy:



the choice of $\{\xi_i\}_i$ composing f_x varies with x

- a function estimate $f_x \in [f(x) - \epsilon_f, f(x) + \epsilon_g]$
- $\mathbf{f_x}$ is the sum of $c^\top x$ and a few $\mathbf{Q_i}(\mathbf{x}) = \mathbf{Q}(\mathbf{x}; \xi_\mathbf{i})$

First inexact oracle: exploiting structure

$$Q_i(x) = \begin{cases} \min_{y \geq 0} q_i^\top y \\ Wy = h_i - T_i x \end{cases} = \begin{cases} \max_u u^\top (h_i - T_i x) \\ W^\top u \leq q_i \end{cases} = u_i^\top (h_i - T_i x)$$

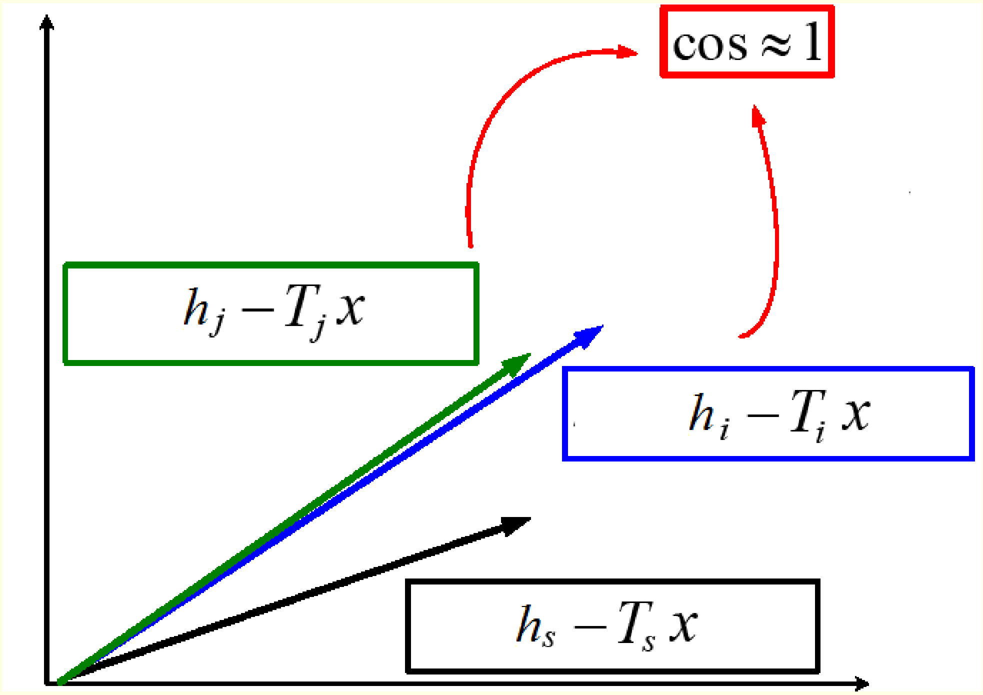
If second-stage cost is deterministic then $q_i = q$ for all i, and

First inexact oracle: exploiting structure

$$Q_i(x) = \begin{cases} \min\limits_{y \geq 0} q_i^\top y \\ Wy = h_i - T_i x \end{cases} = \begin{cases} \max\limits_{u} u^\top (h_i - T_i x) \\ W^\top u \leq \boxed{q_i = q} \end{cases} = u_i^\top (h_i - T_i x)$$

If second-stage cost is deterministic then $q_i = q$ for all i, and

$$Q_j(x) \approx Q_i(x) \qquad \Longleftrightarrow \qquad h_j - T_j x \approx h_i - T_i x$$



Collinearity strategy (cos)

cos approximation:

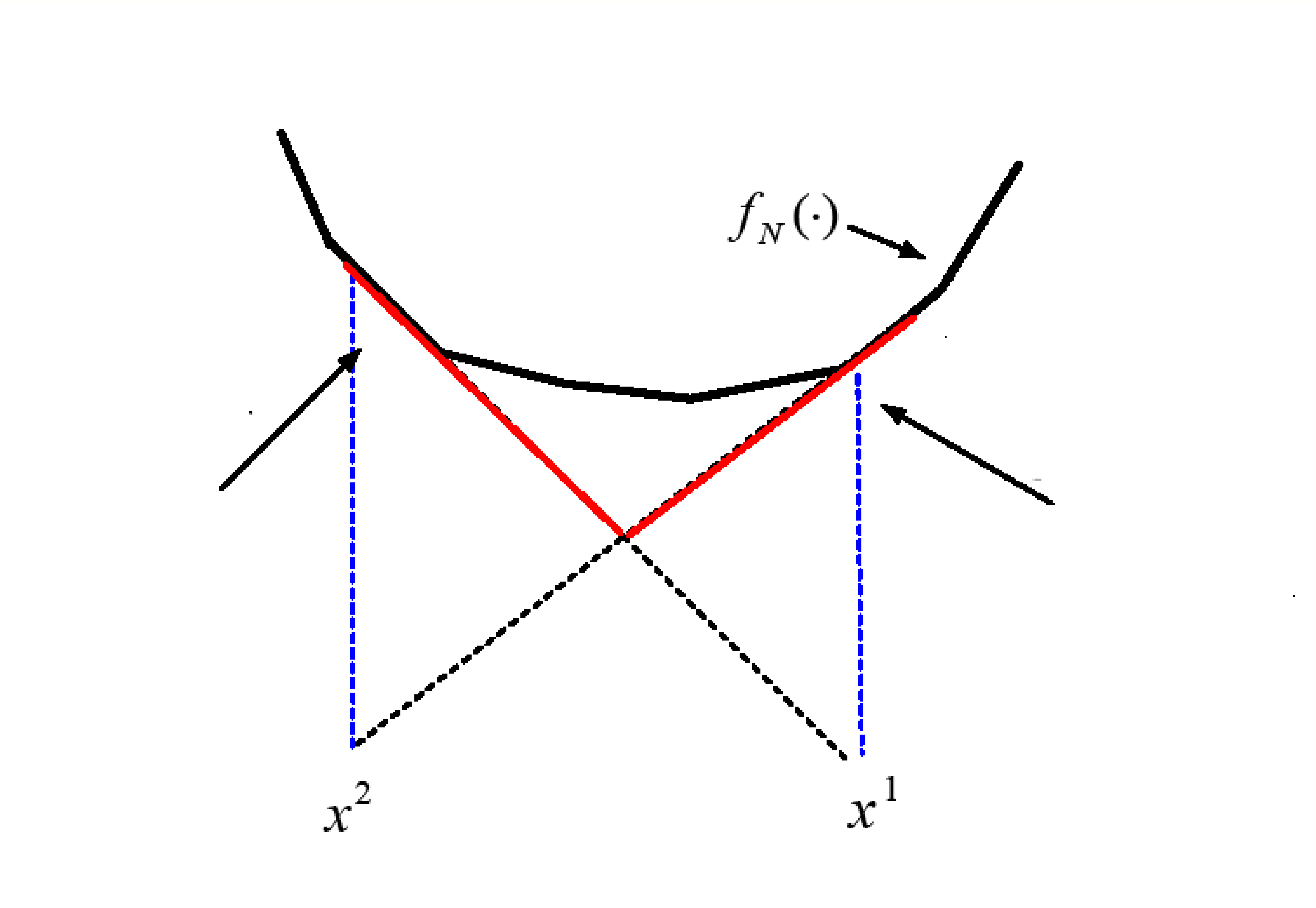
$$
\begin{aligned}
f(x) &= c^\top x + \sum_{\ell=1}^{Nred} p_\ell u_\ell^\top (h_\ell - T_\ell x) + \sum_{j \in \mathcal{I}_\ell, \ell=1}^{Nred} p_j \underbrace{u_{\mathbf{j}}}^\top (h_j - T_j x) \\
f_x &= c^\top x + \sum_{\ell=1}^{Nred} p_i u_\ell^\top (h_\ell - T_\ell x) + \sum_{j \in \mathcal{I}_i, \ell=1}^{Nred} p_j \overbrace{u_\ell}^\top (h_j - T_j x)
\end{aligned}
$$

How NSO method

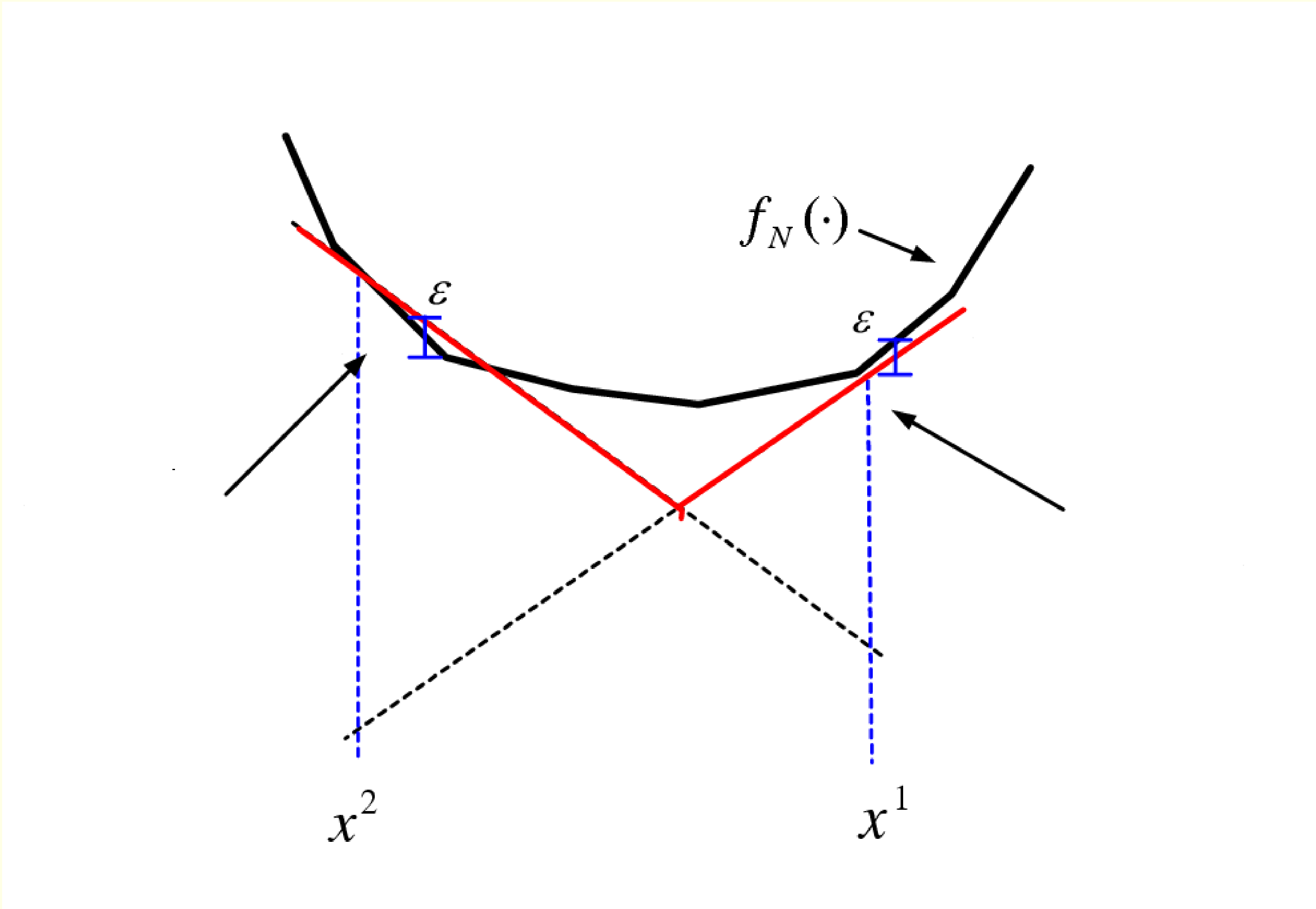
can handle

inaccurate oracles?

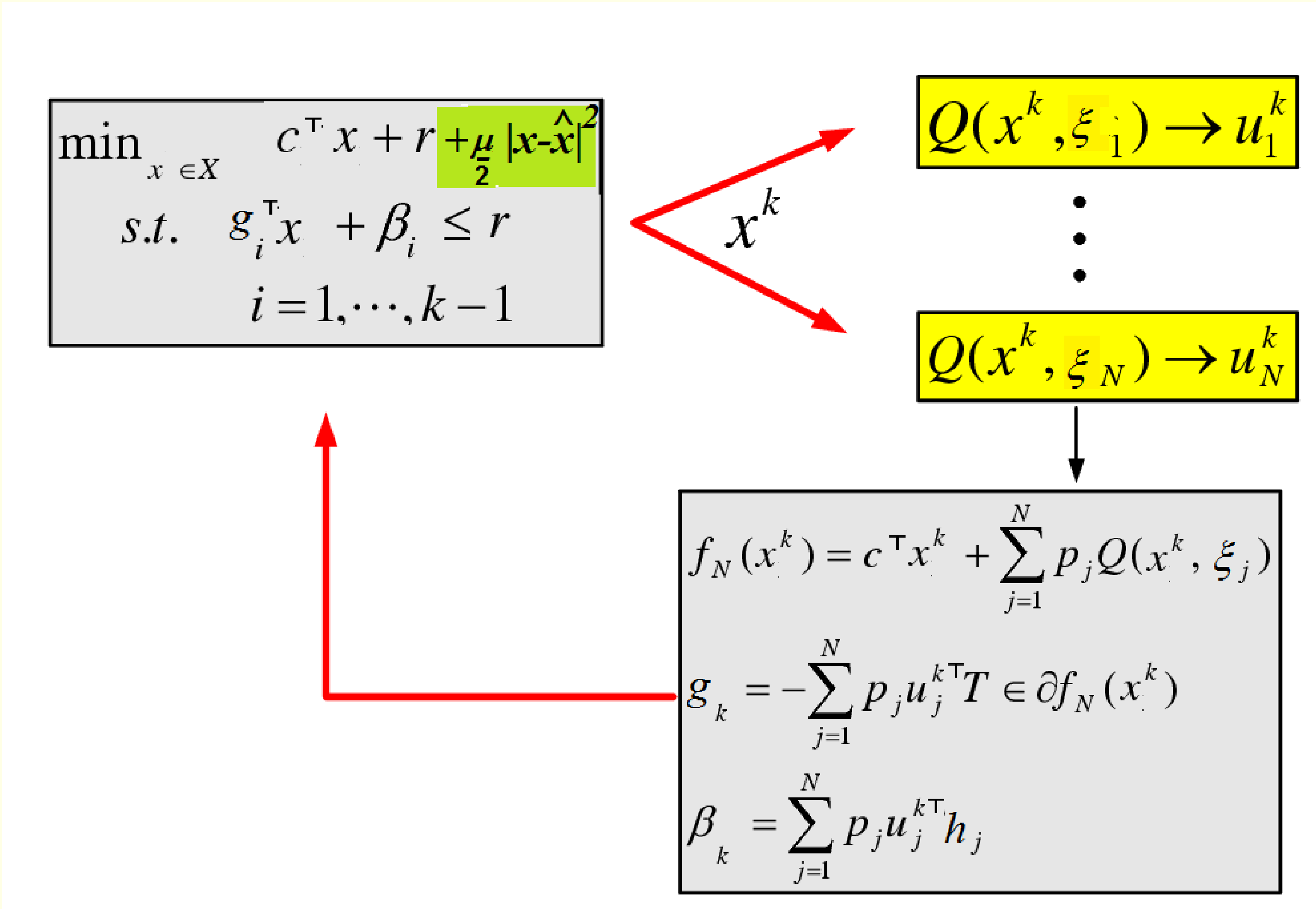
Exact cutting-planes model



Inexact cutting-planes model



The exact bundle method



The inexact bundle method

(Kiwiel 2006)

$$\min_{x \in X} \quad c^\top x + r + \frac{\mu}{2}|x-\hat{x}|^2$$

$$\text{s.t.} \quad g_i^\top x + \beta_i \le r$$

$$i = 1, \cdots, k-1$$

x^k

$$Q(x^k, \xi_1) \to u_1^k$$

$$\vdots$$

$$\cancel{Q(x^k, \xi_N) \to u_N^k}$$

$$f_N(x^k) \approx c^\top x^k + \sum_{j=1}^{N_{red}} p_j Q(x^k, \xi_j)$$

$$g_k = -\sum_{j=1}^{N} p_j u_j^{k\top} T \approx \partial f_N(x^k)$$

$$\beta_k = \sum_{j=1}^{N} p_j u_j^{k\top} h_j$$

modifies μ if inaccuracy becomes cumbersome

converges even if inaccuracy does not vanish

Numerical Results

Proximity between scenarios measured with pseudonorm

$d_\lambda(\xi_i, \xi_j) := \lambda\pi\|\xi_i - \xi_j\| + (1-\lambda)|\mathtt{Q}_i - \mathtt{Q}_j| \;\; \lambda \in [0,1]$

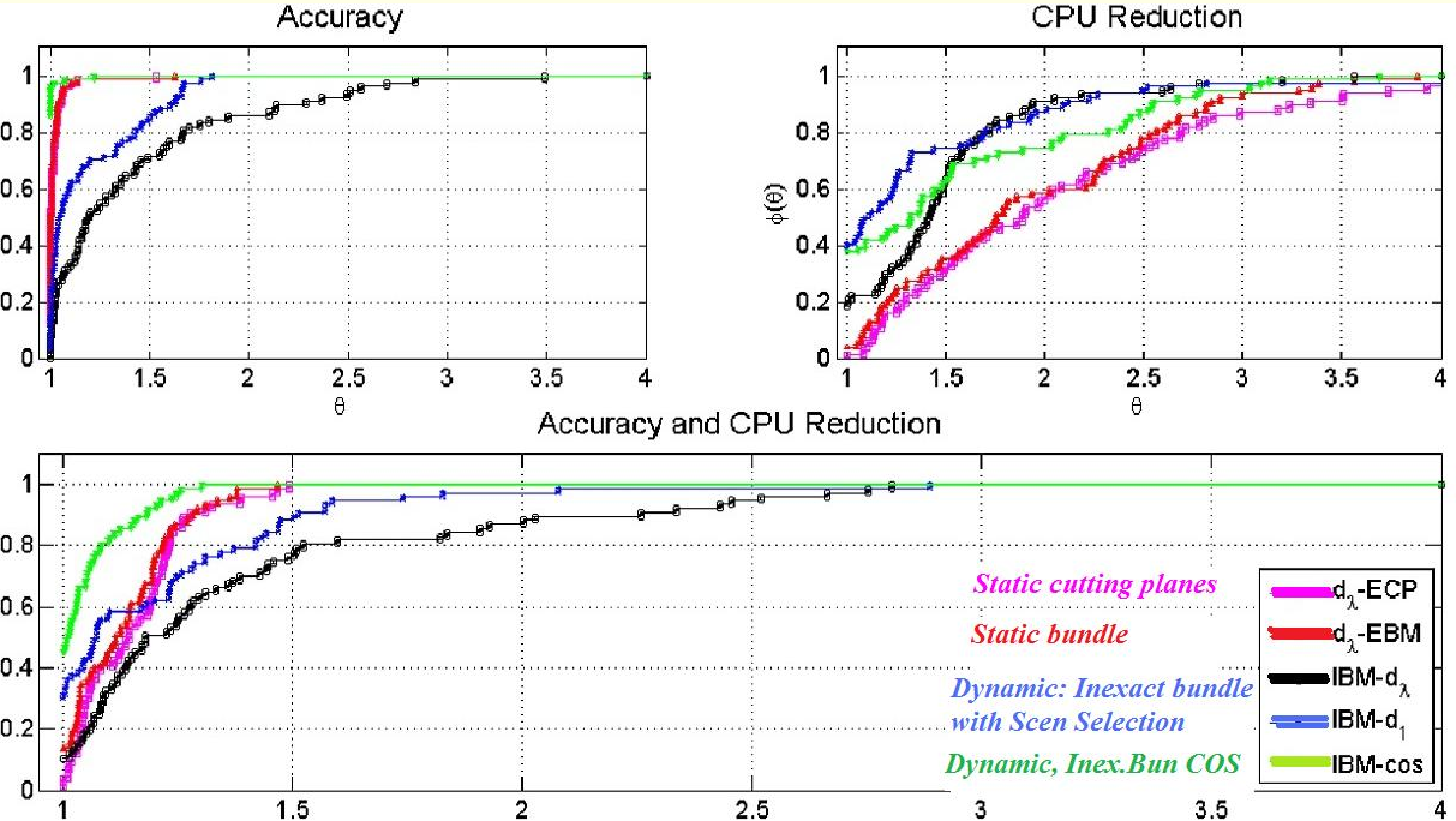
Benchmark of 5 solvers:

$$\text{2 Static} \begin{cases} d_\lambda - ECP & \text{L-shaped on reduced SAA} \\ d_\lambda - EBM & \text{Exact Bundle on reduced SAA} \end{cases}$$

$$\text{3 Dynamic} \begin{cases} IBM - cos & \text{Inexact Bundle method with cos} \\ IBM - d_1 & \text{IBM with scenario selection [DKR03]} \\ IBM - d_\lambda & \text{IBM with pseudonorm} \end{cases}$$

on 10 families of problems ($dim\,\xi \in [2, 200]$), and

$$N = \{100, 200, 300, 500, 800, 1000, 1200, 1500, 1800, 2000, 2500\}$$


Accuracy
CPU Reduction
$\phi(\theta)$
θ
Accuracy and CPU Reduction
Static cutting planes
Static bundle
Dynamic: Inexact bundle
with Scen Selection
Dynamic, Inex.Bun COS
d_λ-ECP
d_λ-EBM
IBM-d_λ
IBM-d_1
IBM-cos

Paris 1, september 1998