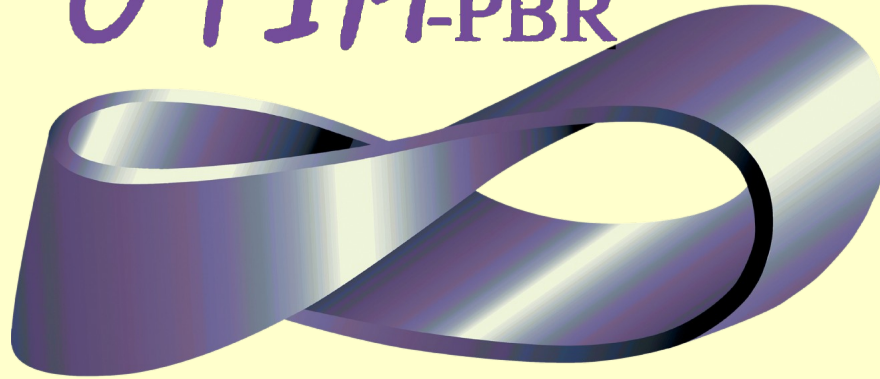


# R&D Project IMPA-Petrobras

*OTIM*-PBR



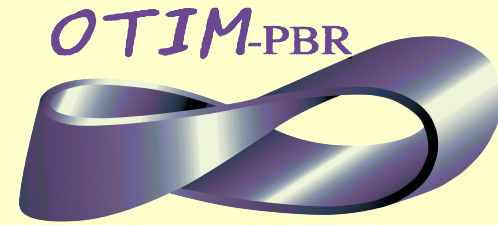
## Introducing Uncertainty in Brazil's Oil Supply Chain

Juan Pablo Luna (UFRJ)

Claudia Sagastizábal (IMPA visiting researcher)

on behalf of OTIM-PBR team

Workshop AASS, April 1st 2016



# OTIM-PBR: main features

General Goal: Introduce uncertainty in “PlanAb”, Petrobras’ planning tool for managing its supply chain (monthly decisions over half a year).

## Specific goals:

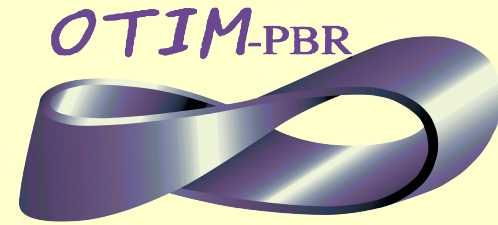
- Representation of stochastic data, focusing on price and volume risk (“statistical” block).
- Definition of a risk-neutral model and direct solution of a small instance.
- Definition of a protocol for evaluating the results.
- The ultimate tool, for huge-scale problems, introducing a decomposition method.

Academic team: IMPA/UFRJ/UERJ/UFSC (10 persons), and a guest from NTNU Norway.

## A Multidisciplinary Team

- IMPA: Mikhail Solodov, Jorge Zubelli
- UFRJ: Laura Bahiense, Carolina Effio, José Herskovits, Juan Pablo Luna
- UERJ: Welington de Oliveira
- UFSC: Marcelo Córdova, Erlon Finardi
- NTNU: Asgeir Tomasgard
- PETROBRAS:
  - Paulo Ribas (Supply chain department)
  - Flavia Schittine (OR department)
  - Sergio Bruno (Corporate risk department)

# OTIM-PBR: main features



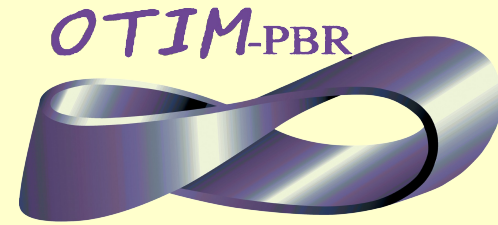
General Goal: Introduce uncertainty in “PlanAb”, Petrobras’ planning tool for managing its supply chain (monthly decisions over half a year).

## Specific goals:

- Representation of stochastic data, focusing on price and volume risk (“statistical” block).
- Definition of a risk-neutral model and direct resolution of a small instance.
- Definition of a protocol for evaluating the results.
- The ultimate tool, for huge-scale problems, introducing a decomposition method.

Academic team: IMPA/UFRJ/UERJ/UFSC (10 persons), and a guest from NTNU Norway.

# Computational model



## Block 1: Statistical

Tool generating scenarios for:

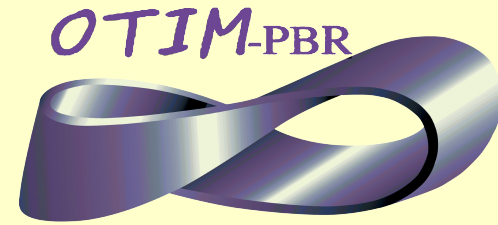
- International prices (oil and derivatives, correlated)
- Oil volume arriving from the platforms each month.

## Block 2: Optimization

Tool to determine the optimal policy for the whole supply chain of the company for the next month. It includes:

- To which extent the company network (of production, transportation, commercialization) can be simplified?
- How to handle uncertainty in the optimization problem (in the cost and in the right hand side)
- Problem solution via decomposition method (huge scale problem)

# Computational model



## Block 1: Statistical

Tool generating scenarios for:

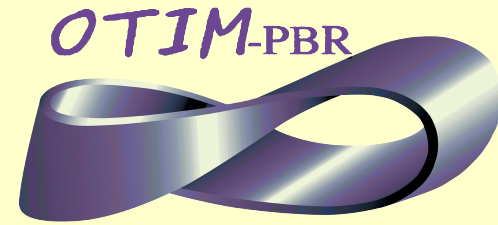
- International prices (oil and derivatives, correlated)
- Oil volume arriving from the platforms each month.

## Block 2: Optimization

Tool to determine the optimal policy for the whole supply chain of the company for the next month. It includes:

- To which extent the company network (of production, transportation, commercialization) can be simplified?
- How to handle uncertainty in the optimization problem (in the cost and in the right hand side)
- Problem solution via decomposition method (huge scale problem)

# Computational model



## **Block 1: Statistical**

Tool generating scenarios for:

- International prices (oil and derivatives, correlated)
- Oil volume arriving from the platforms each month.

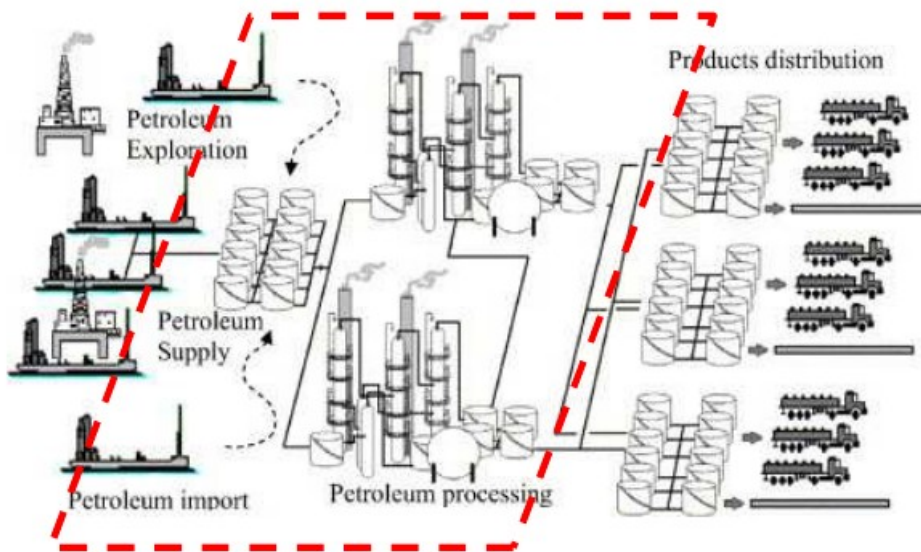
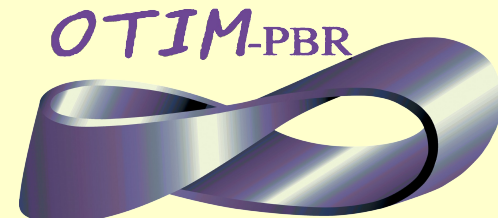
Technique: multivariate model + Kalman filter + SDP NLP solver, assessed by backtesting (Juan Pablo Luna).

## **Block 2: Optimization**

Tool to determine the optimal policy for the whole supply chain of the company for the next month. It includes:

- To which extent the company network (of production, transportation, commercialization) can be simplified?
- How to handle uncertainty in the optimization problem (in the cost and in the right hand side)
- Problem solution via decomposition method (huge scale problem)

# The setting



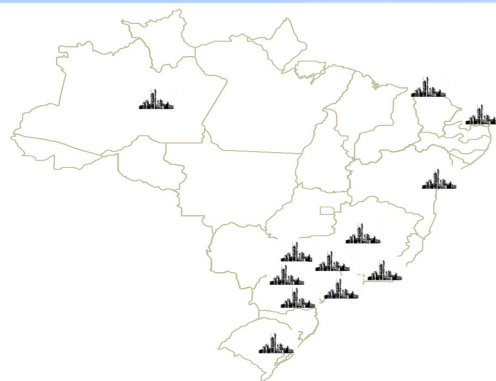
Wide variety of commercial, industrial and logistics operations occur over the midstream segment

Strong dependence among the operations and some gains can only be estimated by considering the whole supply chain

Planning this segment is crucial to achieve the success of all operations.



# Brazil's supply chain



Source: M. Maia,  
Petrobras

## Crude Oil:

- Imports (1)
- Exports (3)
- Production (5)

## Oil Products:

- Imports (2)
- Exports (4)
- Market Selling (6)

~ 200 different crude oils

~ 50 different oil products

10 basins of oil production

23000 km pipelines

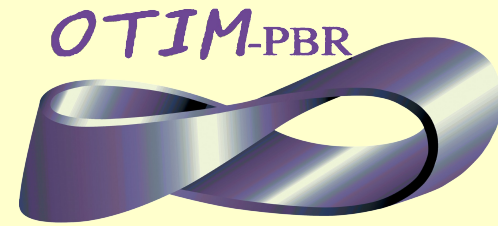
>100 platforms

> 40 terminals in Brazil

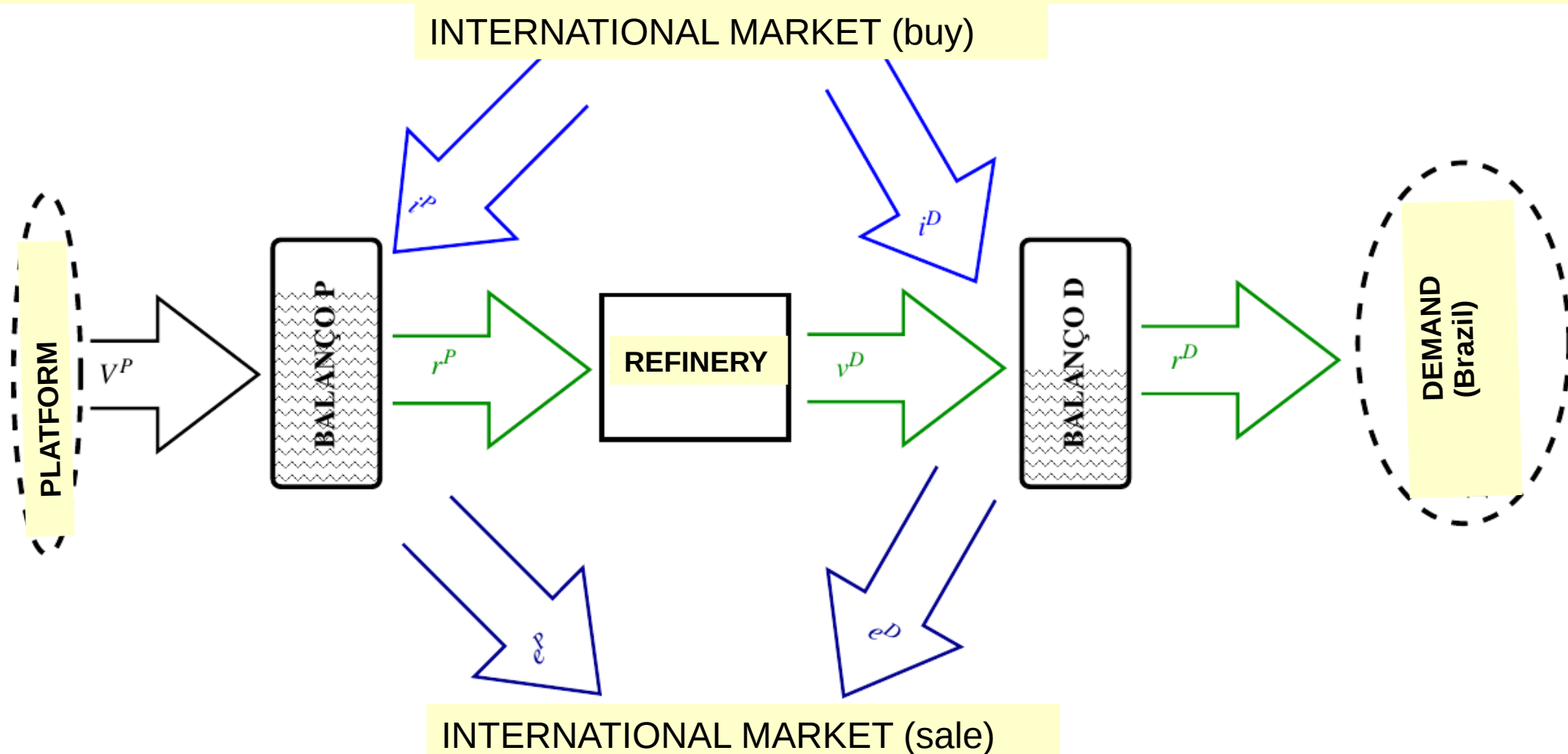
>200 terminals abroad

12 refineries

# Production flow PlanAb

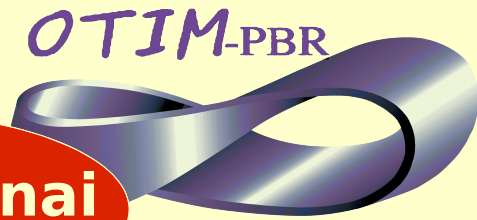


## Simplified Network

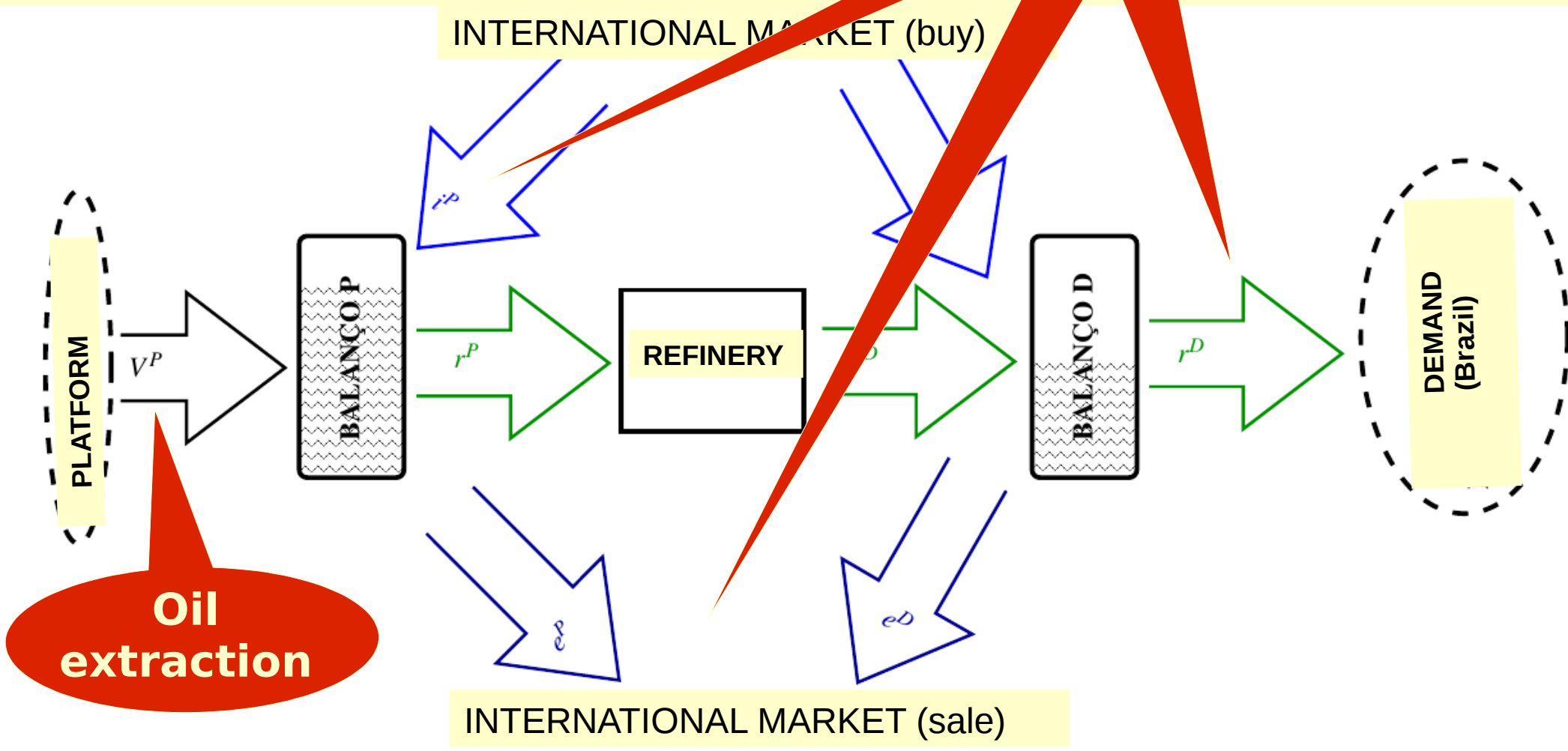


# Production flow

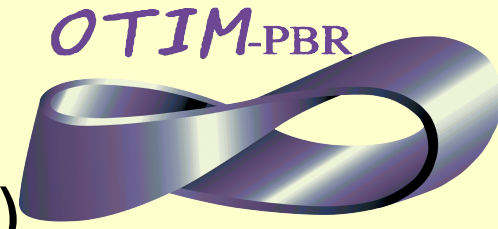
## PlanAb **UNCERTAINTY**



Simplified Network

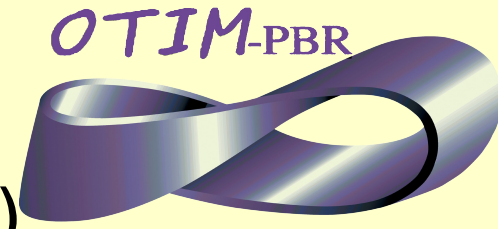


# Some comments



- Uncertainty increases the (already huge) size of the linear program (LP).
- Mounting the data for the deterministic model takes longer than solving the LP.
- Solving an aggregate model is not an option: keeping the level of detail of the deterministic model is important for the company.

# Some comments



- Uncertainty increases the (already huge) size of the linear program (LP).
- Mounting the data for the deterministic model takes longer than solving the LP.
- Solving an aggregate model is not an option: keeping the level of detail of the deterministic model is important for the company.

## **BUT**

- Changing CPLEX stopping tolerance from default  $10^{-8}$  to  $10^{-4}$  provides a good trade-off between accuracy and solving time:
  - mean relative error on variables 0,007%
  - solving times reduced in 18,56%

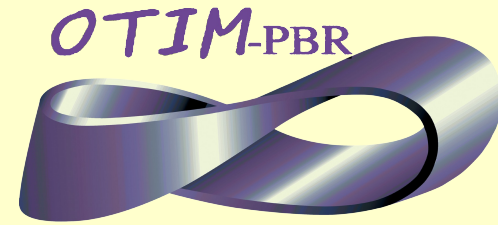
# Statistical Block

*OTIM*-PBR



# Sources of uncertainty

## Calibration of price models

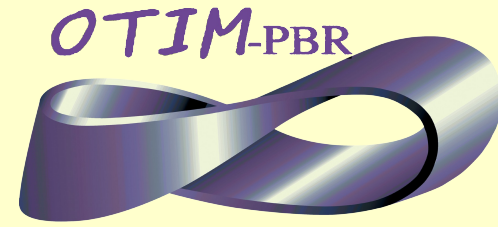


1. Oil, gasoline and diesel international prices
  - The model is a process defined by stochastic differential equations.
  - To determine the parameters defining the model (mean, trend, volatility, correlations) we maximize the proximity to historical data (likelihood) using a Kalman filter.
  - Small scale non-convex optimization problem, solved by a nonlinear programming method of interior feasible directions. Master's thesis supervised by J.P. Luna and J. Herskovits.



# Sources of uncertainty

## Calibration of price models

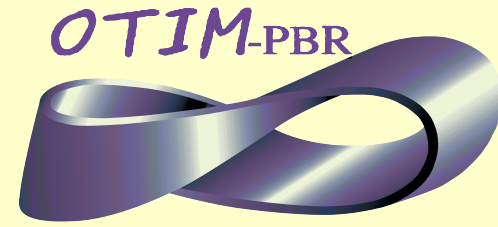


- There are several stochastic processes for modelling commodity prices.
- They should include important oil price features such as picks, seasonality, mean reversion, etc.
- The models are completely described by certain parameters that can be time dependent (mean, trend, volatility, correlations) and whose values need to be estimated (calibration).



# Sources of uncertainty

## Calibration of price models

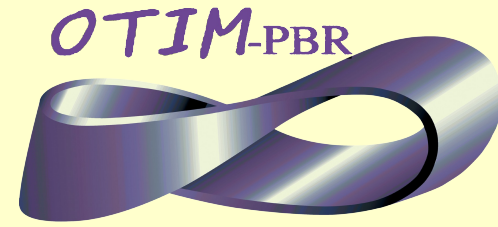


- **Agenda:**

1. Choose a model suitable for our purposes.
2. Calibrate the model for fitting the available historical data.
3. Generate future price scenarios.

# Sources of uncertainty

## Calibration of price models



- Calibration:
- To determine the parameters defining the model we maximize the proximity to historical data (likelihood).
- To compute the likelihood requires computing the joint probability density function that may not be available in closed form but can be estimated through Kalman Filters.
- The likelihood function is nonconvex.

# Schwartz - Smith Model

$$\log(S_t) = a\chi_t + b\xi_t$$

$$d\chi_t = -\text{diag}(\kappa)\chi_t dt + A^\chi dW_t^\chi$$

$$d\xi_t = \mu dt + A^\xi dW_t^\xi$$

$$dW_t^\chi dW_t^\xi = \hat{\rho}^{\chi\xi}$$

$$\delta = \left( a, b, \kappa, A^\chi, A^\xi, \lambda_\xi, \mu_\xi, \mu_\xi^*, \hat{\rho}^{\chi\xi} \right)$$

# Future Contract Price

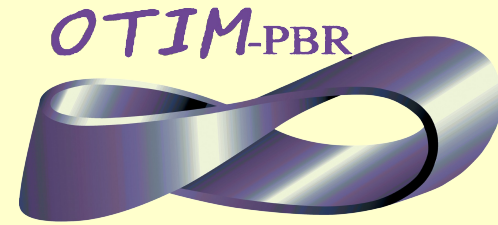
$$F_{t,T}^i = \mathbb{E}^*[S_T^i | \mathcal{F}_t]$$

$$\log(F_{t,T}^i) = a \left[ e^{-k_i(T-t)} \chi_t^i \right] + b \xi_t^i + A^i(t, T)$$

$$\begin{aligned} A^i(t, T) = & b [(\mu_i - \lambda_{\xi_i})(T - t)] - a \left[ \frac{\lambda_{\chi_i}}{k_i} (1 - e^{-k_i(T-t)}) \right] \\ & + \frac{1}{2} \left( (1 - e^{-2k_i(T-t)}) \frac{\|a e_i^\top A^\chi\|^2}{2k_i} + 2(1 - e^{-k_i(T-t)}) \frac{a b e_i^\top A^\chi \hat{\rho}^{\chi\xi} A^\xi{}^\top e_i}{k_i} + \right. \\ & \left. \|b e_i^\top A^\xi\|^2 (T - t) \right) \end{aligned}$$

Note the nonlinear relations between unknown variables

# Discretization of Schwartz – Smith Model



$$\begin{bmatrix} \chi_{t_{i+1}} \\ \xi_{t_{i+1}} \end{bmatrix} = \mathbb{E} \left[ \begin{bmatrix} \chi_{t_{i+1}} \\ \xi_{t_{i+1}} \end{bmatrix} \mid \mathcal{F}_{t_i} \right] + w_{t_i}$$

$$\mathbb{E} \left[ \begin{bmatrix} \chi_{t_{i+1}} \\ \xi_{t_{i+1}} \end{bmatrix} \mid \mathcal{F}_{t_i} \right] = \begin{bmatrix} \text{diag}(e^{-k\Delta t}) \chi_{t_i} \\ \xi_{t_i} + \mu \Delta t \end{bmatrix}$$

$$\text{Cov}(w_{t_i}) = \begin{bmatrix} \tilde{\Sigma}^{\chi} & \text{diag} \left( \left( \frac{1-e^{-k_i\Delta t}}{k_i} \right)_i \right) A^{\chi} \hat{\rho}^{\chi\xi} A^{\xi\top} \\ A^{\xi} (\hat{\rho}^{\chi\xi})^{\top} A^{\chi\top} \text{diag} \left( \left( \frac{1-e^{-k_i\Delta t}}{k_i} \right)_i \right) & \Delta t \Sigma^{\xi} \end{bmatrix}$$

$$\tilde{\Sigma}_{ij}^{\chi} = \frac{1 - e^{-(k_i+k_j)\Delta t}}{k_i + k_j} \Sigma_{ij}^{\chi}$$

We need to keep positive definite this nonlinear matrix

# Implementation Issues

1. Objective function (evaluated through a Kalman Filter) is highly nonlinear.
2. Objective function is defined only on certain set: feasible optimization methods must be used
3. Nonlinear constraints must include the positive definiteness of correlation matrices.
4. Optimization methods are highly sensitive to gradient values: an accurate implementation is needed.

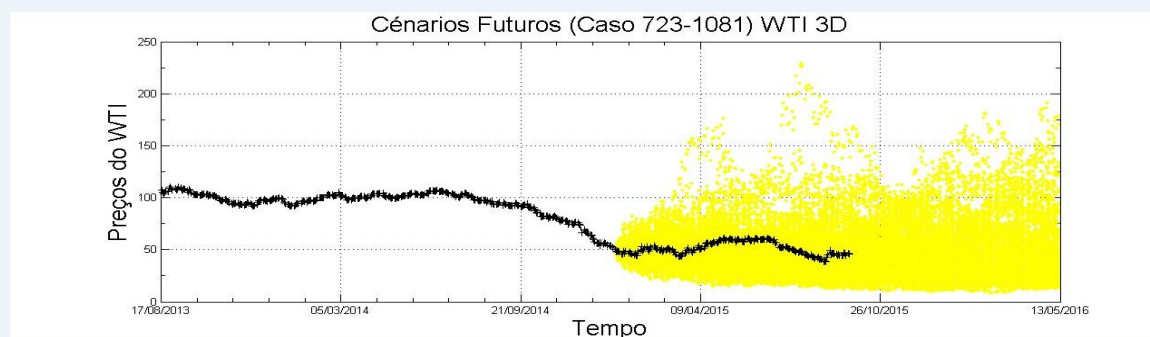
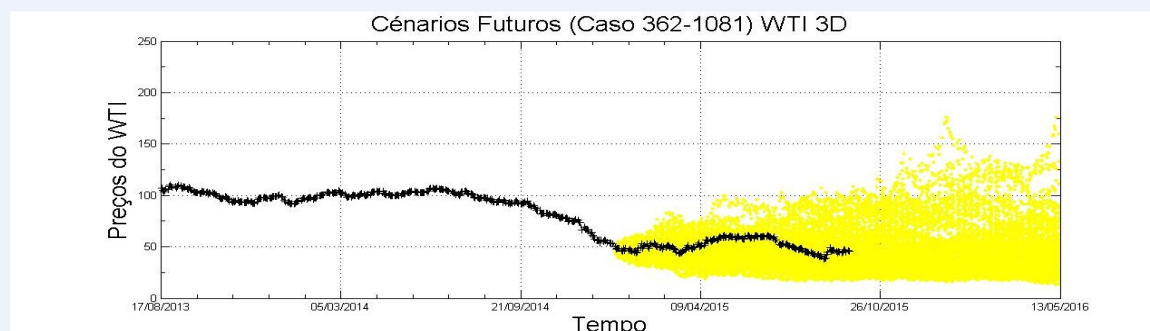
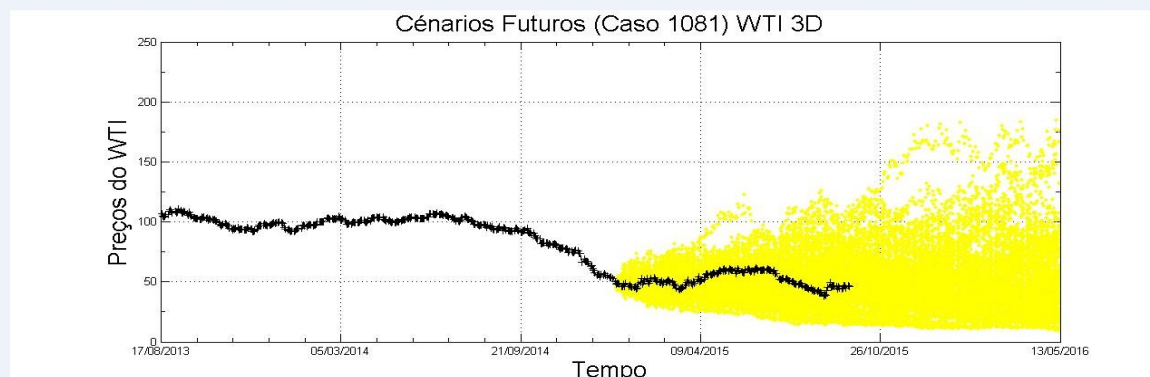
# Numerical Results

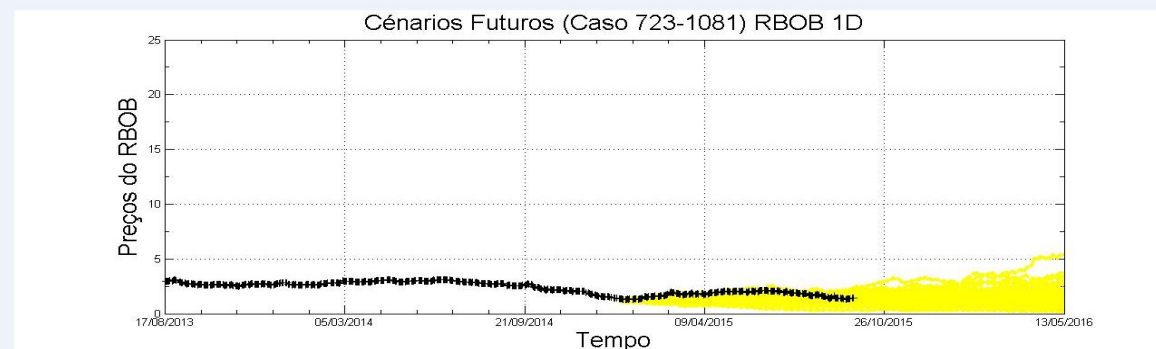
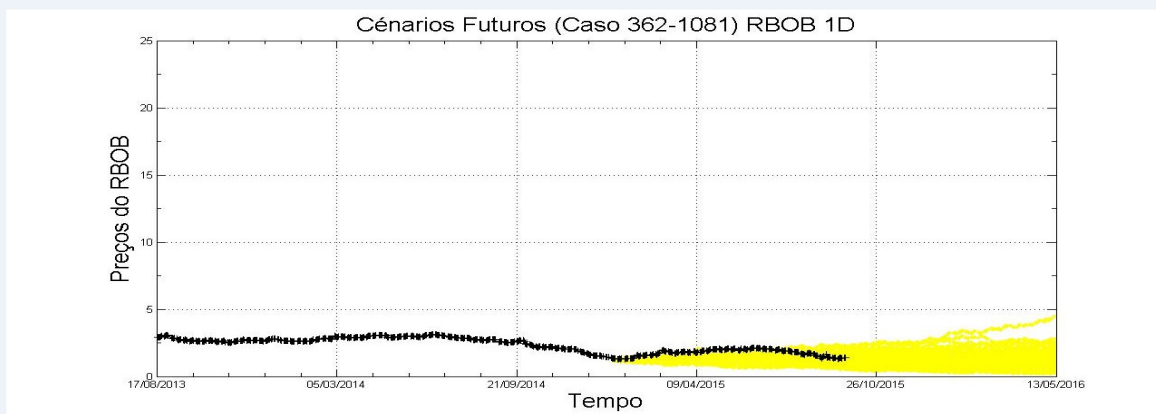
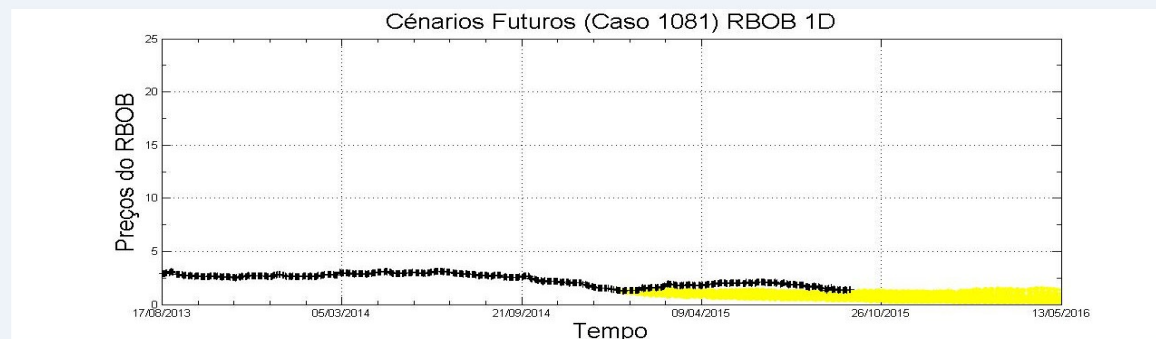
- WTI, HO, RBOB future contracts (with 1 to 4 months of maturity) from Energy Information Administration-EIA.  
([http://www.eia.gov/dnav/pet/xls/PET\\_PR\\_FUT\\_S1-D.xls](http://www.eia.gov/dnav/pet/xls/PET_PR_FUT_S1-D.xls))
- Time period considered 1081 days (10/06/2011 to 22/09/2015)
- Three decks (1081, 720 and 359 days) for calibrating Schwartz – Smith.

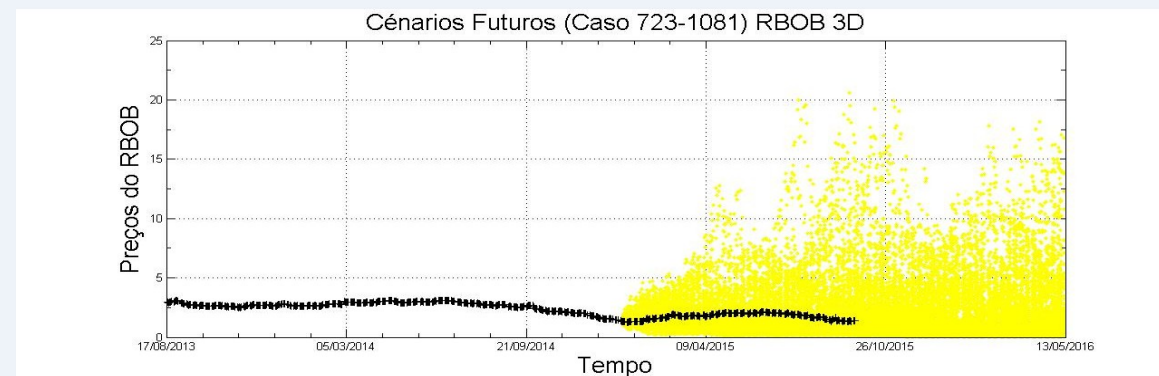
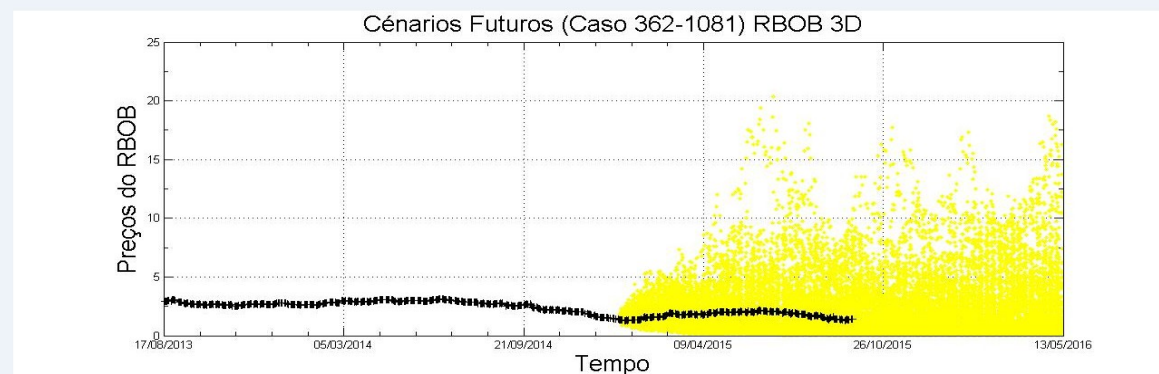
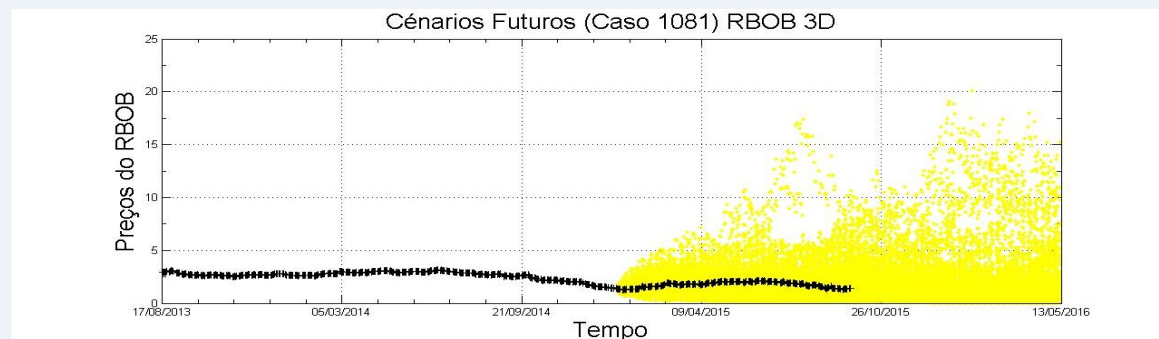
# Numerical Results

- We considered 1D and 3D Schwartz – Smith processes.
- The optimization problems were solved using FDIPA and FDIPA-SDP non linear programming algorithms.
- Numerical approximations of likelihood functions versus its exact evaluation.









# Conclusions

- As expected, considering together all three assets produces better simulations of prices.
- Considering Schwartz – Smith models for more than one dimension leads to more challenging numerical problems that need sophisticated optimization solvers (and expertise from the optimization community).

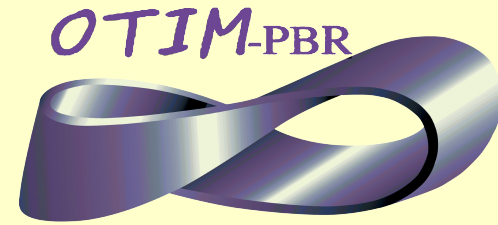
# Optimization Block

*OTIM*-PBR



# Supply chain planning tool

## Computational model



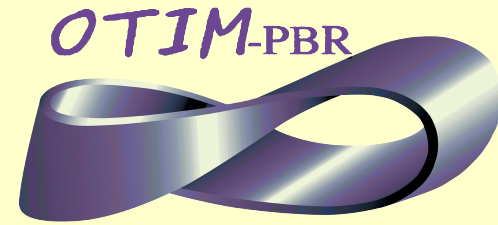
### PlanAb

- Computational model for Petrobras supply chain
- Tool to determine the optimal planning for the whole supply chain of the company for the next month.
- Currently PlanAb is a deterministic model that solves a large scale linear program

$$\left\{ \begin{array}{ll} \min_{x,y} & c^\top x + q^\top y \\ \text{s.t} & Ax = b \\ & Tx + Wy = h \\ & \underline{x} \leq x \leq \bar{x} \\ & \underline{y} \leq y \leq \bar{y} \\ & y_i = \bar{y}_i \quad i \in I_V \end{array} \right.$$

# Supply chain planning tool

## Computational model



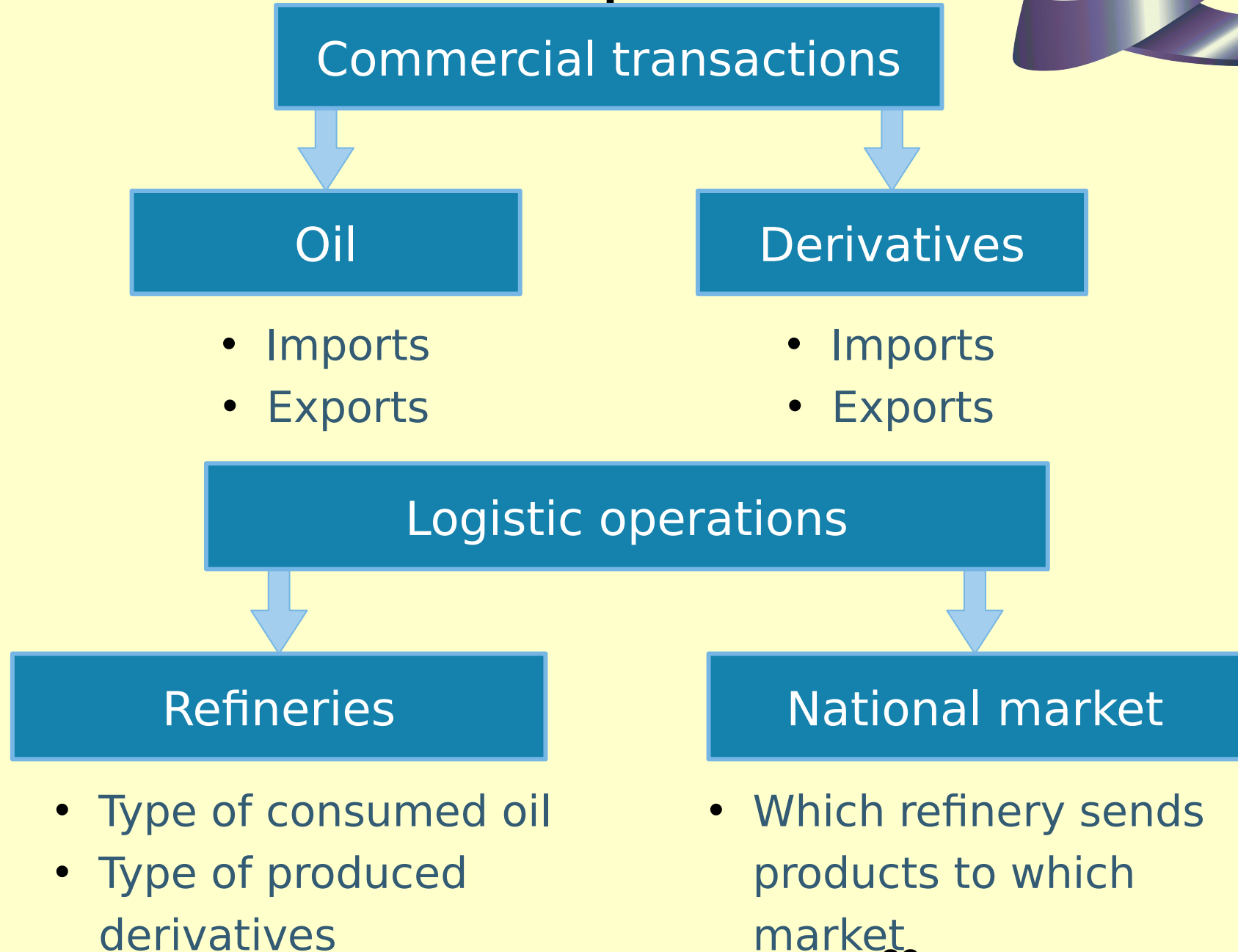
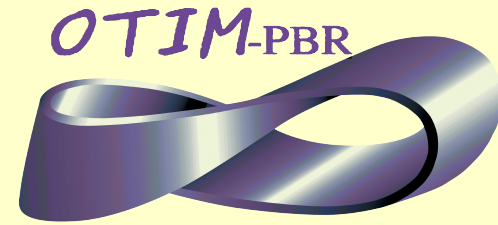
### PlanAb

$$\left\{ \begin{array}{ll} \min_{x,y} & c^\top x + q^\top y \\ \text{s.t} & Ax = b \\ & Tx + Wy = h \\ & \underline{x} \leq x \leq \bar{x} \\ & \underline{y} \leq y \leq \bar{y} \\ & y_i = \bar{y}_i \quad i \in I_V \end{array} \right.$$

Availability of oil volume from the plataforms

- Variable  $x$  represents commercial transactions such as imports along the planning horizon (four months)
- Variable  $y$  represents exports and logistic operations such transportation and refinery oil/derivative production

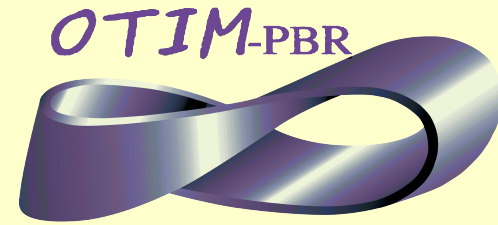
# Most relevant output





# Objective function

## PlanAb model



Current model:  $\min_{x,y} c^\top x + q^\top y$

Model (objective function)

Costs

- Oil imports
- Derivative imports
- Derivative storage
- Refining unit operations
- Transportation

Revenues

- Oil exports
- Derivative exports
- Derivative national sales

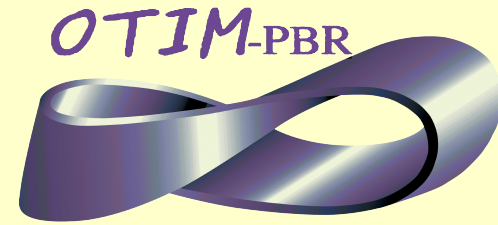
Prices are uncertain....

Proposed model:

$$\min_{x,y(\xi)} c^\top x + \mathbb{E}[q(\xi)^\top y(\xi)]$$

# Variables and constraints

## PlanAb model



### 1. Oil and derivatives balance

- Includes all constraints involving the system balance of oil/derivatives and how they flow in the network
- Balancing the derivatives in each terminal and refinery

$$\begin{array}{rcl} Ax & & = b \\ Tx + Wy & = & h \end{array}$$

- Considering adjustments on the **available oil** for:
  - Processing in the refineries
  - Exportation, when the volume of oil arriving from the platforms is larger than **foreseen**

$$\begin{array}{rcl} \underline{y} \leq y \leq \bar{y} \\ y_i = \bar{y}_i & i \in I_V \end{array}$$

# Sources of uncertainty

1. Oil, gasoline and diesel international prices (about 30 in total)

Stochastic process: multivariate Schwartz-Smith (Juan Pablo's talk)

This uncertainty is represented by



2. Availability of national oil

The ratio between the foreseen and observed volumes follows a log-logistic probability distribution

This uncertainty is represented by



**Two different sources of uncertainty!**

# PlanAb

- Deterministic

How large are these models?

$$\left\{ \begin{array}{ll} \min_{x,y} & c^\top x + q^\top y \\ \text{s.t} & Ax = b \\ & Tx + Wy = h \\ & \underline{x} \leq x \leq \bar{x} \\ & \underline{y} \leq y \leq \bar{y} \\ & y_i = \bar{y}_i \quad i \in I_V \end{array} \right.$$

- Stochastic

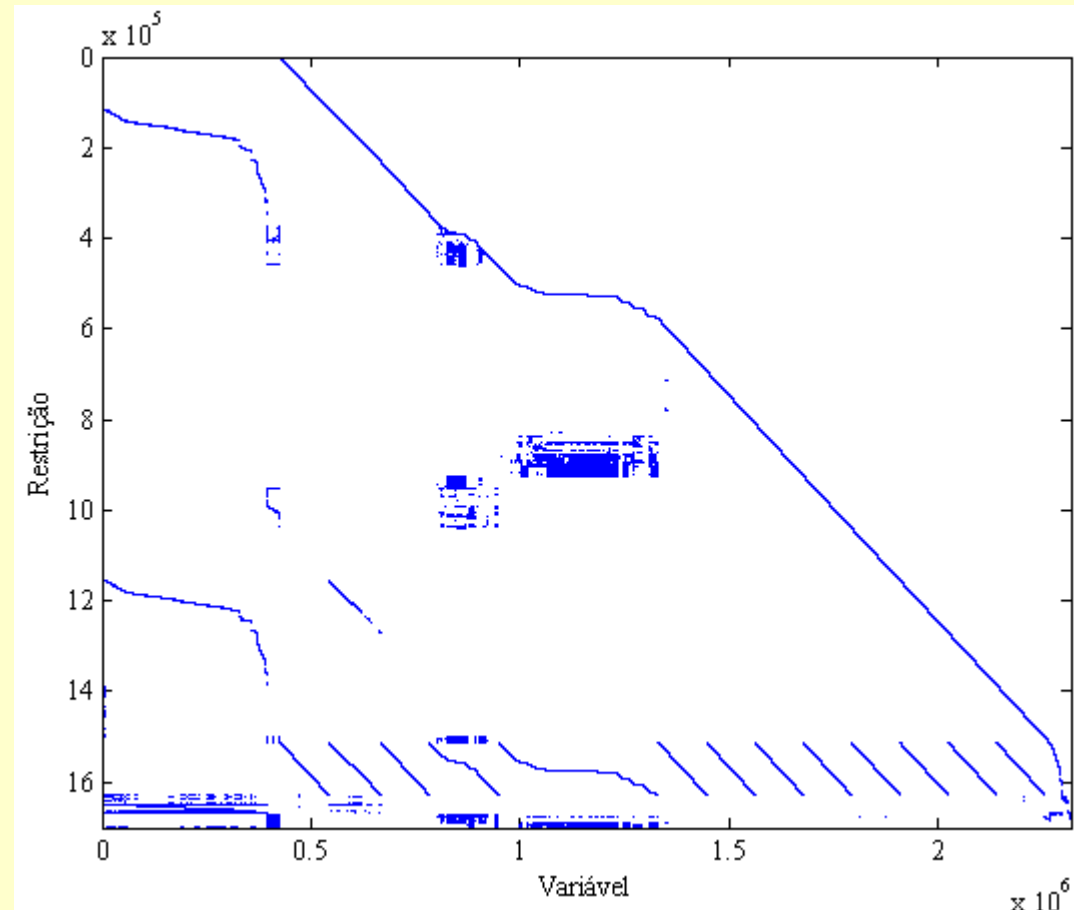
$$\left\{ \begin{array}{ll} \min_{x,y} & c^\top x + \mathbb{E}[q(\xi)^\top y(\xi, \omega)] \\ \text{s.t} & Ax = b \\ & Tx + Wy(\xi, \omega) = h \\ & \underline{x} \leq x \leq \bar{x} \\ & \underline{y} \leq y(\xi, \omega) \leq \bar{y} \\ & y_i(\xi, \omega) = \bar{y}_i(\omega) \quad i \in I_V \\ & \forall \xi \in \Xi \quad \omega \in \Omega \end{array} \right.$$

# Deterministic PlanAb

Implemented in AIMMS

- LP's dimension:
  - 2.3 million of variables
  - 1.7 million of constraints
- Constraints' matrix:
  - Sparse (0,0002% non-zero elements; approximately 7.3 million of elements)
  - Block-diagonal structure

$$\left\{ \begin{array}{ll} \min_{x,y} & c^\top x + q^\top y \\ \text{s.t} & Ax = b \\ & Tx + Wy = h \\ & \underline{x} \leq x \leq \bar{x} \\ & \underline{y} \leq y \leq \bar{y} \\ & y_i = \bar{y}_i \quad i \in I_V \end{array} \right.$$



# Deterministic PlanAb

- PlanAb LP:
  - 2.3 million of variables, 1.7 million of constraints
  - 7.3 million of non-zero elements (0.0002% - **sparse**)
  - After presolve, size reduced approximately by 80%

Model	Variables	Constraints	Non-zero elements	Presolve time (sec)
Original	2.3 million	1.7 million	7.3 million	0
Conservative presolve	644,000	307,000	2.88 million	2.7
Automatic presolve	385,000	94,000	2.19 million	11.1
Aggressive presolve	385,000	92,000	2.14 million	13.2

# Deterministic PlanAb

- Solution of PlanAb by CPLEX solver, using different methods

Method	Time AIMMS (sec)	Memory
Primal simplex	600*	800 MB
Dual simplex	600*	800 MB
Network + Primal	600*	770 MB
Network + Dual	600*	820 MB
Barrier	180	1.25 GB
Barrier - Primal crossover	195	1.25 GB
Barrier - Dual crossover	195	1.25 GB
Sifting	600*	1.1 GB
Concurrent	352	1.8 GB

\* Limit for time execution

# Stochastic PlanAb

$$\left\{ \begin{array}{ll} \min_{x,y} & c^\top x + \mathbb{E}[q(\xi)^\top y(\xi, \omega)] \\ \text{s.t.} & Ax = b \\ & Tx + Wy(\xi, \omega) = h \\ & \underline{x} \leq x \leq \bar{x} \\ & \underline{y} \leq y(\xi, \omega) \leq \bar{y} \\ & y_i(\xi, \omega) = \bar{y}_i(\omega) \quad i \in I_V \\ & \forall \xi \in \Xi \quad \omega \in \Omega \end{array} \right.$$

- Problem's size depends on the number of scenarios of price and oil volume
  - Number of variables is N times 2.3 million
  - Number of constraints is N times 1.7 million

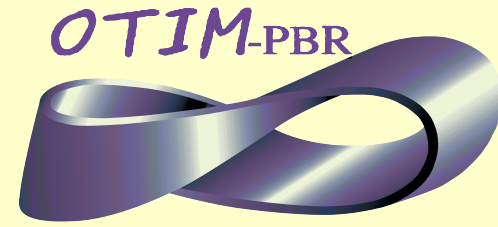
$$N = |\Xi| * |\Omega|$$

- Impossible to load the problem in a (powerful) computer for  $N > 10!$  (Memory issues)



# Two-stage decomposition

## Stochastic PlanAb



- Consider finitely many scenarios of price and oil volume
- The problem is decomposed into two decision levels

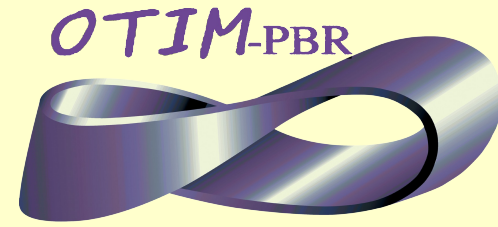
$$\begin{cases} \min_x c^\top x + \mathbb{E}[Q(x; \xi, \omega)] \\ \text{s.t.} & Ax = b \\ & \underline{x} \leq x \leq \bar{x} \end{cases}$$

with

$$Q(x; \xi, \omega) \begin{cases} \min_y q(\xi)^\top y \\ \text{s.t.} & Wy = h - T x \\ & \underline{y} \leq y \leq \bar{y} \\ & y_i = \bar{y}_i(\omega) \quad i \in I_V \end{cases}$$

# Two-stage decomposition

## Stochastic PlanAb



- Consider finitely many scenarios of price and oil volume
- The problem is decomposed into two decision levels

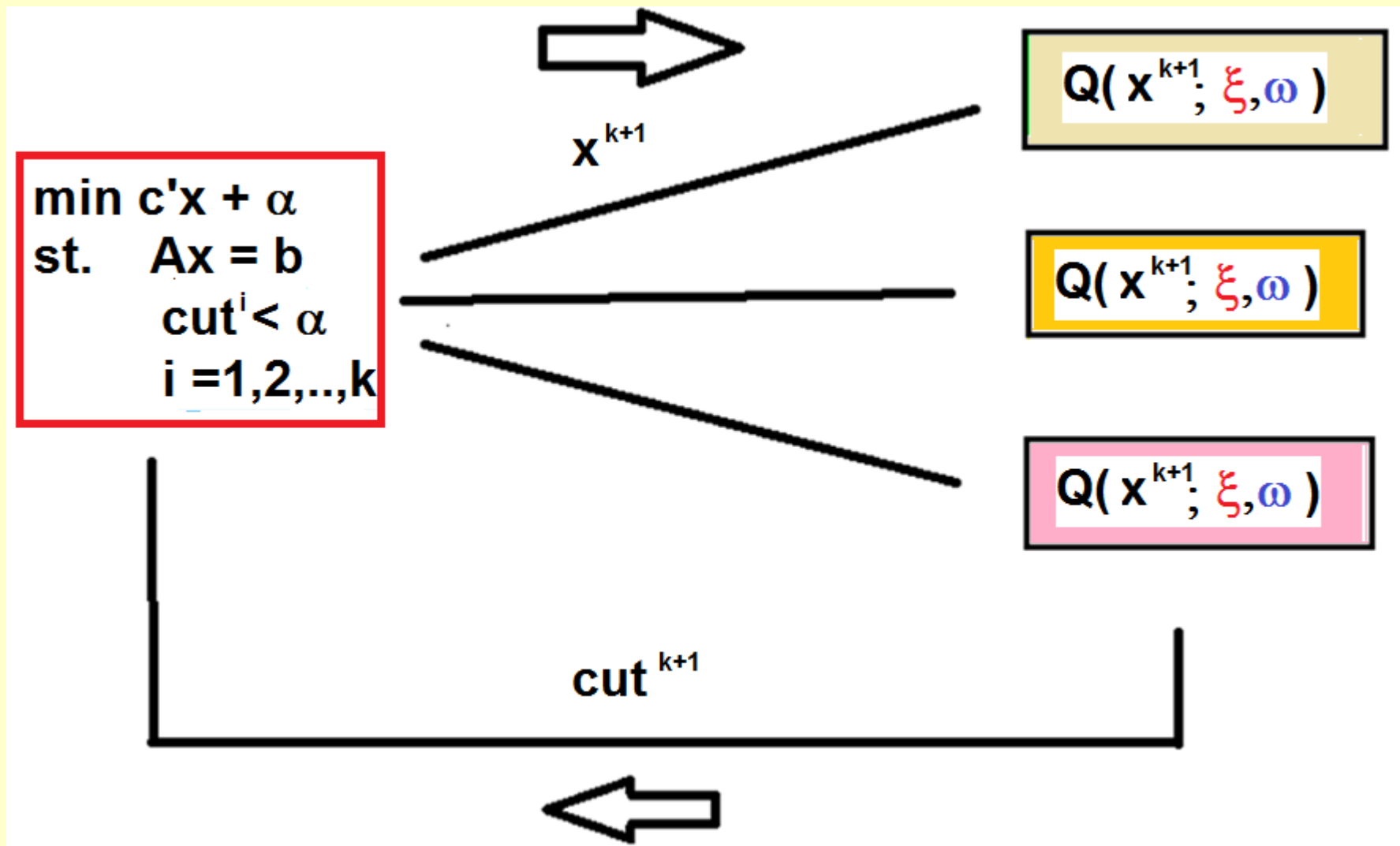
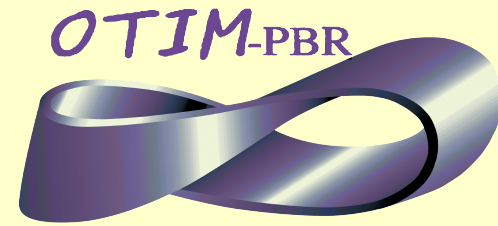
$$\begin{cases} \min_x c^\top x + \sum_{i=1}^N p_i [Q(x; \xi_i, \omega_i)] \\ \text{s.t.} & Ax = b \\ & \underline{x} \leq x \leq \bar{x} \end{cases}$$

with

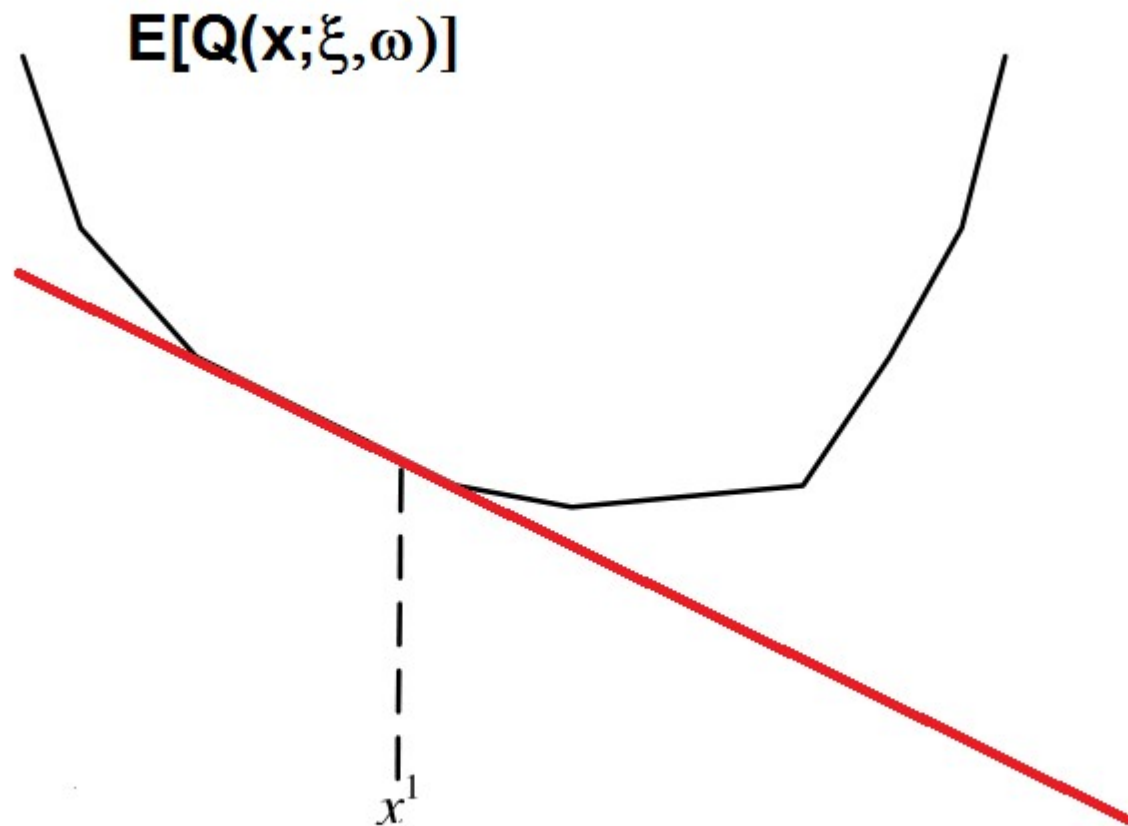
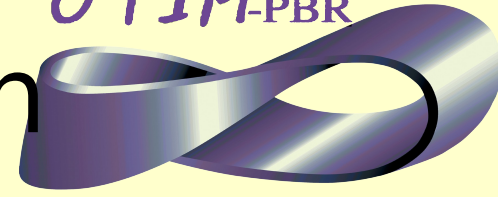
$$Q(x; \xi, \omega) \begin{cases} \min_y q(\xi)^\top y \\ \text{s.t.} & Wy = h - T x \\ & \underline{y} \leq y \leq \bar{y} \\ & y_i = \bar{y}_i(\omega) \quad i \in I_V \end{cases}$$

# Two-stage decomposition

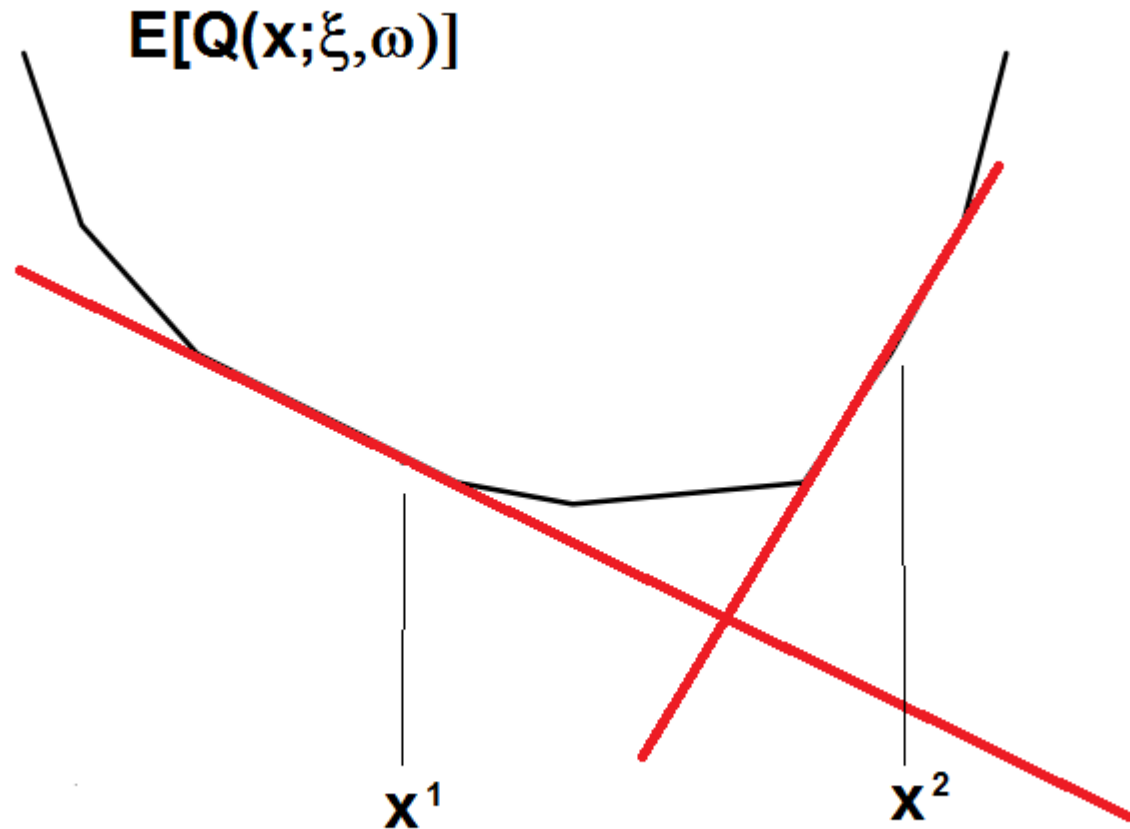
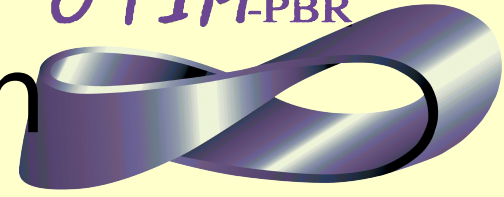
## Stochastic PlanAb



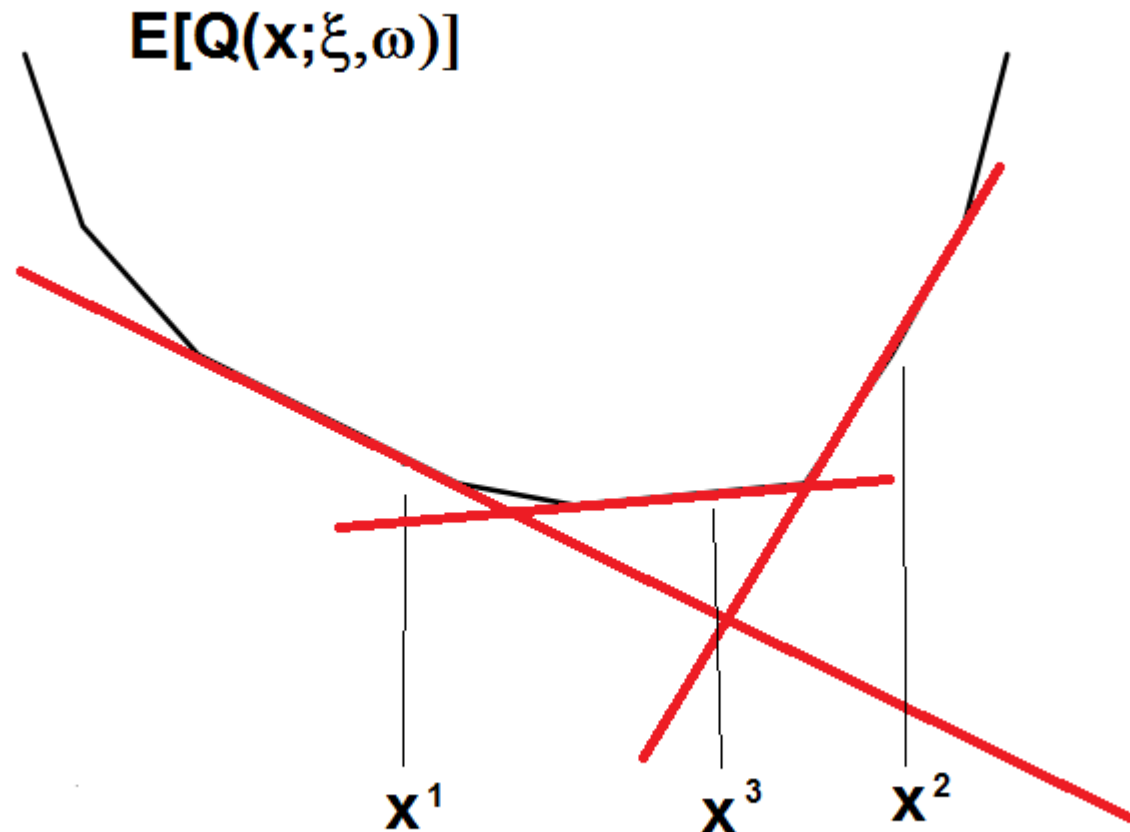
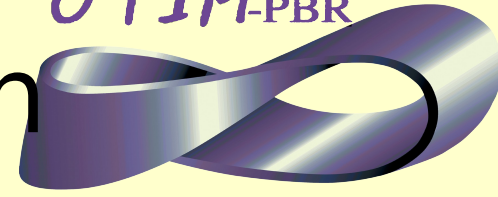
# Cutting-plane approximation



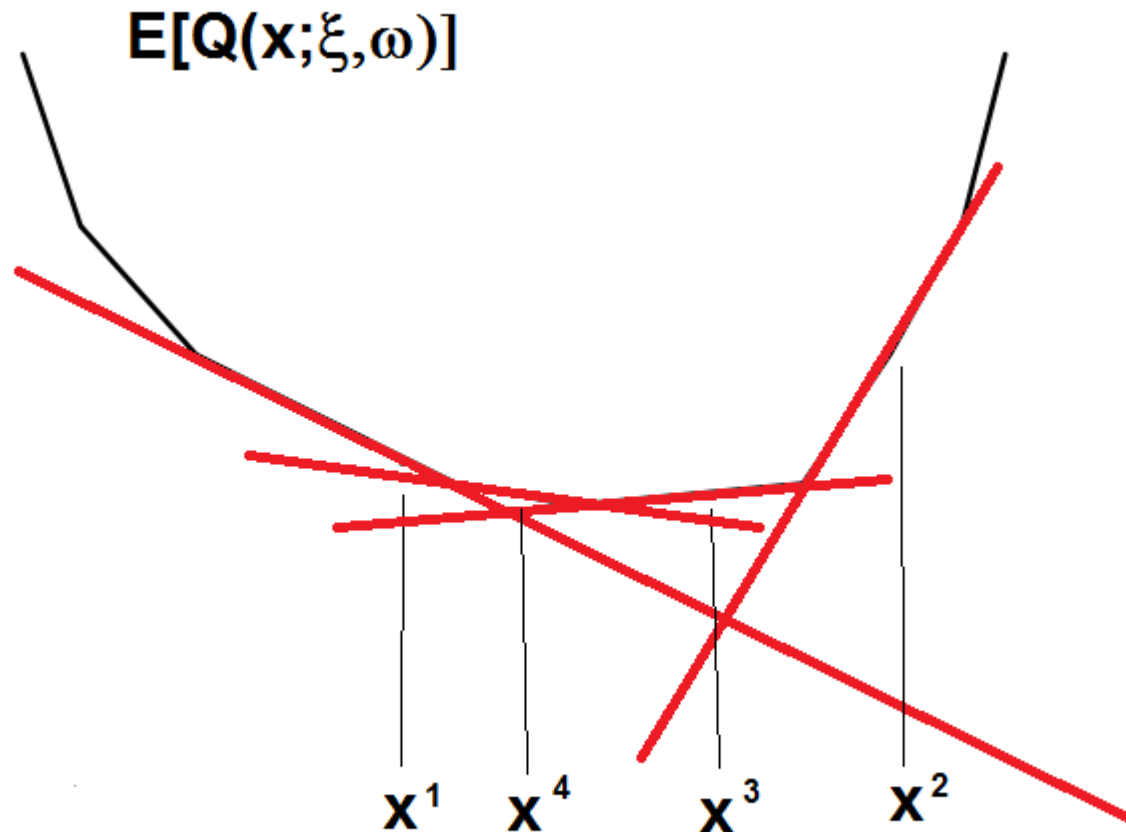
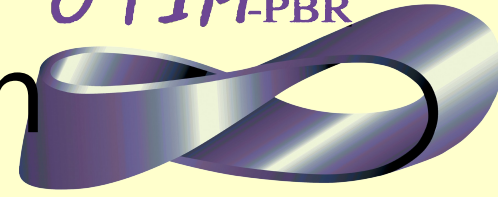
# Cutting-plane approximation



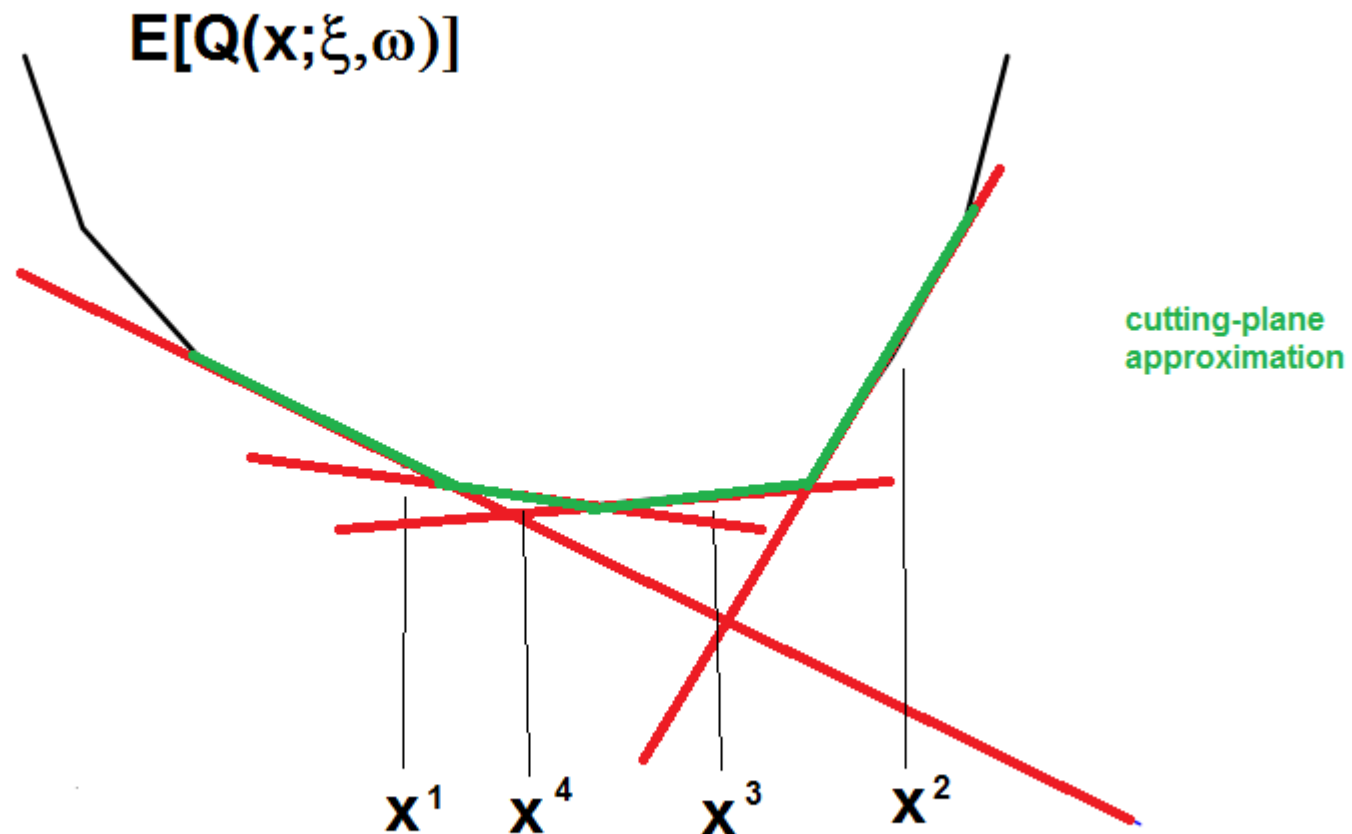
# Cutting-plane approximation



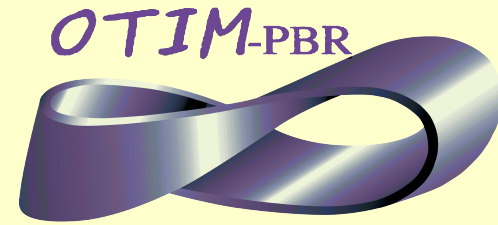
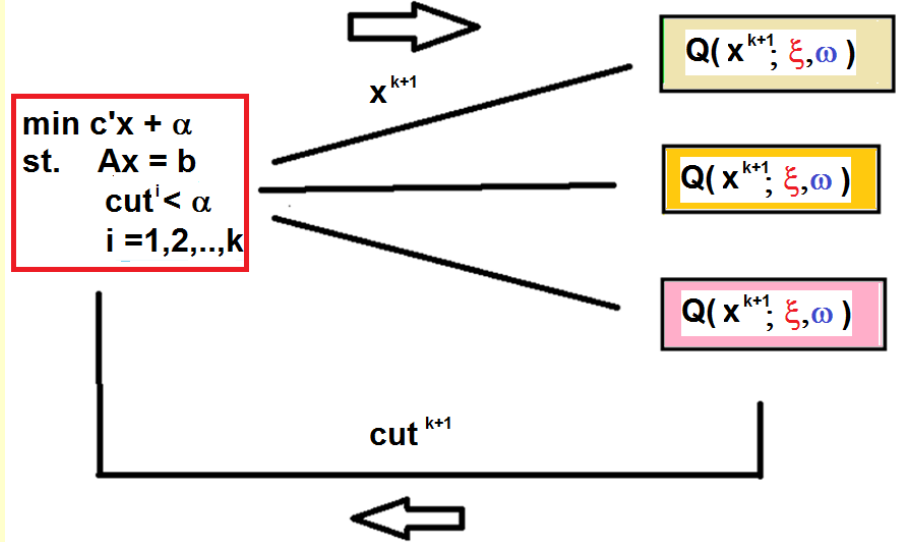
# Cutting-plane approximation



# Cutting-plane approximation







- $N$  linear programming problems must be solved for every first-stage decision

$$Q(x; \xi, \omega) \begin{cases} \min_y & q(\xi)^\top y \\ \text{s.t.} & Wy = h - Tx \\ & \underline{y} \leq y \leq \bar{y} \\ & y_i = \bar{y}_i(\omega) \quad i \in I_V \end{cases}$$

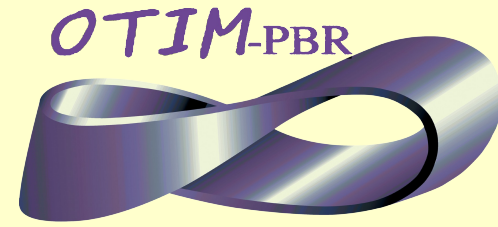
- This is a difficult task for large values of  $N$ 
  - We may solve the LPs in an approximate manner (inexact cuts)
  - Use more efficient cutting-plane methods, such as Bundle Methods:

Convex proximal bundle methods in depth: a unified analysis for inexact oracles.

Math. Programming, 2014, 148, 1-2, pp 241-277

W. de Oliveira, C Sagastizábal and C. Lemaréchal. 49

# PlanAb with chance-constraints



- One manner to prevent the number of scenarios  $N = |\Xi| * |\Omega|$  to be large is to handle the oil volume uncertainty by chance constraints

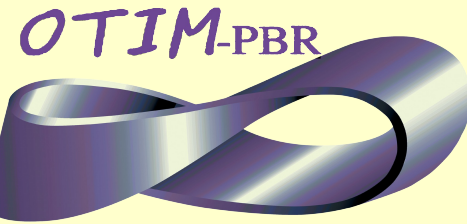
~~$$Q(x; \xi, \omega) = \begin{cases} \min_y & q(\xi)^\top y \\ \text{s.t.} & Wy = h - Tx \\ & \underline{y} \leq y \leq \bar{y} \\ & y_i = \bar{y}_i(\omega) \quad i \in I_V \end{cases}$$~~

$$Q(x; \xi) = \begin{cases} \min_y & q(\xi)^\top y \\ \text{s.t.} & Wy = h - Tx \\ & \underline{y} \leq y \leq \bar{y} \\ & \underline{y}_i^{LL} \leq y_i \leq \bar{y}_i^{LL} \quad i \in I_V \end{cases}$$

The bounds  $\underline{y}_i^{LL}, \bar{y}_i^{LL}$  are such that  $\mathbb{P}[\underline{y}_i^{LL} \leq y_i(\omega) \leq \bar{y}_i^{LL}] = 1 - p \quad (p \in (0, 1))$

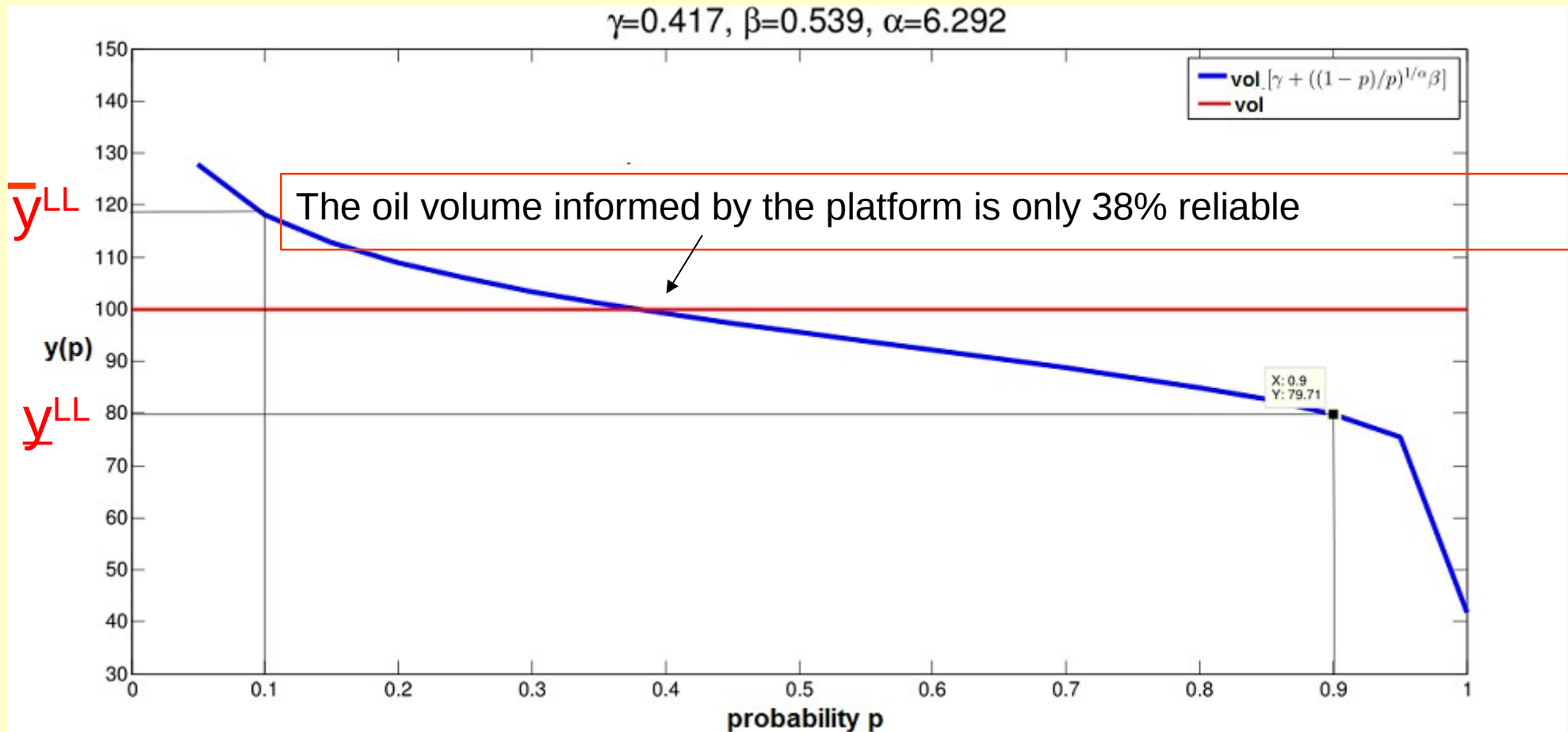
Determining the bounds is not a difficult task, since the oil volume of one platform is independent from the other platforms

# Confiability curve

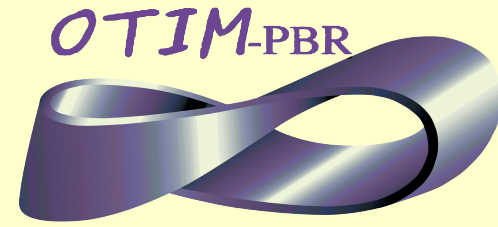


$$\frac{\bar{y}(\omega)}{\text{Vol}} \sim LL(\beta, \alpha, \gamma) \Rightarrow \mathbb{P} \left[ \frac{\bar{y}_i(\omega)}{\text{Vol}} \leq y \right] = \frac{(y - \gamma)^\alpha}{\beta^\alpha + (y - \gamma)^\alpha}, \text{ for } y \geq \gamma$$

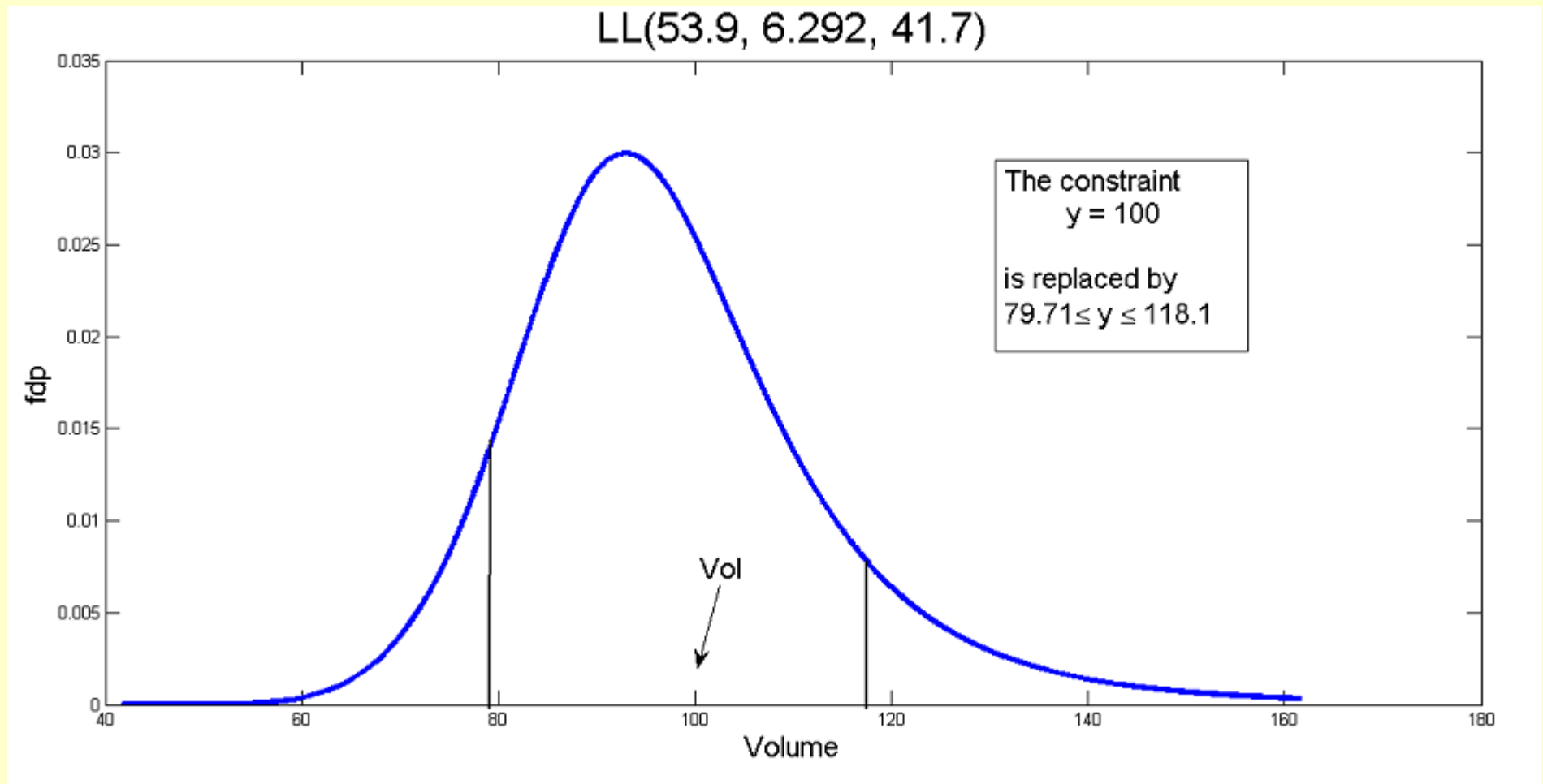
$$\mathbb{P} \left[ \frac{\bar{y}_i(\omega)}{\text{Vol}} \leq y \right] = 1 - p \Rightarrow y(p) = \text{Vol} \left[ \gamma \left( \frac{1-p}{p} \right)^{1/\alpha} \beta \right]$$



# PlanAb with chance-constraints

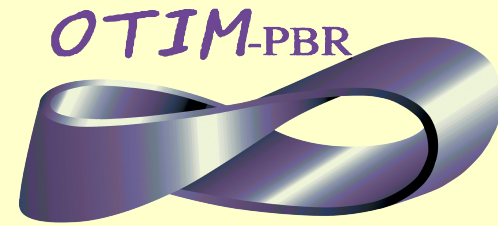


$$\bar{y}_i(\omega) \sim LL(\beta, \alpha, \gamma) \Rightarrow \mathbb{P}[\bar{y}_i(\omega) \leq y] = \frac{(y - \gamma)^\alpha}{\beta^\alpha + (y - \gamma)^\alpha}, \text{ for } y \geq \gamma$$



# Two-stage decomposition

Stochastic PlanAb + chance constraint



- Consider finitely many scenarios of **prices**
- The problem is decomposed into two decision levels

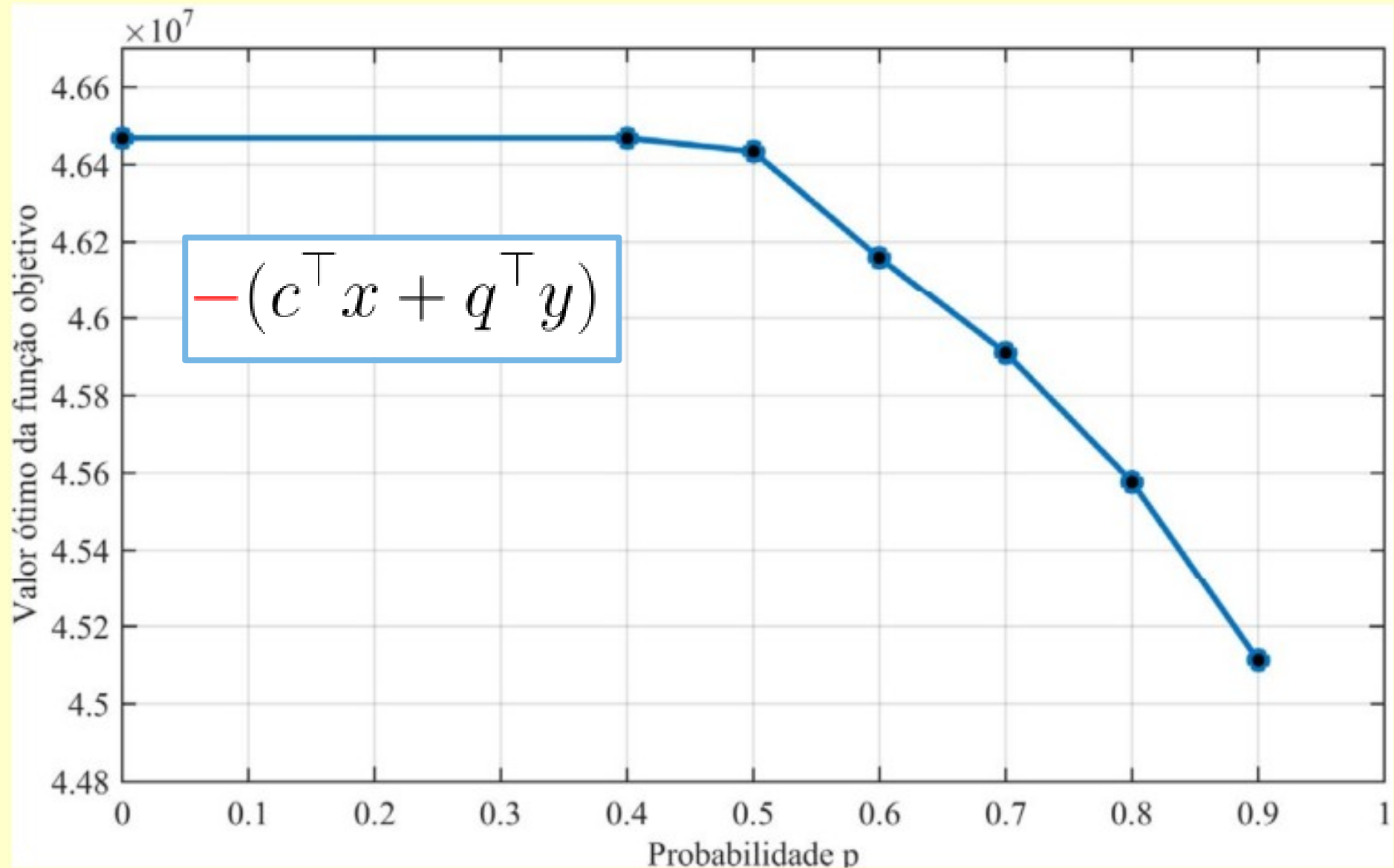
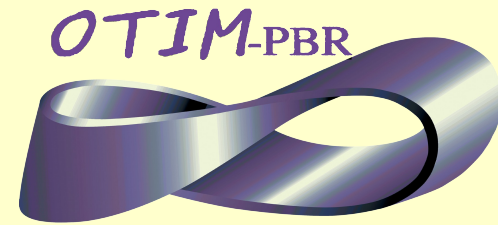
$$\begin{cases} \min_x c^\top x + \sum_{i=1}^N p_i [Q(x; \xi_i)] \\ \text{s.t.} & Ax = b \\ & \underline{x} \leq x \leq \bar{x} \end{cases}$$

with

$$Q(x; \xi) = \begin{cases} \min_y q(\xi)^\top y \\ \text{s.t.} & Wy = h - T x \\ & \underline{y} \leq y \leq \bar{y} \\ & \underline{y}_i^{LL} \leq y_i \leq \bar{y}_i^{LL} \quad i \in I_V \end{cases}$$

The bounds  $\underline{y}_i^{LL}, \bar{y}_i^{LL}$  are such that  $\mathbb{P}[\underline{y}_i^{LL} \leq y_i(\omega) \leq \bar{y}_i^{LL}] = 1 - p \quad (p \in (0, 1))$

# PlanAb + chance constraint



# Conclusions

- Stochastic PlanAb
  - Price scenarios
  - Oil volume
    - can be modelled either by using scenarios or chance-constraints
    - follow independent log-logistic probability distributions
- The computational implementation of the stochastic PlanAb model, with price scenarios and chance-constraints for oil volumes is in progress



Good bye  
and thank you for coming

