A new strategy to start Interior Point Methods for Stochastic Linear Programming problems

Marco Colombo Jacek Gondzio Andreas Grothey

School of Mathematics
The University of Edinburgh

Apmod 2006

What is stochastic programming

- Model uncertainty through the analysis of possible future scenarios
- Alternating sequence of decisions and random realisations
- Robust decision making

Assumptions for this talk:

- We do not have knowledge on the stochastic processes governing the uncertainties
- Stochastic problems are given in the SMPS format

SMPS format

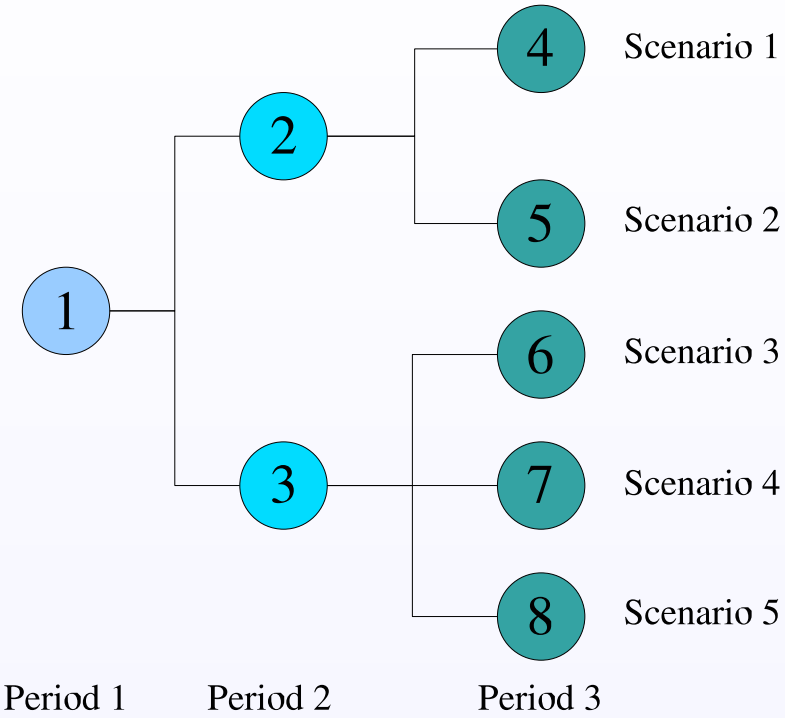
Standard formulation of multistage stochastic programs.

A problem in SMPS format is defined through 3 files:

- Core file: underlying deterministic problem in MPS format;
- Time file: information about the breaking up in stages;
- Stoch file: list of variations to the core data for each scenario.

The SMPS format provides all information we need about the structure of the problem.

Event tree



To each node of the tree we associate:

- a set of constraints
- an objective function
- the conditional probability of visit from the parent node

Notation

t period

l_t index of a node of period t

$a(l_t)$ ancestor of node l_t

n^{l_t} node data: $\{T^{l_t}, W^{l_t}, h^{l_t}, q^{l_t}, p^{l_t}\}$

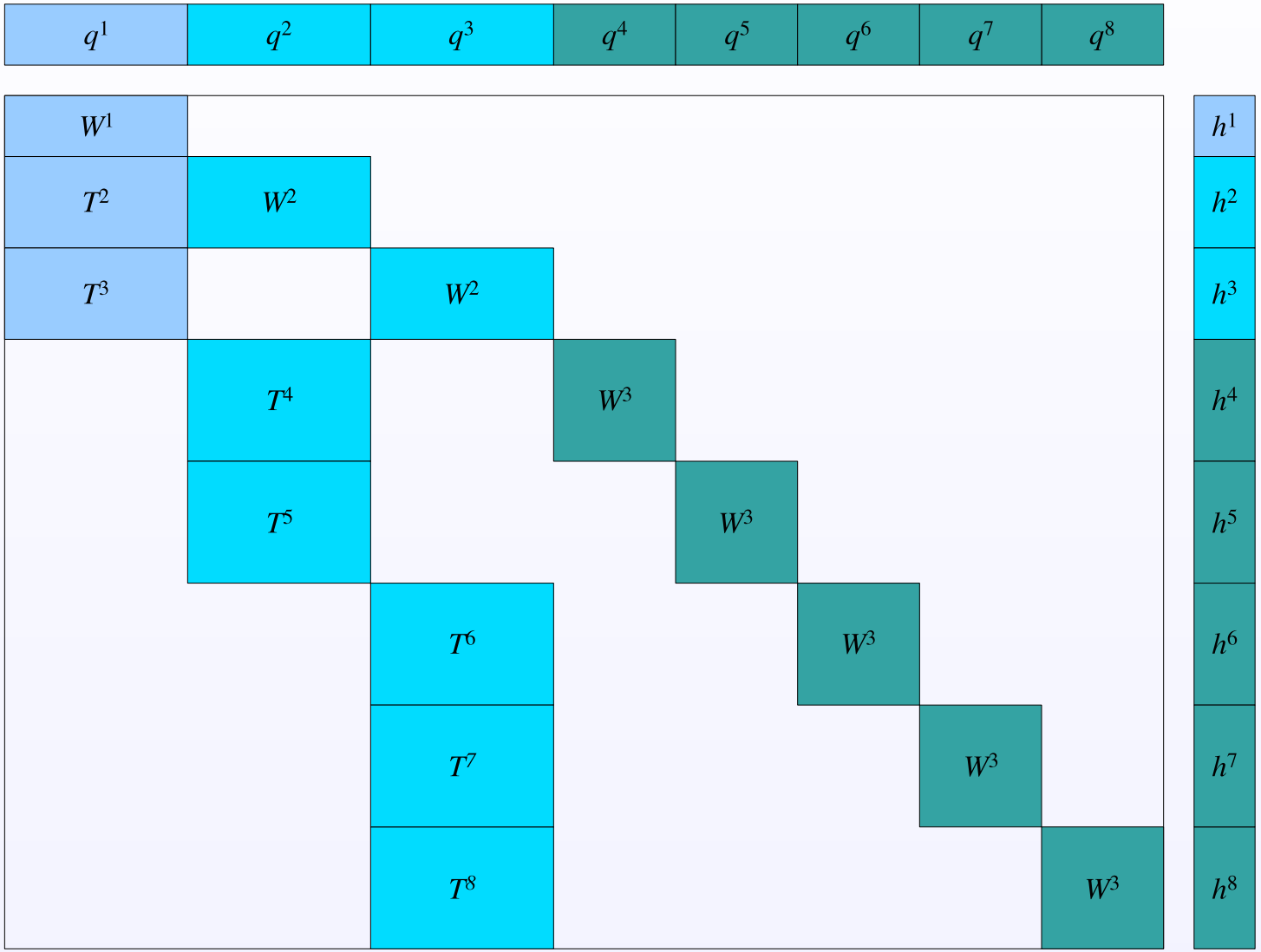
Model of the dynamics of the system (at node l_t):

$$\begin{aligned} \min \quad & \sum_{l_t} p^{l_t} (q^{l_t})^T x^{l_t} \\ \text{s.t.} \quad & T^{l_t} x^{a(l_t)} + W^{l_t} x^{l_t} = h^{l_t} \\ & x^{l_t} \geq 0 \end{aligned}$$

Complete deterministic equivalent formulation

$$
\begin{aligned}
\min \quad & (q^{l_1})^T x^{l_1} + \sum_{l_2=L_1+1}^{L_2} p^{l_2} (q^{l_2})^T x^{l_2} + \ldots + \sum_{l_T=L_{T-1}+1}^{L_T} p^{l_T} (q^{l_T})^T x^{l_T} \\
\text{s.t.} \quad & W^{l_1} x^{l_1} = h^{l_1}, \\
& T^{l_2} x^1 + W^{l_2} x^{l_2} = h^{l_2}, \qquad l_2 = L_1+1, \ldots, L_2, \\
& \vdots \qquad\qquad \vdots \\
& T^{l_T} x^{a(l_T)} + W^{l_T} x^{l_T} = h^{l_T}, \qquad l_T = L_{T-1}+1, \ldots, L_T, \\
& x^{l_t} \geq 0, \qquad l_t = 1, \ldots, L_T.
\end{aligned}
$$

Deterministic equivalent



Issues with the deterministic equivalent approach

The deterministic equivalent formulation produces problems of extremely large size

Example: fxm

Core matrix	330x600	2,755 nonzeros
3 stages, 6 nodes:	6,200x12,625	57,722 nonzeros
4 stages, 16 nodes:	386,940x822,855	4,823,612 nonzeros

- A detailed description produces robust decisions
- Detailed event trees can be very large
- The dimensions involved explode

The way forward

However:

- Interior point solvers are available in the community (CPLEX Barrier, PCx, HOPDM, etc.)
- Competitiveness of IPM grows with the problem size
- Parallel implementations are possible

And:

- Structure-exploiting (parallel) software (OOPS)
- Exploiting the stochastic structure to warm-start the IPMs

Primal–dual interior point methods

Basic structure of an IPM iteration

- Given an iterate (x, y, s) for which $(x, s) > 0$
- Solve the perturbed KKT conditions with Newton's method

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} b - Ax \\ c - A^T y - s \\ -XSe + \mu e \end{bmatrix}$$

- Move to the next point with stepsize α such that

$$(x + \alpha \Delta x, s + \alpha \Delta s) > 0$$

Considerations

- Keep centrality

$$XSe \approx \mu e$$

Considerations

- Keep centrality

$$XSe \approx \mu e$$

- Keep infeasibilities under control

Considerations

- Keep centrality

$$XSe \approx \mu e$$

- Keep infeasibilities under control
- How to find the starting point?
  - Mehrotra's starting point heuristic
  - Use of warm-start strategies

Warm-start issues

More difficult to implement successfully than with the simplex

Warm-start issues

More difficult to implement successfully than with the simplex

- "Contradictory" requirements:
  - Point should be close to the solution
  - Point should be away from the boundary

Warm-start issues

More difficult to implement successfully than with the simplex

- "Contradictory" requirements:
  - Point should be close to the solution
  - Point should be away from the boundary

Current attempts:

- Store an "advanced" iterate (3–4 digits of accuracy)
- Take special care of centrality
- Restore primal and dual feasibility with independent directions

Reduced event tree

Observation:

Very detailed event trees provide a fine-grained solution to a problem that could have been solved more coarsely with a much smaller tree.

Idea:

Use the solution to a smaller instance of the problem to generate a warm-start point

Reduced tree generation

Two complementary strategies:

1. Span the breadth of the tree
   - Choose some of the nodes at stage k (where k is small)
   - Choose all their ancestors up to the root node

Reduced tree generation

Two complementary strategies:

1. Span the breadth of the tree
   - Choose some of the nodes at stage k (where k is small)
   - Choose all their ancestors up to the root node
2. Choose the most representative scenario in each subtree
   - Define a "scenario distance"
   - Minimize the distance from an average scenario

Scenario distance

Distance between two nodes at period t:

$$d(n^{i_t}, n^{j_t}) = \|T^{i_t} - T^{j_t}\| + \|W^{i_t} - W^{j_t}\| + \|h^{i_t} - h^{j_t}\| + \|q^{i_t} - q^{j_t}\|$$

Scenario distance

Distance between two nodes at period t:

$$d(n^{i_t}, n^{j_t}) = \|T^{i_t} - T^{j_t}\| + \|W^{i_t} - W^{j_t}\| + \|h^{i_t} - h^{j_t}\| + \|q^{i_t} - q^{j_t}\|$$

Distance between two scenarios:

$$D(s_i, s_j) = \sum_{t=1}^{T} d(n^{i_t}, n^{j_t}), \quad i_t \in s_i, j_t \in s_j$$

Representative scenario

In each subtree S:

1. Create an artificial node $\bar{n}^t$

$$\bar{n}^t = \frac{1}{|S_t|} \sum_{l_t \in S_t} (T^{l_t}, W^{l_t}, h^{l_t}, q^{l_t})$$

(S_t: set of nodes of stage t in subtree S)

Representative scenario

In each subtree S:

1. Create an artificial node $\bar{n}^t$

$$\bar{n}^t = \frac{1}{|S_t|} \sum_{l_t \in S_t} (T^{l_t}, W^{l_t}, h^{l_t}, q^{l_t})$$

(S_t: set of nodes of stage t in subtree S)

2. Create an artificial scenario $\bar{s}$ as a sequence of artificial nodes

$$\bar{n}^1, \ldots, \bar{n}^T$$

Representative scenario

In each subtree S:

1. Create an artificial node $\bar{n}^t$

$$\bar{n}^t = \frac{1}{|S_t|} \sum_{l_t \in S_t} (T^{l_t}, W^{l_t}, h^{l_t}, q^{l_t})$$

(S_t: set of nodes of stage t in subtree S)

2. Create an artificial scenario $\bar{s}$ as a sequence of artificial nodes

$$\bar{n}^1, \ldots, \bar{n}^T$$

3. Representative scenario s^* is the one that minimizes the weighted distance from the artificial scenario

$$s^* = s_k, \quad k = \arg\min_{i \in S} (1 - p_i) D(s_i, \bar{s}).$$

Main steps of the algorithm

Exploit the structure of the stochastic program

- Find a reduced event tree

Main steps of the algorithm

Exploit the structure of the stochastic program

- Find a reduced event tree
- Solve the corresponding deterministic equivalent with a large tolerance

Main steps of the algorithm

Exploit the structure of the stochastic program

- Find a reduced event tree
- Solve the corresponding deterministic equivalent with a large tolerance
- Use this advanced point to generate a warm-start iterate for the complete problem

Main steps of the algorithm

Exploit the structure of the stochastic program

- Find a reduced event tree
- Solve the corresponding deterministic equivalent with a large tolerance
- Use this advanced point to generate a warm-start iterate for the complete problem
- Solve the complete problem

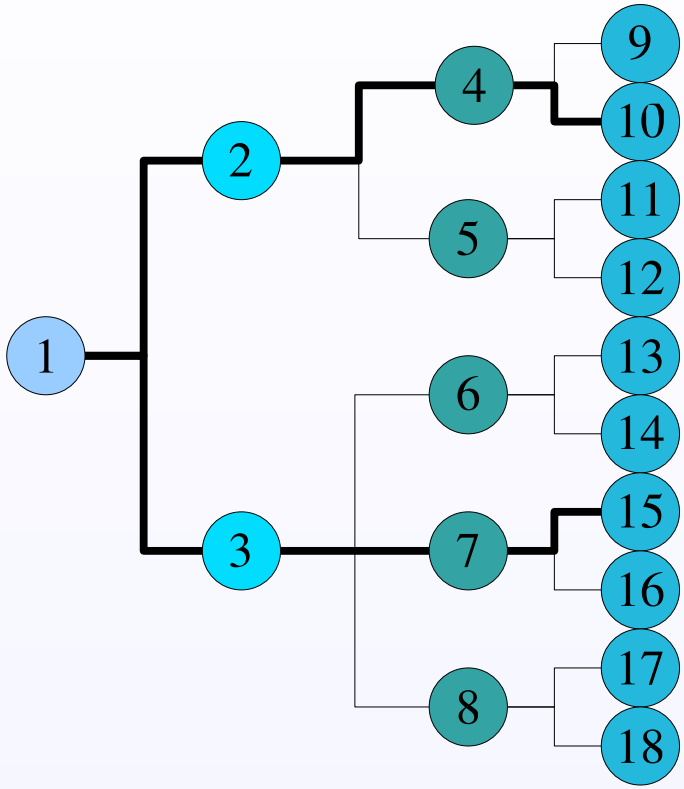
Main steps of the algorithm

Exploit the structure of the stochastic program

- Find a reduced event tree
- Solve the corresponding deterministic equivalent with a large tolerance
- Use this advanced point to generate a warm-start iterate for the complete problem
- Solve the complete problem

Good news: The reduced problem is very easy to solve!

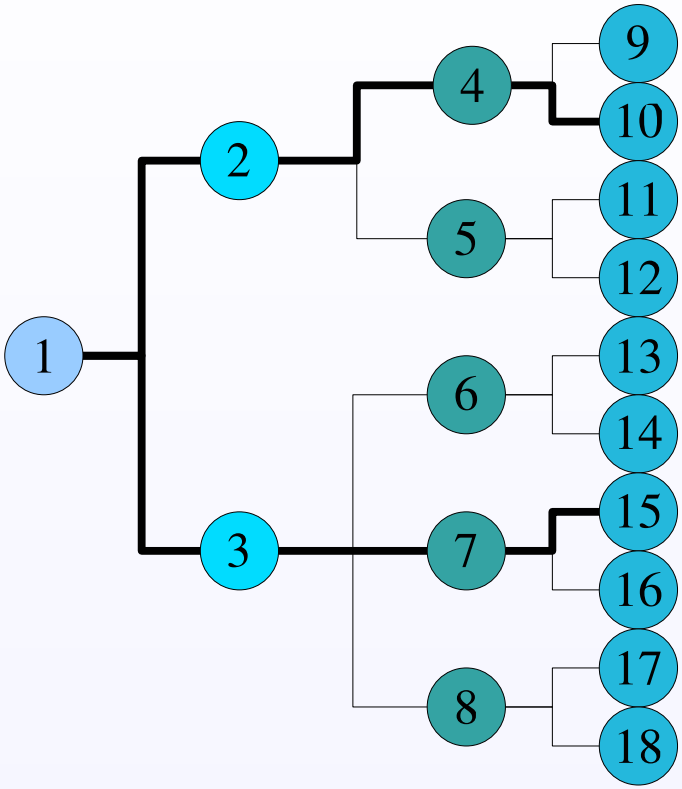
Construction of the warm-start iterate



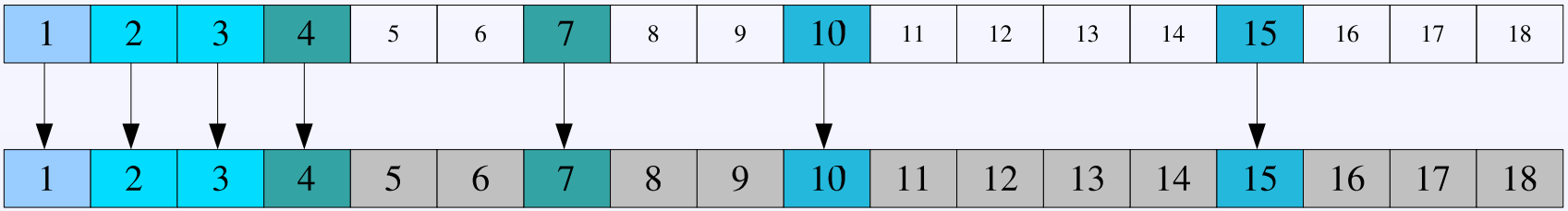
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

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

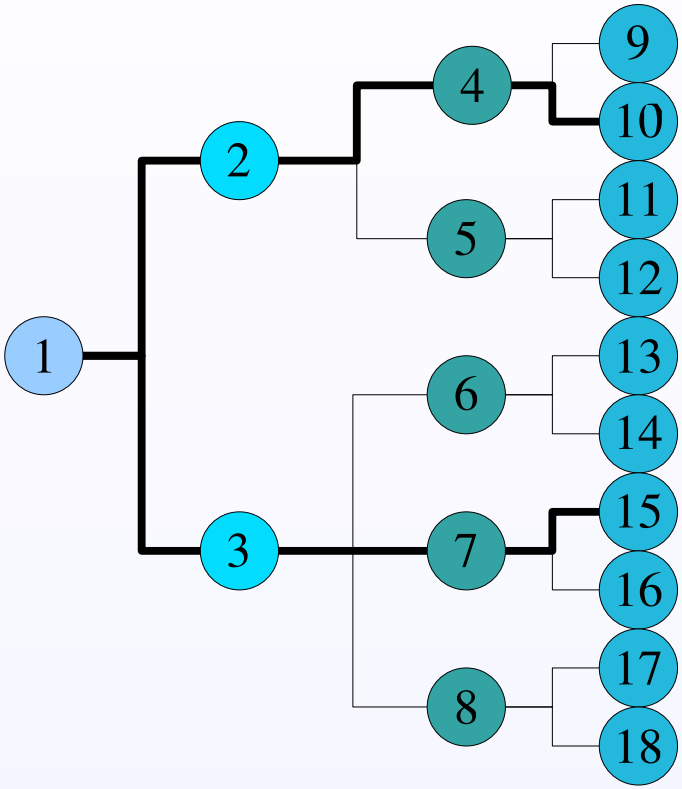
Construction of the warm-start iterate



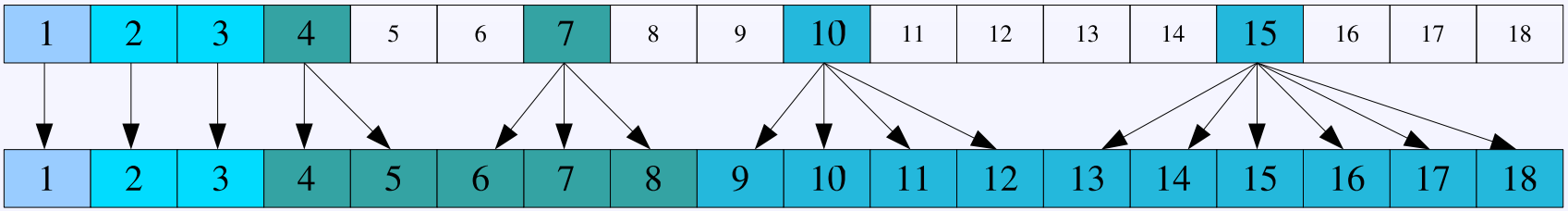
- Nodes in the reduced tree: provide already the wanted solution



Construction of the warm-start iterate



- Nodes in the reduced tree: provide already the wanted solution
- Remaining nodes: find the closest reduced-tree node of the same period and copy its solution



Computational experience

- Results from a collection of standard SMPS problems solved with HOPDM
- Presolve and scaling disabled
- 2 scenarios in the reduced tree
- Optimality tolerance for the reduced problem: 5.0×10^{-1}
- Optimality tolerance for the complete problem: 5.0×10^{-8}

Results with HOPDM (2 scenarios)

Problem data			Cold start		Warm start	
Name	Stgs	Scens	Iters	Seconds	Iters	Seconds
fxm2_16	2	16	22	1.2	14	1.1
fxm3_6	3	36	30	1.5	17	1.3
fxm3_16	3	256	40	31.1	20	20.7
fxm4_6	4	216	30	8.2	22	8.3
fxm4_16	4	4096	41	218.3	27	182.6
pltexpA3_16	3	256	26	153.8	14	87.8
pltexpA4_6	4	216	36	55.8	16	27.5
pltexpA5_6	5	1296	81	772.0	30	311.5
storm27	2	27	47	127.1	26	73.4
storm125	2	125	110	182.3	43	80.0
storm1000	2	1000	EXC	>1762.5	38	661.5
stocfor2	2	64	19	0.9	18	1.1
stocfor3	7	512	29	4.0	24	4.2

Conclusions

- Reduced tree solutions contain valuable information to construct a good warm-start iterate
- In this case, interior point methods can be used successfully in warm-start situations
- Exploiting the structure gives once again an additional advantage