School of Mathematics

Unblocking Heuristics for Warmstarting Interior Point Methods

Andreas Grothey, Jacek Gondzio, Marco Colombo

Warmstarting Interior Point Methods

- Many applications require the solution of a series of optimization problems (SQP, B&B, MPC, etc)
- Simplex/Active Set Methods are easy and efficient to warmstart
- IPM Warmstart is seen as difficult to impossible
- One of the main arguments against IPM

Warmstarting Interior Point Methods

- Many applications require the solution of a series of optimization problems (SQP, B&B, MPC, etc)
- Simplex/Active Set Methods are easy and efficient to warmstart
- IPM Warmstart is seen as difficult to impossible
- One of the main arguments against IPM

Some research (among others)

- Mitchell, Todd '92
- Hippolito '93
- Lustig, Marsten, Shanno '94
- Gondzio '98
- Gondzio, Vial '99
- Yıldırım, Wright '02
- Gondzio, G. '03
- John,Yıldırım '06
- Benson,Shanno '06

⇒ Aim:

- Warmstarting IPMs is possible!

Warmstarting Interior Point Methods: Overview

- Interior Point Methods
- Warmstarting
  - Modification Steps
  - Unblocking Strategies
  - Numerical Results
- Towards Structured Warmstarts
  - Example: Stochastic Programming
  - Results

Interior Point Methods (for QP)

$$\min c^\top x + \tfrac{1}{2}x^\top Qx \text{ s.t. } \begin{array}{rcl} Ax &=& b \\ x &\geq& 0 \end{array} \tag{QP}$$

Optimality conditions:

$$\begin{array}{rcl} c + Qx - A^\top y - z &=& 0 \\ Ax &=& b \\ XZe &=& 0 \quad (\mu e) \\ x, z &\geq& 0 \end{array}$$

$\Rightarrow$ Newton Step:

$$\begin{bmatrix} -Q & A^\top & I \\ A & 0 & 0 \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} \xi_c \\ \xi_b \\ r_{xz} \end{bmatrix} = \begin{bmatrix} c + Qx - A^\top y - z \\ b - Ax \\ \mu e - XZe \end{bmatrix} \tag{NS-QP}$$

where $X = \text{diag}(x), \quad Z = \text{diag}(z)$

Interior Point Methods II

- choose $x_0, y_0, z_0 > 0$
- solve (NS-QP) for $\Delta x, \Delta z, \Delta y$
- compute stepsizes

$$\begin{aligned} \alpha_P &= \max_{\alpha > 0}\{\alpha : x + \alpha\Delta x \geq 0\} \\ \alpha_D &= \max_{\alpha > 0}\{\alpha : z + \alpha\Delta z \geq 0\} \end{aligned}$$

- take step

$$\begin{aligned} x_+ &= x + 0.995\alpha_P\Delta x \\ y_+ &= y + 0.995\alpha_D\Delta y \\ z_+ &= z + 0.995\alpha_D\Delta z \end{aligned}$$

- update μ:

$$\mu_+ = \sigma\frac{x_+^\top z_+}{n}, \quad 0 < \sigma < 1$$

Corrector steps:

- Given: *Predictor* Step $(\Delta x_p, \Delta y_p, \Delta z_p)$
- Solve:

$$\begin{bmatrix} -Q & A^\top & I \\ A & 0 & 0 \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x_c \\ \Delta y_c \\ \Delta z_c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \Delta x_p^\top \Delta z_p \end{bmatrix} \quad \left(= \begin{bmatrix} \xi_c \\ \xi_b \\ \mu e - XZe \end{bmatrix} \right)$$

  Note that $\Delta x_p^\top \Delta z_p = (x + \Delta x_p)^\top (z + \Delta z_p) - \mu e$
  (= residual in $\mu e - XZe = 0$ after predictor step is taken)

- Use search direction $\Delta x = \Delta x_p + \Delta x_c$

$$\Delta z = \Delta z_p + \Delta z_c$$

- Repeat (as long as improvement)

Warmstarting Interior Point Methods

Aim: Use information from solution process of

$$\min c^\top x + \tfrac{1}{2}x^\top Q x \quad \text{s.t.} \quad \begin{aligned} Ax &= b \\ x &\geq 0 \end{aligned} \tag{QP}$$

to construct a starting point for (nearby problem)

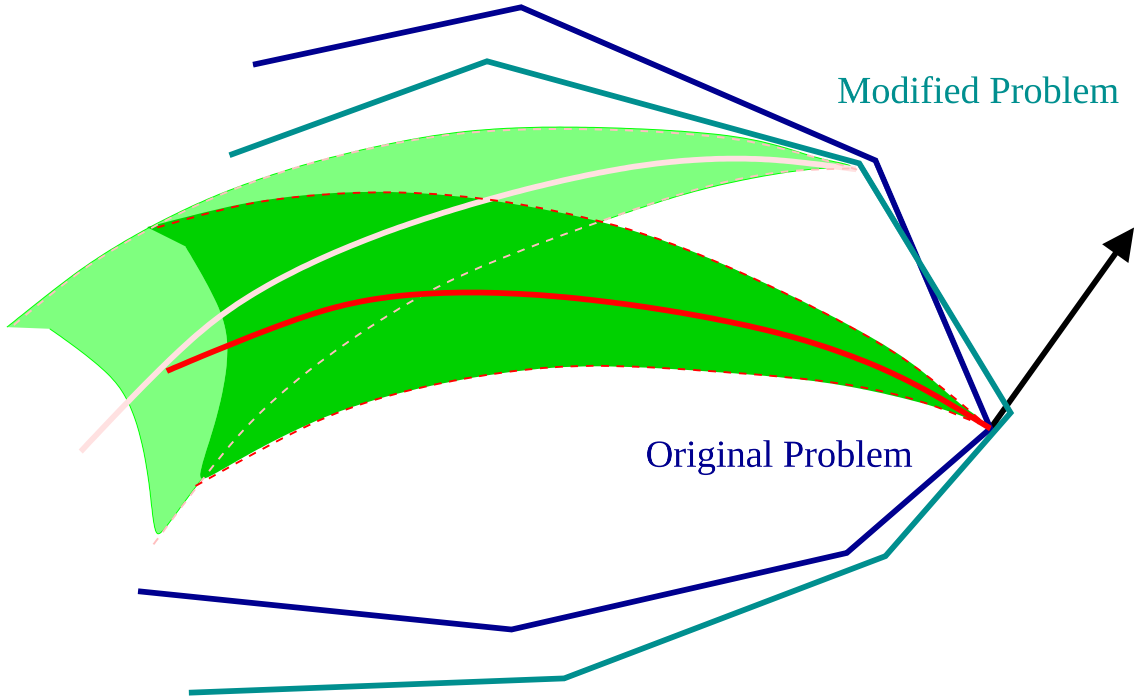
$$\min \tilde{c}^\top x + \tfrac{1}{2}x^\top \tilde{Q} x \quad \text{s.t.} \quad \begin{aligned} \tilde{A}x &= \tilde{b} \\ x &\geq 0 \end{aligned} \tag{$\widetilde{\text{QP}}$}$$

where $\tilde{A} \approx A, \tilde{Q} \approx Q, \tilde{b} \approx b, \tilde{c} \approx c$

- It is **not** a good idea to use the solution of (QP) to start ($\widetilde{\text{QP}}$).
- *Unlike for the Simplex/Active Set Method!*

Why?

Hippolito (1993): Search direction is parallel to nearby constraints

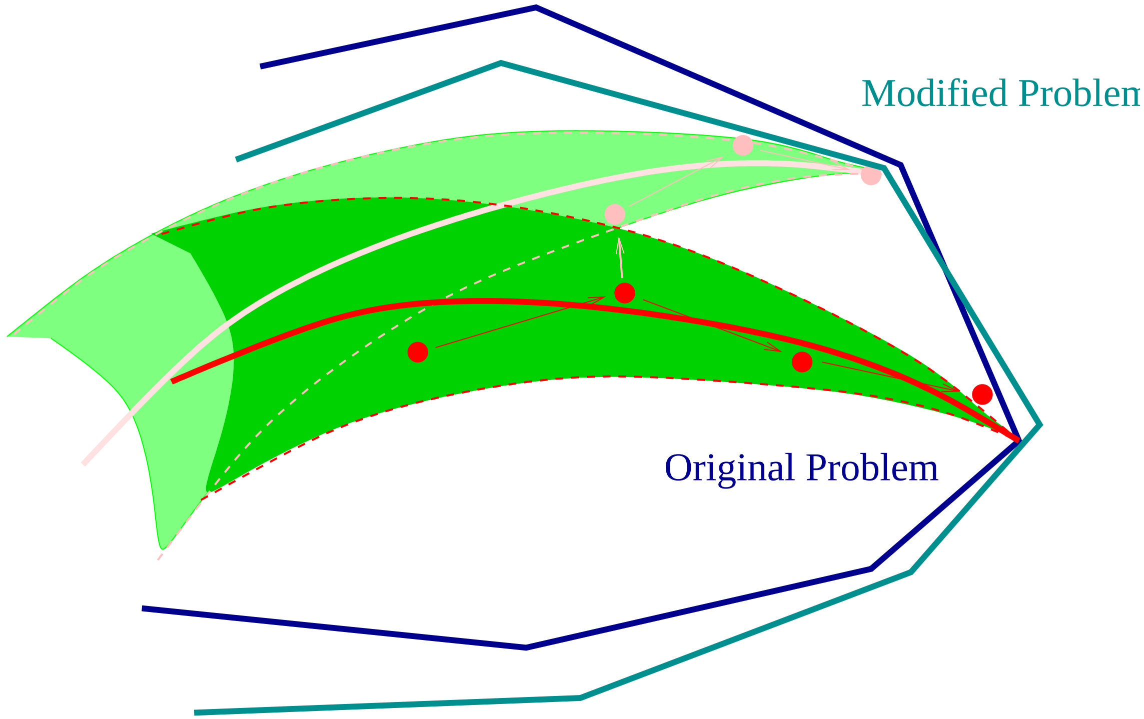


⇒ only small step in search direction can be taken

⇒ *blocking* of step

Warmstarting Heuristics

Idea: Start close to the (new) central path, not close to the (old) solution



$\Rightarrow$ Start from a previous iterate and do additional *modification* step.

$\Rightarrow$ *Blocking* might still be a problem.

Warmstarting Strategies

Three components:

- Choose a $\hat{\mu}$-**iterate** from previous solve
  Central ($x_i z_i \approx \hat{\mu}$),
  but **not** *primal/dual feasible w.r.t. new problem.*
- Do a **modification** step
  To gain primal/dual feasibility in new problem (might loose centrality).
- **Unblocking** strategy
  To recover from bad starting point (regain centrality)

Questions:

- How to choose $\hat{\mu}$?
- What is an appropriate modificator step?
- What to do about (un)blocking?

Warmstarting IPM: Theoretical Guidance

Solution of Newton System can be interpreted as projection:

$\Rightarrow$ Condition for non-blocking step (full step $\alpha_P = \alpha_D = 1$ is feasible):

$$-X^{1/2}Z^{1/2}e \leq P_{AD^{-1/2}\Delta x=\xi_b}(D^{-1/2}(X^{-1}r_{xz} - \xi_c)) \leq \mu X^{-1/2}Z^{-1/2}e$$

Warmstarting IPM: Theoretical Guidance

Solution of Newton System can be interpreted as projection:

$\Rightarrow$ Condition for non-blocking step (full step $\alpha_P = \alpha_D = 1$ is feasible):

$$-X^{1/2}Z^{1/2}e \leq P_{AD^{-1/2}\Delta x=\xi_b}(D^{-1/2}(X^{-1}r_{xz} - \xi_c)) \leq \mu X^{-1/2}Z^{-1/2}e$$

Step is non-blocking

- Current point (x, z) is well centered (r_{xz} small).

Warmstarting IPM: Theoretical Guidance

Solution of Newton System can be interpreted as projection:

$\Rightarrow$ Condition for non-blocking step (full step $\alpha_P = \alpha_D = 1$ is feasible):

$$-X^{1/2}Z^{1/2}e \leq P_{AD^{-1/2}\Delta x=\xi_b}(D^{-1/2}(X^{-1}r_{xz} - \xi_c)) \leq \mu X^{-1/2}Z^{-1/2}e$$

Step is non-blocking

- Current point (x, z) is well centered (r_{xz} small).
- Primal/dual residuals (ξ_b, ξ_c) are small.

Warmstarting IPM: Theoretical Guidance

Solution of Newton System can be interpreted as projection:

$\Rightarrow$ Condition for non-blocking step (full step $\alpha_P = \alpha_D = 1$ is feasible):

$$-X^{1/2}Z^{1/2}e \leq P_{AD^{-1/2}\Delta x=\xi_b}(D^{-1/2}(X^{-1}r_{xz} - \xi_c)) \leq \mu X^{-1/2}Z^{-1/2}e$$

Step is non-blocking

- Current point (x, z) is well centered (r_{xz} small).
- Primal/dual residuals (ξ_b, ξ_c) are small.
- Larger μ allows for larger residuals (Note: $XZe = \mu e$ for centered point)

Warmstarting IPM: Theoretical Guidance

Solution of Newton System can be interpreted as projection:

$\Rightarrow$ Condition for non-blocking step (full step $\alpha_P = \alpha_D = 1$ is feasible):

$$-X^{1/2}Z^{1/2}e \leq P_{AD^{-1/2}\Delta x=\xi_b}(D^{-1/2}(X^{-1}r_{xz} - \xi_c)) \leq \mu X^{-1/2}Z^{-1/2}e$$

Step is non-blocking

- Current point (x, z) is well centered (r_{xz} small).
- Primal/dual residuals (ξ_b, ξ_c) are small.
- Larger μ allows for larger residuals (Note: $XZe = \mu e$ for centered point)

$\Rightarrow$ **Heuristics:**

- Retrace iterates of original problem until appropriate μ-center is found (large enough to absorb residuals ξ_b, ξ_c.
- Aim to reduce residuals ξ_b, ξ_c, r_{xz}.

Modification Steps

Proposed steps fall in two categories:

- Variations of the standard IPM step (NS-QP)
  - Newton Correction Step [Wright, Yıldırim '03]
  - Recentering Step/Splitting Directions [Gondzio, G. '03/'06]
- "Projection" of the current point onto the new primal/dual feasible set
  - Weighted Least Squares (WLS) step [Wright, Yıldırim '03]
  - z-Adjustment [Gondzio, G. '06]

Modification Steps (Variations of standard IPM step)

The standard IPM step
(for the new problem ($\widetilde{\text{QP}}$), from the warmstart point (x, y, z) with $\hat{\mu} = x^\top z/n$)
is:

$$\begin{bmatrix} -\tilde{Q} & \tilde{A}^\top & I \\ \tilde{A} & 0 & 0 \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} \xi_c \\ \xi_b \\ r_{xz} \end{bmatrix} = \begin{bmatrix} \tilde{c} - \tilde{A}^\top y - z \\ \tilde{b} - \tilde{A}x \\ \sigma\hat{\mu}e - XZe \end{bmatrix} \qquad \text{(NS-QP)}$$

⇒ **Newton Correction Step:**

$$\begin{bmatrix} -\tilde{Q} & \tilde{A}^\top & I \\ \tilde{A} & 0 & 0 \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} \xi_c \\ \xi_b \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{c} - \tilde{A}^\top y - z \\ \tilde{b} - \tilde{A}x \\ 0 \end{bmatrix} \qquad \text{(NCS)}$$

- Restore primal/dual feasibility. Don't change centrality.

Modification Steps (Variations of standard IPM step)

The standard IPM step
(for the new problem $(\widetilde{\text{QP}})$, from the warmstart point (x, y, z) with $\hat{\mu} = x^\top z/n$)
is:

$$\begin{bmatrix} -\tilde{Q} & \tilde{A}^\top & I \\ \tilde{A} & 0 & 0 \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} \xi_c \\ \xi_b \\ r_{xz} \end{bmatrix} = \begin{bmatrix} \tilde{c} - \tilde{A}^\top y - z \\ \tilde{b} - \tilde{A}x \\ \sigma\hat{\mu}e - XZe \end{bmatrix} \tag{NS-QP}$$

⇒ **Recentering Steps:**

$$\begin{bmatrix} -\tilde{Q} & \tilde{A}^\top & I \\ \tilde{A} & 0 & 0 \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} \tilde{\xi}_c \\ \tilde{\xi}_b \\ \tilde{r}_{xz} \end{bmatrix} = \begin{bmatrix} \tilde{c} - \tilde{A}^\top y - z \\ \tilde{b} - \tilde{A}x \\ \hat{\mu}e - XSe \end{bmatrix}$$

- Combined with additional centering step in the **original** problem before the warmstarting point is returned. ($\Rightarrow \hat{\mu}e - XSe \approx 0$)

⇒ Warmstart point should be more central (hopefully).

Unblocking Heuristics

- Modification step may not be feasible ($x, z > 0$ not satisfied after step)
- Modification step may worsen centrality

For these reasons the (modified) warmstarting point may still lead to small steps to be taken (blocking)

$\Rightarrow$ Rather than abandon the warmstart attempt, try **unblocking** heuristic.

- **Higher Order Correctors**
- Unblocking by **Sensitivity Analysis**

Higher Order Correctors as unblocking device

Typical situation:

- Modification step leads to primal/dual feasible point (hopefully) **but** may sacrifice centrality.

⇒ Centrality difficult to achive in one step

⇒ Bad starting point.

Idea:

- Relax requirement for full centrality
- Only correct for worst components
  - Easier to achieve
  - These are the only ones that really hurt

⇒ Project complementarity products onto *neighbourhood* of central path
⇒ And use this as target

Reminder: Corrector steps

- Given: *Predictor* Step $(\Delta x_p, \Delta y_p, \Delta z_p)$
- Solve:

$$\begin{bmatrix} -Q & A^\top & I \\ A & 0 & 0 \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x_c \\ \Delta y_c \\ \Delta z_c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \Delta x_p^\top \Delta z_p \end{bmatrix}$$

  Note that $\Delta x_p^\top \Delta z_p = (x + \Delta x_p)^\top (z + \Delta z_p) - \mu e$
  ($=$ residual in $\mu e - XZe = 0$ after predictor step is taken)

- Use search direction $\Delta x = \Delta x_p + \Delta x_c$
  $\Delta z = \Delta z_p + \Delta z_c$
- Repeat (as long as improvement)

Higher Order Correctors to Unblock Directions

$$\begin{bmatrix} -Q & A^\top & I \\ A & 0 & 0 \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x_c \\ \Delta y_c \\ \Delta z_c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ m(\Delta x_p^\top \Delta z_p) \end{bmatrix}$$

$$m(\Delta x_{p,i} \Delta z_{p,i}) = \begin{cases} \Delta x_{p,i} \Delta z_{p,i} & (x_i + \Delta x_{p,i})(z_i + \Delta z_{p,i}) < \mu/10 \\ -5 * \mu & (x_i + \Delta x_{p,i})(z_i + \Delta z_{p,i}) > 10\mu \\ 0 & \text{else} \end{cases}$$

Conflicting goals:

- Attempt to reach central path
  - can take larger step at next iteration
  - central path is away from boundary (reduce blocking)
- But using large r_{xz} will lead to blocking step

Higher Order Correctors:

- Aim towards central path
- But do not force if adjustment to large

Sensitivity Analysis

Idea:

- Small change in current point (x, y, z) might lead to a large change in resulting direction
- Can we influence this?

$\Rightarrow$ Obtain derivatives of Primal/Dual direction w.r.t current point

Sensitivity Analysis

The step equation

$$\begin{bmatrix} -Q & A^\top & I \\ A & 0 & 0 \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} \xi_c \\ \xi_b \\ r_{xz} \end{bmatrix} = \begin{bmatrix} c - A^\top y - z \\ b - Ax \\ \mu e - XZe \end{bmatrix} \qquad \text{(NS-QP)}$$

Implies a functional relationship

$$(\Delta x, \Delta y, \Delta z) = F(x, y, z, \mu)$$

$\Rightarrow$ Can get derivatives (sensitivity information) $\frac{\partial \Delta x_i}{\partial x_j}, \frac{\partial \Delta z_i}{\partial x_j}, \frac{\partial \Delta x_i}{\partial z_j}, \frac{\partial \Delta z_i}{\partial z_j}$

Use this to unblock the step

- How to get derivatives ?
- How to use them to unblock the direction ?

Sensitivity Analysis: Derivatives

Differentiate the step equation (w.r.t. x_i) to obtain

$$\begin{bmatrix} -Q & A^\top & I \\ A & 0 & 0 \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \frac{\partial \Delta x}{\partial x_i} \\ \frac{\partial \Delta y}{\partial x_i} \\ \frac{\partial \Delta z}{\partial x_i} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_i}\xi_c \\ \frac{\partial}{\partial x_i}\xi_b \\ -\Delta Z e_i + \frac{\partial}{\partial x_i} r_{xz} \end{bmatrix} = \begin{bmatrix} Qe_i \\ -Ae_i \\ -Ze_i - \Delta Z e_i \end{bmatrix}$$

and similarly for $\frac{\partial}{\partial y_i}, \frac{\partial}{\partial z_i}$ gives

$$\begin{bmatrix} -Q & A^T & I \\ A & 0 & 0 \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \frac{\mathrm{d}\Delta x}{\mathrm{d}x} & \frac{\mathrm{d}\Delta x}{\mathrm{d}y} & \frac{\mathrm{d}\Delta x}{\mathrm{d}z} \\ \frac{\mathrm{d}\Delta y}{\mathrm{d}x} & \frac{\mathrm{d}\Delta y}{\mathrm{d}y} & \frac{\mathrm{d}\Delta y}{\mathrm{d}z} \\ \frac{\mathrm{d}\Delta z}{\mathrm{d}x} & \frac{\mathrm{d}\Delta z}{\mathrm{d}y} & \frac{\mathrm{d}\Delta z}{\mathrm{d}z} \end{bmatrix} = \begin{bmatrix} Q & -A^T & -I \\ -A & 0 & 0 \\ -Z - \Delta Z & 0 & -X - \Delta X \end{bmatrix}$$

$$= -\begin{bmatrix} -Q & A^T & I \\ A & 0 & 0 \\ Z & 0 & X \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \Delta Z & 0 & \Delta X \end{bmatrix}$$

If A has full row rank $\Rightarrow$ System matrix is invertible

Sensitivity Analysis: Derivatives

$$\begin{bmatrix} \frac{\mathrm{d}\Delta x}{\mathrm{d}x_i} \\ \frac{\mathrm{d}\Delta y}{\mathrm{d}x_i} \\ \frac{\mathrm{d}\Delta z}{\mathrm{d}x_i} \end{bmatrix} = \begin{bmatrix} -e_i \\ 0 \\ 0 \end{bmatrix} + \Delta z_i \begin{bmatrix} -Q & A^T & I \\ A & 0 & 0 \\ Z & 0 & X \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ -e_i \end{bmatrix}$$

$$\begin{bmatrix} \frac{\mathrm{d}\Delta x}{\mathrm{d}y_i} \\ \frac{\mathrm{d}\Delta y}{\mathrm{d}y_i} \\ \frac{\mathrm{d}\Delta z}{\mathrm{d}y_i} \end{bmatrix} = \begin{bmatrix} 0 \\ -e_i \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \frac{\mathrm{d}\Delta x}{\mathrm{d}z_i} \\ \frac{\mathrm{d}\Delta y}{\mathrm{d}z_i} \\ \frac{\mathrm{d}\Delta z}{\mathrm{d}z_i} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -e_i \end{bmatrix} + \Delta x_i \begin{bmatrix} -Q & A^T & I \\ A & 0 & 0 \\ Z & 0 & X \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ -e_i \end{bmatrix}$$

$\Rightarrow$ • System matrix is the same as for step equation

• one additional backsolve yields sensitivity for one x_i, z_i component.

? will this succeed in unblocking the step?

Sensitivity Analysis

Result:

Let

$$\mathcal{C}_l := \{(x,y,z) : x + \Delta x(x,z) >= -le, z + \Delta z(x,z) >= -le\}$$

Then $\exists l > 0 : \forall l \in C_l$ there exists an open nonempty set $\Omega_{l,u}$ such that

- $\Omega_{l,u} \subset B(0,u)$
- Every $(\mathrm{d}x, \mathrm{d}z) \in \Omega_{l,u}$ will unblock the step:

  $$x + \mathrm{d}x + \Delta x(x + \mathrm{d}x, z + \mathrm{d}z) >= 0, z + \mathrm{d}z + \Delta z(x + \mathrm{d}x, z + \mathrm{d}z) >= 0$$

Sensitivity Analysis: Unblocking

Computation of complete sensitivity is computationally expensive

⇒ **Heuristic:**

For every blocking component i of $\Delta x, \Delta z$:

- Get sensitivity for change of x_i, z_i
- Find *most efficient* unblocking change to current point (x_i, z_i) subject to *safeguards*:
  - No change outside $[x_i/10, 10x_i]$
  - No change larger than $\mathcal{O}(\max\{\|\xi_c\|, \|\xi_b\|\})$.
  - No introduction of new blocking components

Numerical Experiments: NETLIB

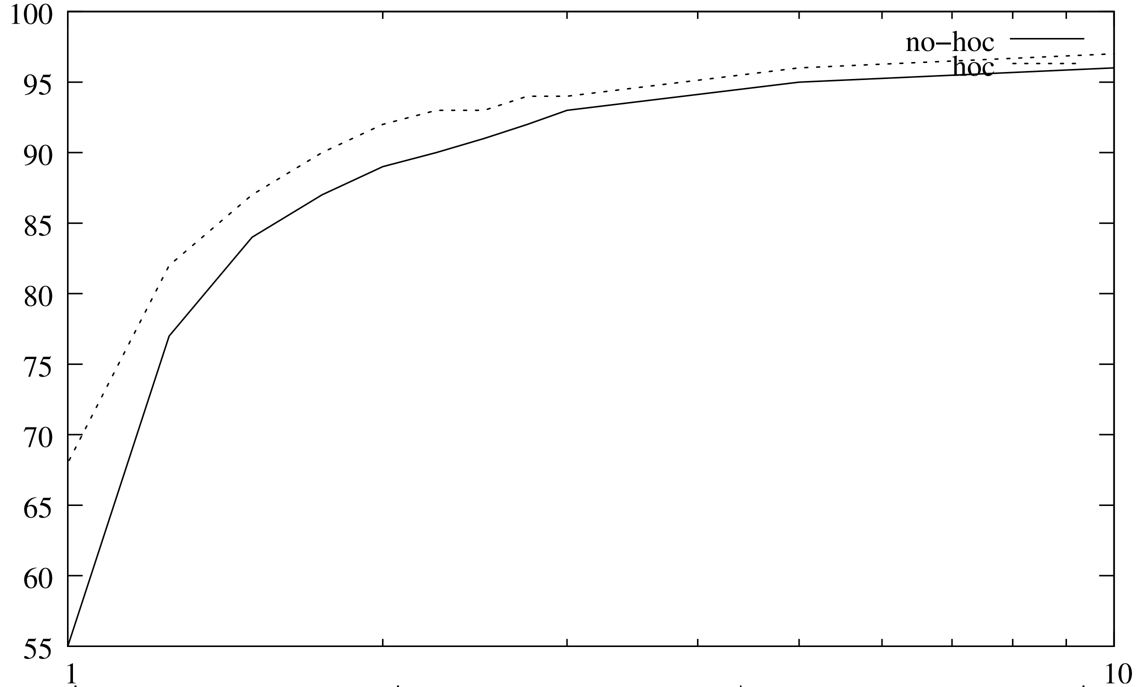
Testset proposed by Benson, Shanno '06:

- Take small to mid-scale instances of NETLIB LP library
- Randomly perturb problem data in b, c or A
- Perturb 10% (at most 20) of components on average.
- Perturb by $0.001, 0.01, 0.1$.

Our results:

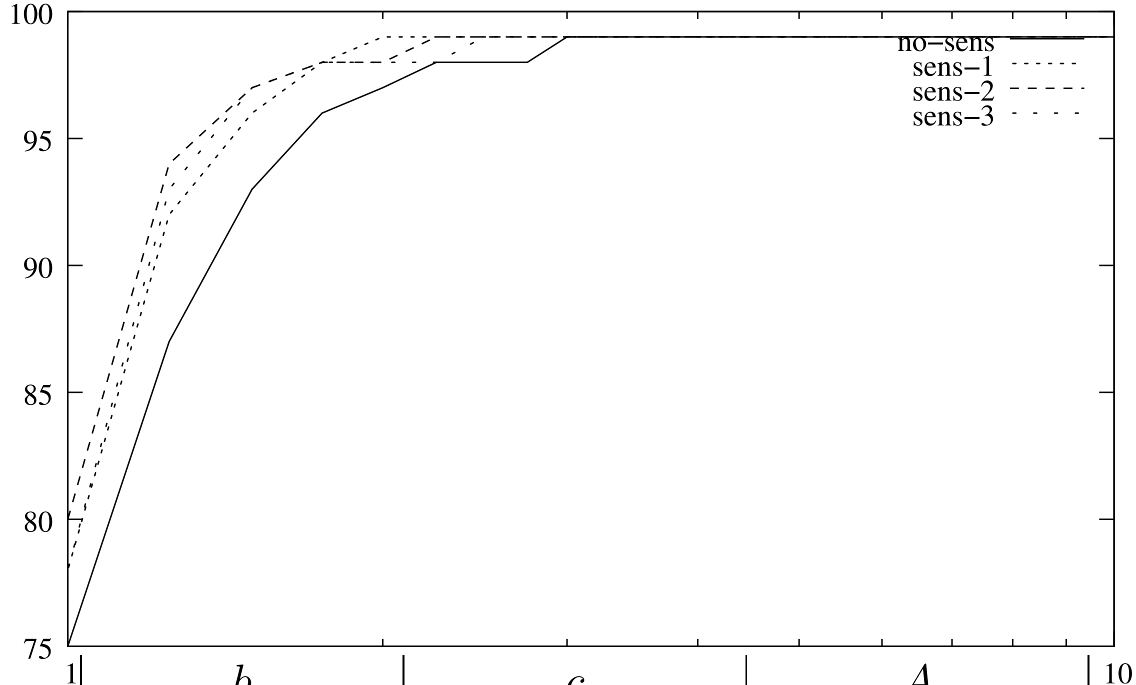
- Choose 10 random instances (and take average)
- All problems warmstarted with $\hat{\mu} = 10^{-2}$.
- All test run within the OOPS/HOPDM IPM solver

Results: Higher Order Correctors



	b			c			A			
	0.1	0.01	0.001	0.1	0.01	0.001	0.1	0.01	0.001	total
no-hoc	6.1	5.0	5.0	14.9	8.7	6.2	11.5	7.0	6.6	8.0
hoc	5.7	4.7	5.2	11.6	7.1	5.8	9.4	6.4	6.5	7.0

Results: Sensitivity Steps

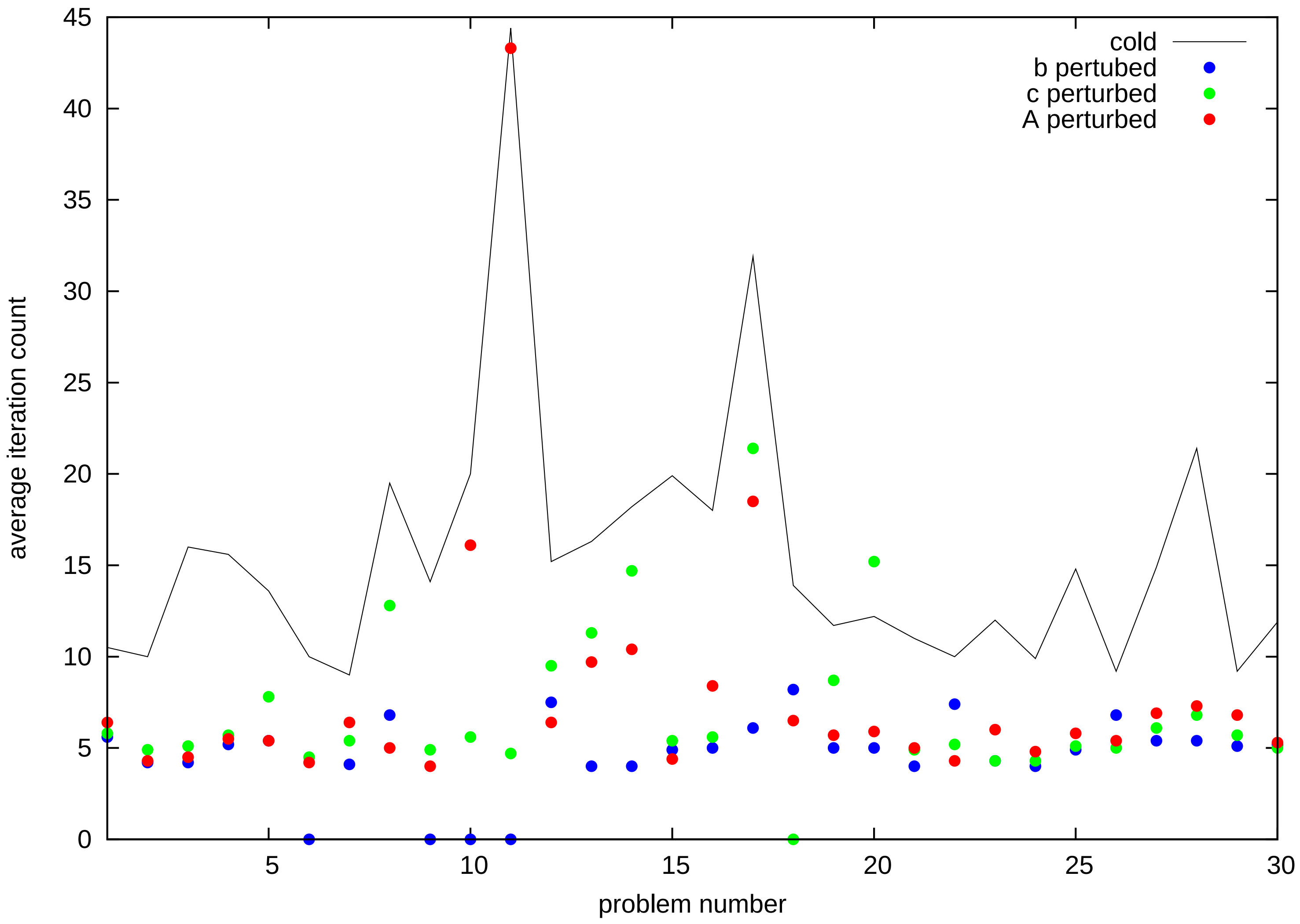


	b			c			A			
	0.1	0.01	0.001	0.1	0.01	0.001	0.1	0.01	0.001	total
no-sens	6.1	5.0	5.0	14.9	8.7	6.2	11.5	7.0	6.6	8.0
sens-1	5.9	4.9	5.0	13.2	7.9	6.0	10.4	6.8	6.7	7.6
sens-2	5.8	5.0	5.0	11.9	8.0	6.1	9.7	6.7	6.9	7.4
sens-3	5.8	5.2	5.1	11.5	7.7	5.8	9.7	6.8	6.8	7.3

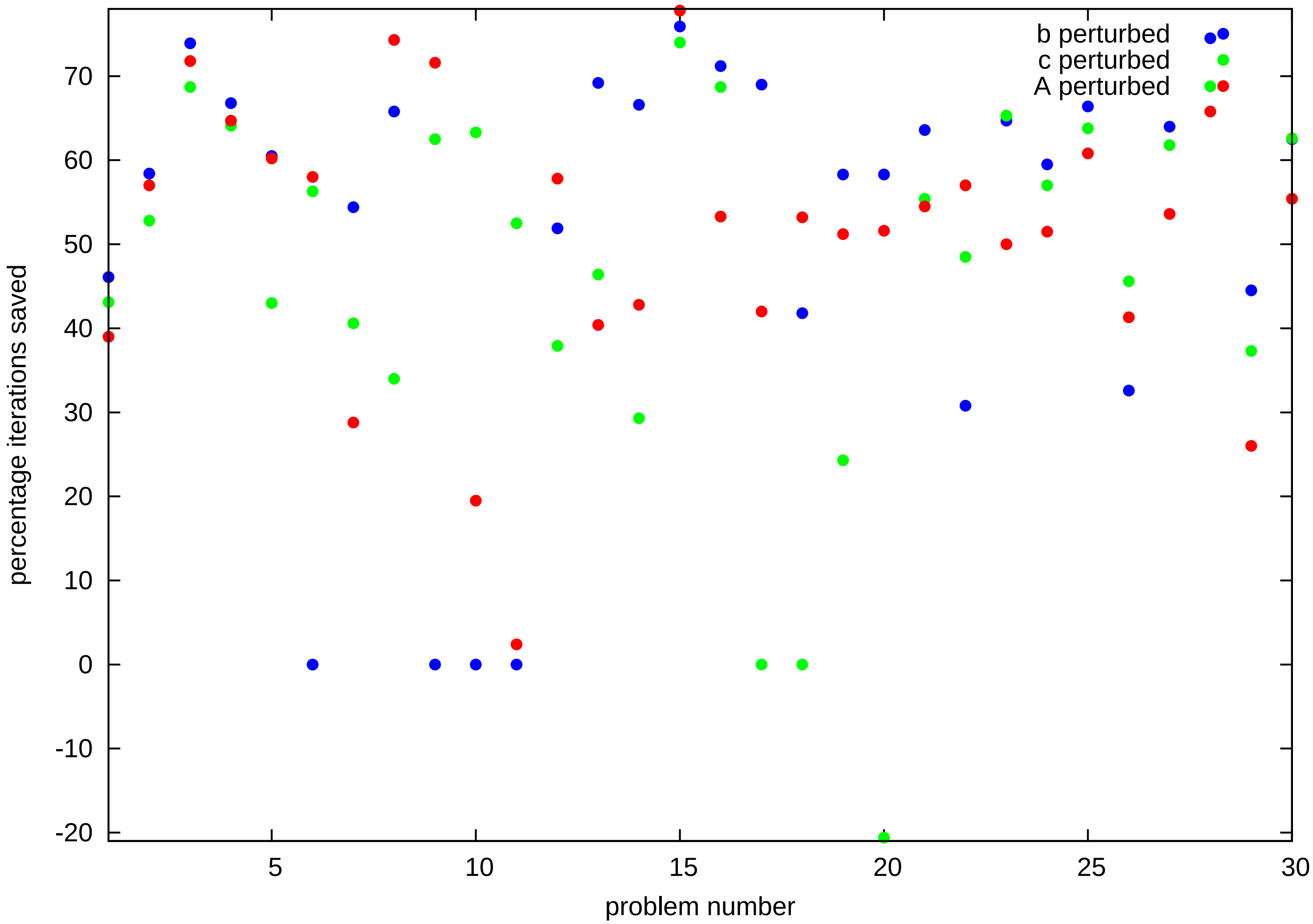
Results (Best Warmstart) - all perturbations

Problem	b			c			A		
	cold	warm	per	cold	warm	perc	cold	warm	per
ADLITTLE	10.4	5.6	46.1	10.2	5.8	43.1	10.5	6.4	39.0
AFIRO	10.1	4.2	58.4	10.4	4.9	52.8	10.0	4.3	57.0
AGG2	16.1	4.2	73.9	16.3	5.1	68.7	16.0	4.5	71.8
AGG3	15.7	5.2	66.8	15.9	5.7	64.1	15.6	5.5	64.7
BANDM	13.7	5.4	60.5	13.7	7.8	43.0	13.6	5.4	60.2
BEACONFD	-	-	-	10.3	4.5	56.3	10.0	4.2	58.0
BLEND	9.0	4.1	54.4	9.1	5.4	40.6	9.0	6.4	28.8
BOEING1	19.9	6.8	65.8	19.4	12.8	34.0	19.5	5.0	74.3
BORE3D	-	-	-	13.1	4.9	62.5	14.1	4.0	71.6
BRANDY	-	-	-	15.3	5.6	63.3	20.0	16.1	19.5
DEGEN2	-	-	-	9.9	4.7	52.5	44.4	43.3	2.4
E226	15.6	7.5	51.9	15.3	9.5	37.9	15.2	6.4	57.8
GROW15	13.0	4.0	69.2	21.1	11.3	46.4	16.3	9.7	40.4
GROW7	12.0	4.0	66.6	20.8	14.7	29.3	18.2	10.4	42.8
ISRAEL	20.4	4.9	75.9	20.8	5.4	74.0	19.9	4.4	77.8
KB2	17.4	5.0	71.2	17.9	5.6	68.7	18.0	8.4	53.3
LOTFI	19.7	6.1	69.0	21.4	21.4	0.0	31.9	18.5	42.0
RECIPELP	14.1	8.2	41.8	-	-	-	13.9	6.5	53.2
SC105	12.0	5.0	58.3	11.5	8.7	24.3	11.7	5.7	51.2
SC205	12.0	5.0	58.3	12.6	15.2	-20.6	12.2	5.9	51.6
SC50A	11.0	4.0	63.6	11.0	4.9	55.4	11.0	5.0	54.5
SC50B	10.7	7.4	30.8	10.1	5.2	48.5	10.0	4.3	57.0
SCAGR25	12.2	4.3	64.7	12.4	4.3	65.3	12.0	6.0	50.0
SCAGR7	9.9	4.0	59.5	10.0	4.3	57.0	9.9	4.8	51.5
SCFXM1	14.6	4.9	66.4	14.1	5.1	63.8	14.8	5.8	60.8
SCSD1	10.1	6.8	32.6	9.2	5.0	45.6	9.2	5.4	41.3
SCTAP1	15.0	5.4	64.0	16.0	6.1	61.8	14.9	6.9	53.6
SHARE1B	21.2	5.4	74.5	21.8	6.8	68.8	21.4	7.3	65.8
SHARE2B	9.2	5.1	44.5	9.1	5.7	37.3	9.2	6.8	26.0
STOCFOR1	13.9	5.2	62.5	13.4	5.0	62.6	11.9	5.3	55.4
Average	13.8	5.3	59.6	14.2	7.3	48.4	15.5	8.0	50.9

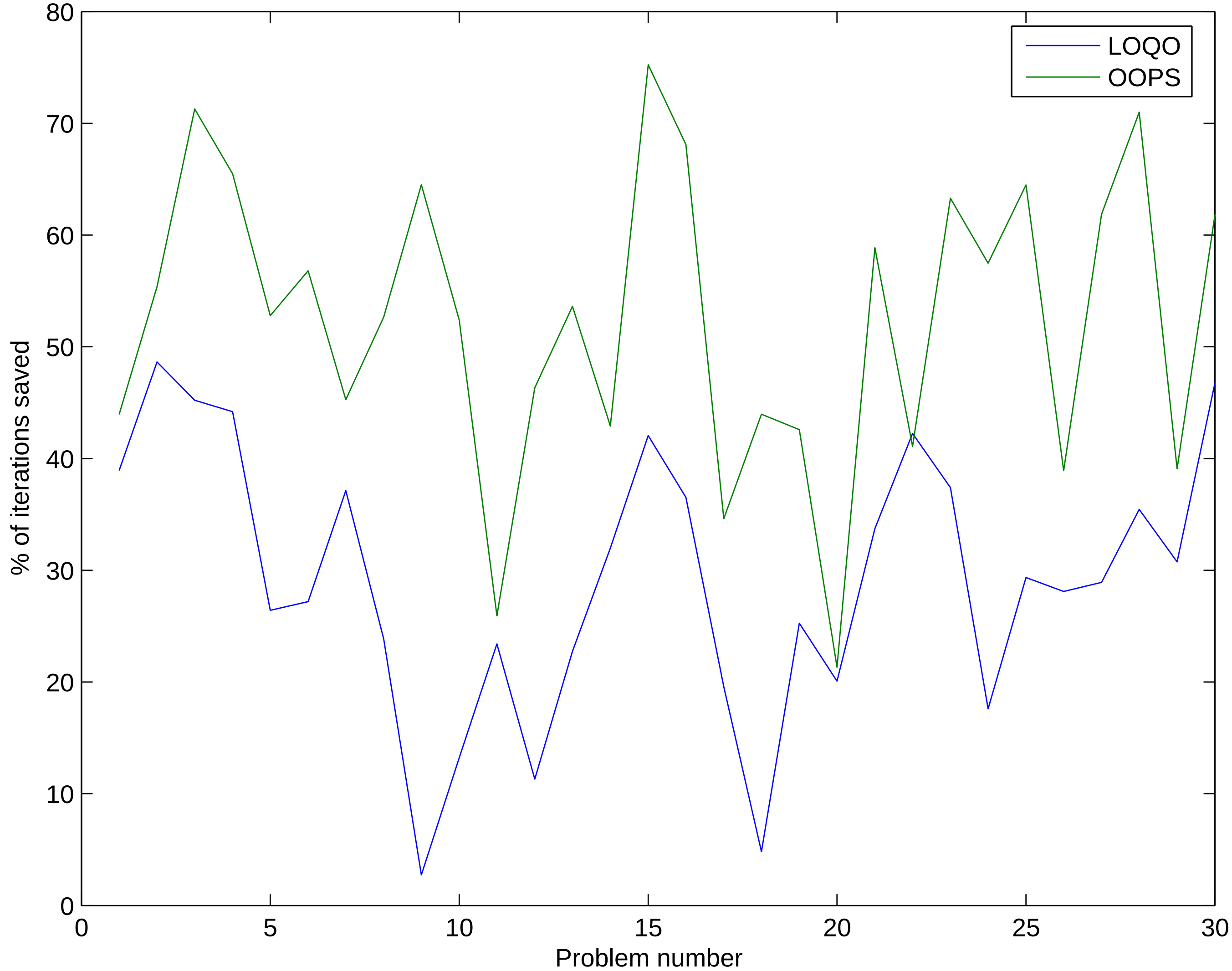
Results for LP problems (NETLIB)



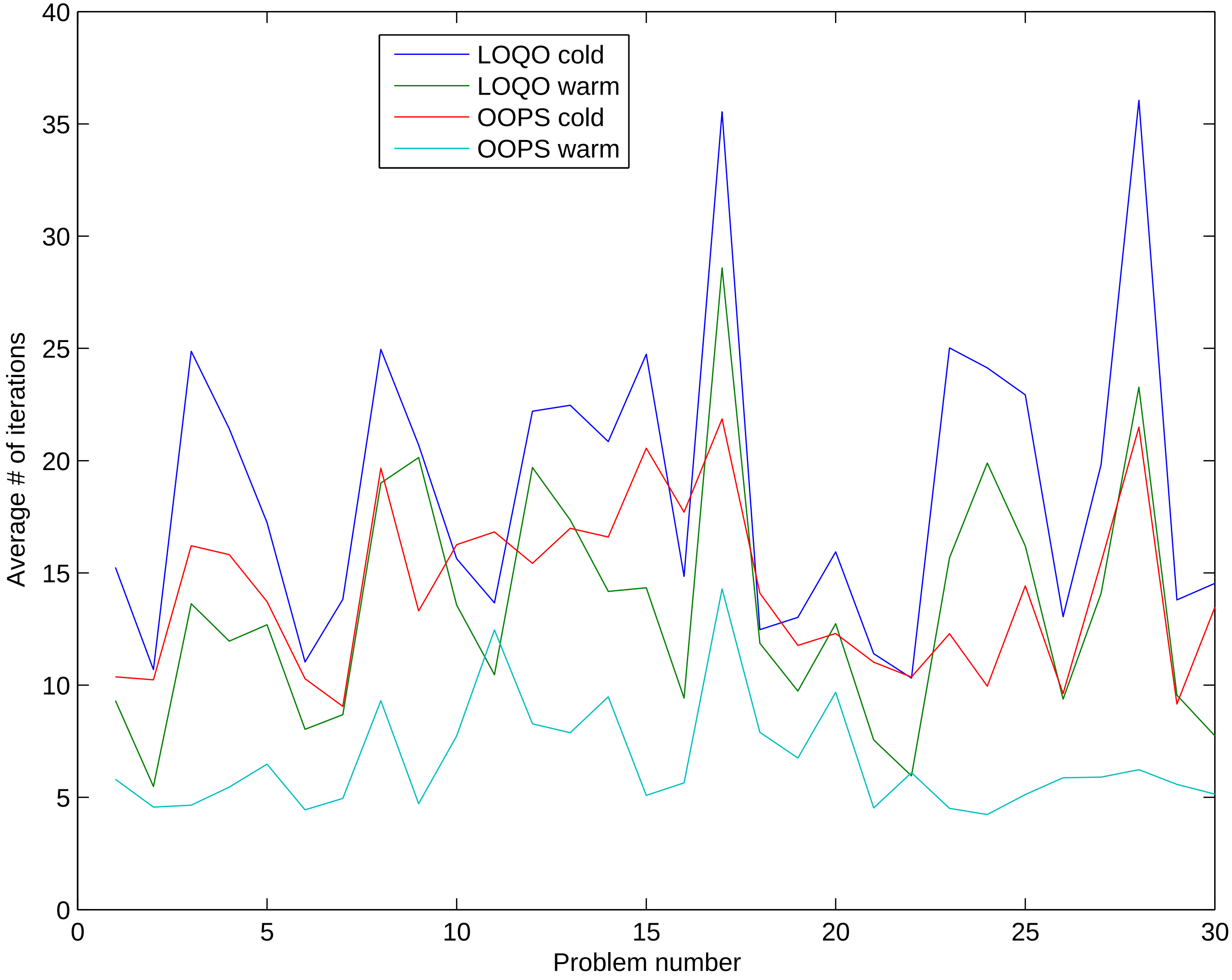
Results for LP problems (NETLIB)



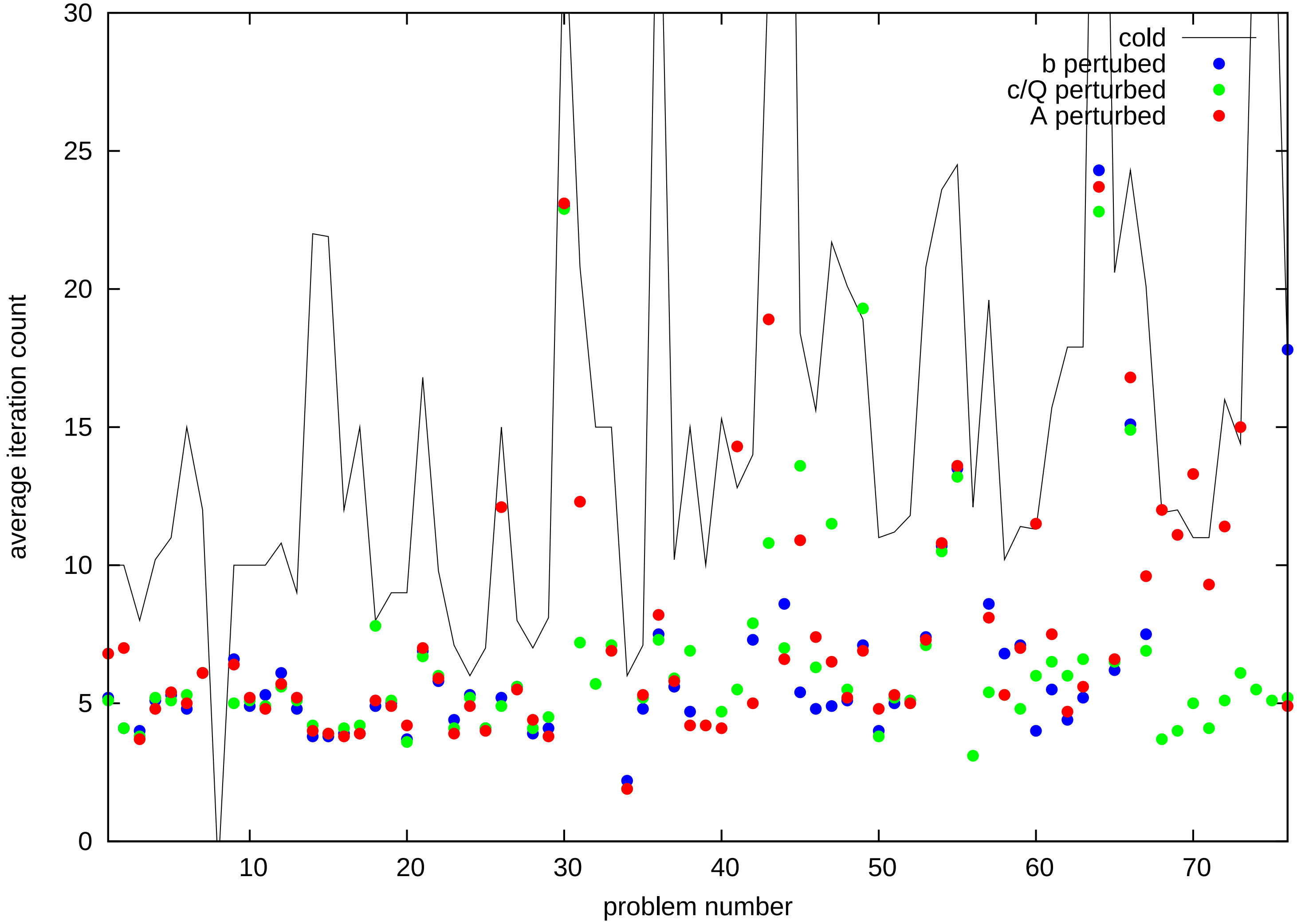
Comparisons with LOQO warmstart (Benson, Shanno '06)



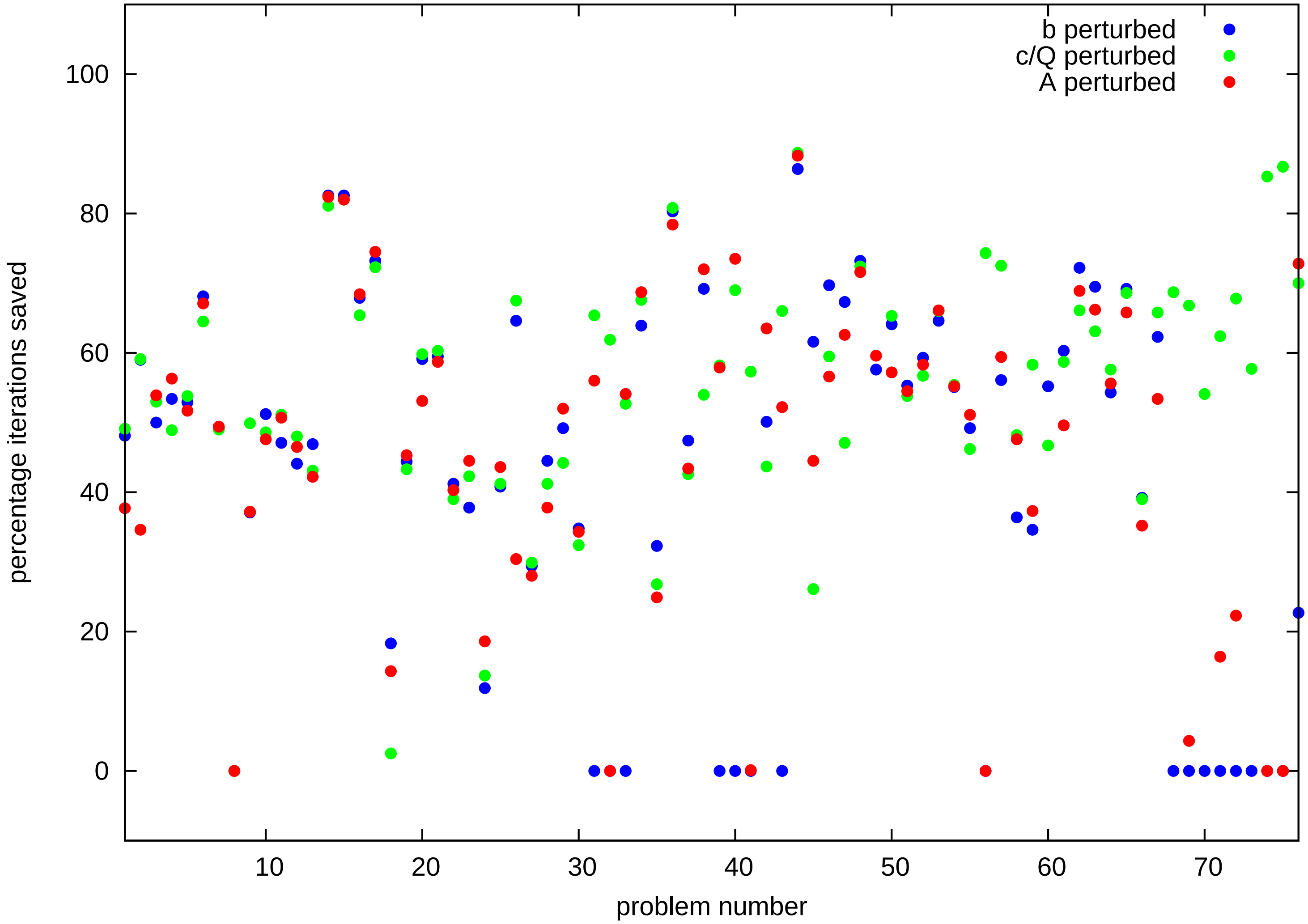
Comparisons with LOQO warmstart (Benson, Shanno '06)



Results for QP problems (Maros/Mezaros: CUTE & NETLIB)



Results for QP problems (Maros/Mezaros: CUTE & NETLIB)



Warmstarting Large Scale Problems

SQP: Capacitated MCNF:

$$\begin{array}{lll} \min & \sum_{(i,j)\in\mathcal{E}} \frac{x_{ij}}{K_{ij}-x_{ij}} & \\ \text{s. t.} & \sum_{k\in\mathcal{D}} x_{ij}^{(k)} \le K_{ij}, & \forall (i,j)\in\mathcal{E}, \\ & Nx^{(k)} = d^{(k)}, & \forall k\in\mathcal{D}, \\ & x^{(k)} \ge 0, & \forall k\in\mathcal{D}. \end{array}$$

- Nonlinear objective, linear constraints
- Solve by SQP, subproblems solved by IPM
- Use warmstarts for each QP problem in sequence

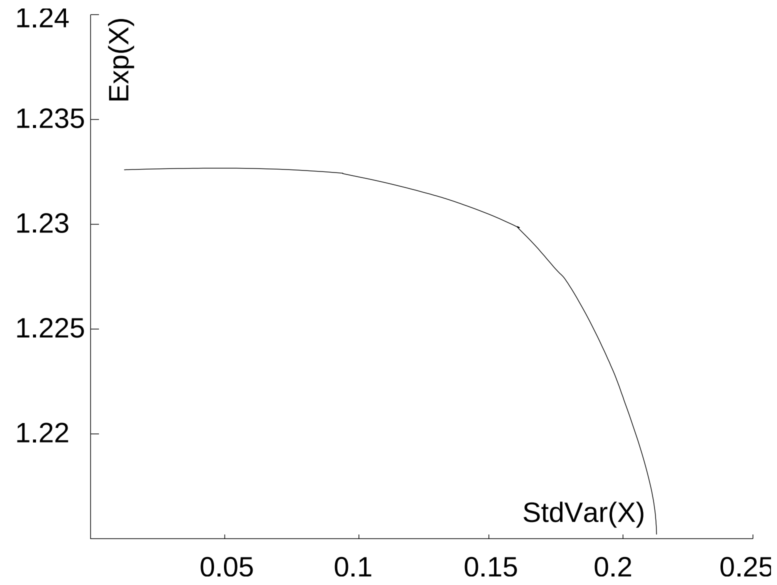
Application: Efficient Frontier (Portfolio Optimization)

- Investors differ in their attitude to risk.
- Captured by the *risk-aversion* parameter ρ in the mean-variance model.
- Efficient Frontier provides a more complete understanding

solve

$$\max \mathbb{E}(X) - \rho \mathrm{Var}(X)$$

for a range of values for ρ



Gives max achievable expected gain for a given risk, or vice versa.

Results: Capacitated MCNF by SQP

- Reported averages over 9 different network
- 4-300 nodes, Up to 600 arcs, 7021 commodities.
- up to 353.400 variables

iter	1	2	3	4	5	6	7	8	9	10
cold	12.7	11.9	13.7	15.8	16.2	15.6	14.9	14.6	14.5	15.0
warm	12.7	7.0	6.0	5.8	6.4	7.0	7.0	6.7	6.2	6.0
per	0.0	41.2	56.2	63.3	60.5	55.1	53.0	54.1	57.2	60.0

Reported are average iterations for first 10 QPs solved

Efficient Frontier (Results)

Problem: Multistage SP formulation of Mean-Variance Model

constraints	variables	ρ =0.001	0.005	0.01	0.05	0.1	0.5	1	5	10
223.321	76.881	14	14	14	14	14	13	17	16	17
		14	5	5	5	4	5	5	8	8
		0.0	64.2	64.2	64.2	71.4	61.5	70.5	50.0	52.9
533.725	198.525	14	14	14	14	14	15	18	18	17
		14	5	5	5	6	5	5	9	10
		0.0	64.3	64.3	64.3	57.1	66.7	72.2	50.0	41.2
5.982.604	16.316.191	24	23	24	23	25	22	24	23	24
		24	8	11	13	11	13	12	12	14
		0.0	65.2	54.2	43.5	56.0	40.9	50.0	47.8	41.7
70.575.308	192.478.111	52	53	45	43	44	42	44	46	46
		52	13	13	15	15	16	16	23	25
		0.0	75.5	71.1	65.1	65.9	61.9	63.6	50.0	45.6

Towards Structured Warmstarts:

Many/most large scale problems are structured

Aim: Use problem structure

- to facilitate warmstarting.
- to speed up the overall solution by **constructing** warmstart point

Towards Structured Warmstarts:

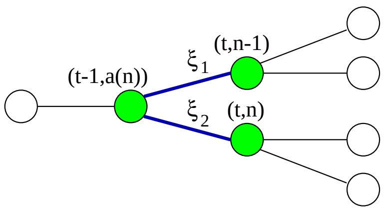
Many/most large scale problems are structured

Aim: Use problem structure

- to facilitate warmstarting.
- to speed up the overall solution by **constructing** warmstart point

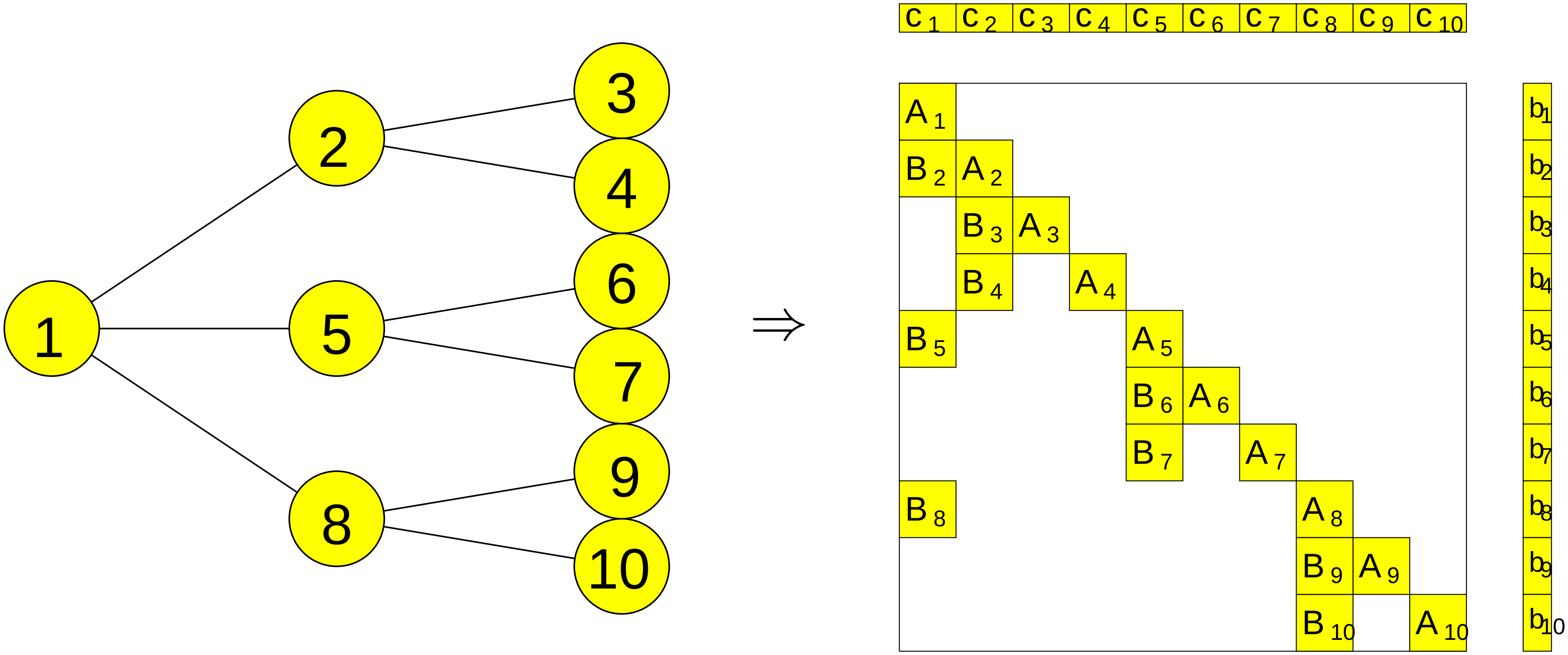
Example: Stochastic programming

- Some of the problem data is uncertain
- Stochastic programming uses an event tree to discretise different possible future values.

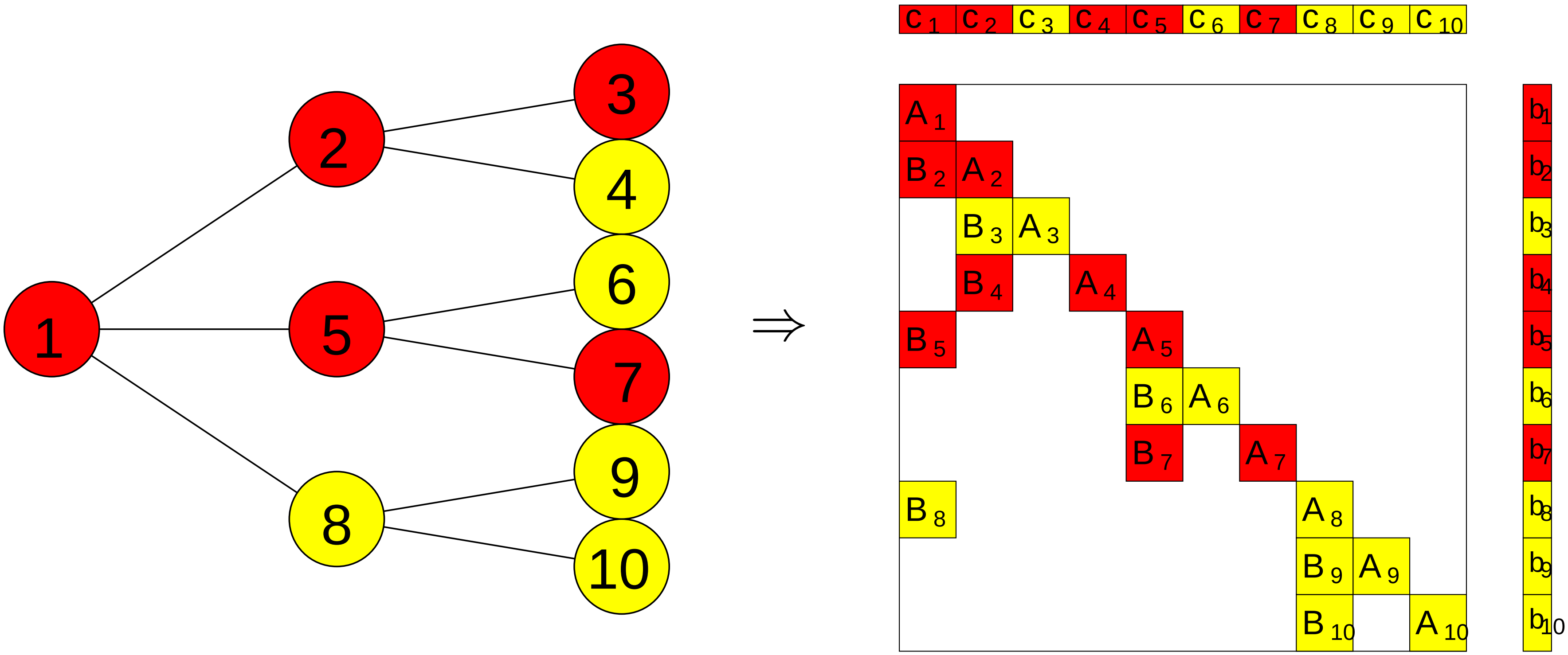


- Problems quickly become very large (100.000- 100.000.000 variables)
- But structure is ameniable to IPM

Towards Structured Warmstarts: Stochastic Programming

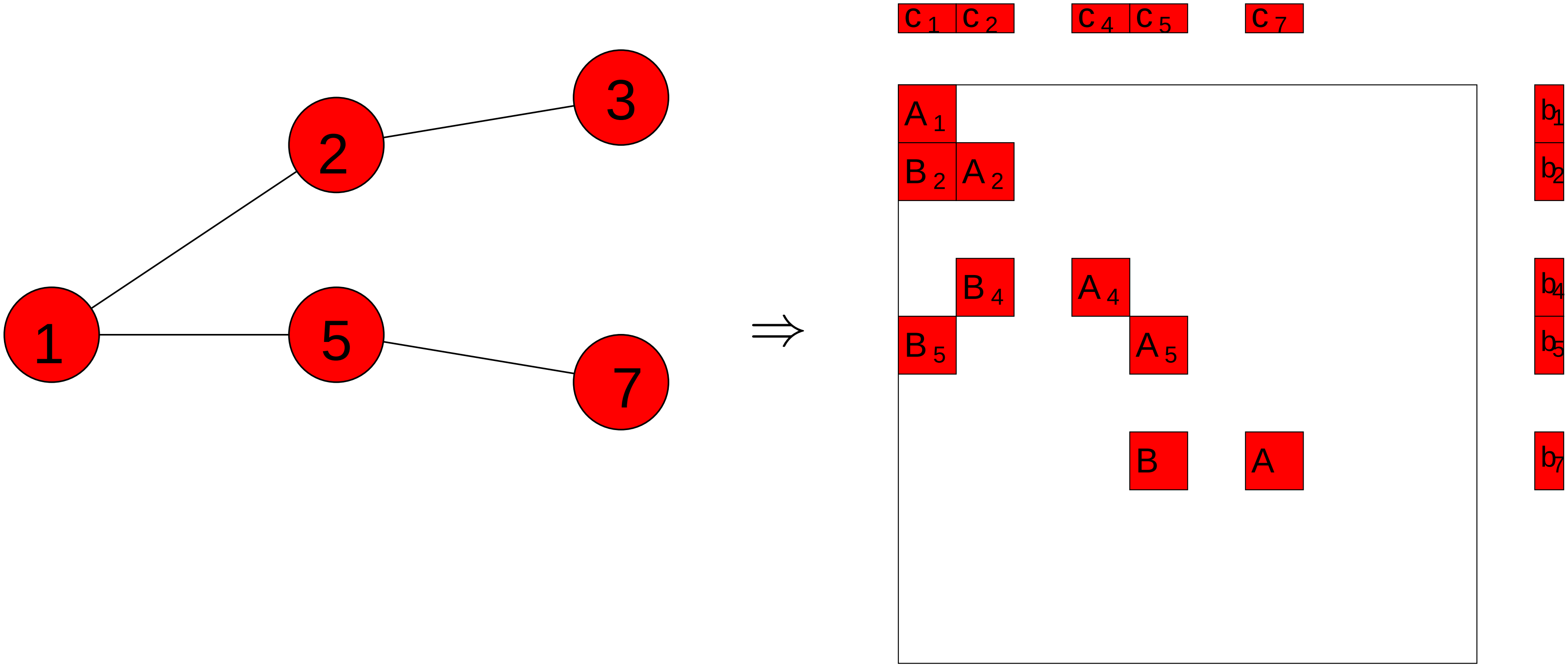


Towards Structured Warmstarts: Stochastic Programming



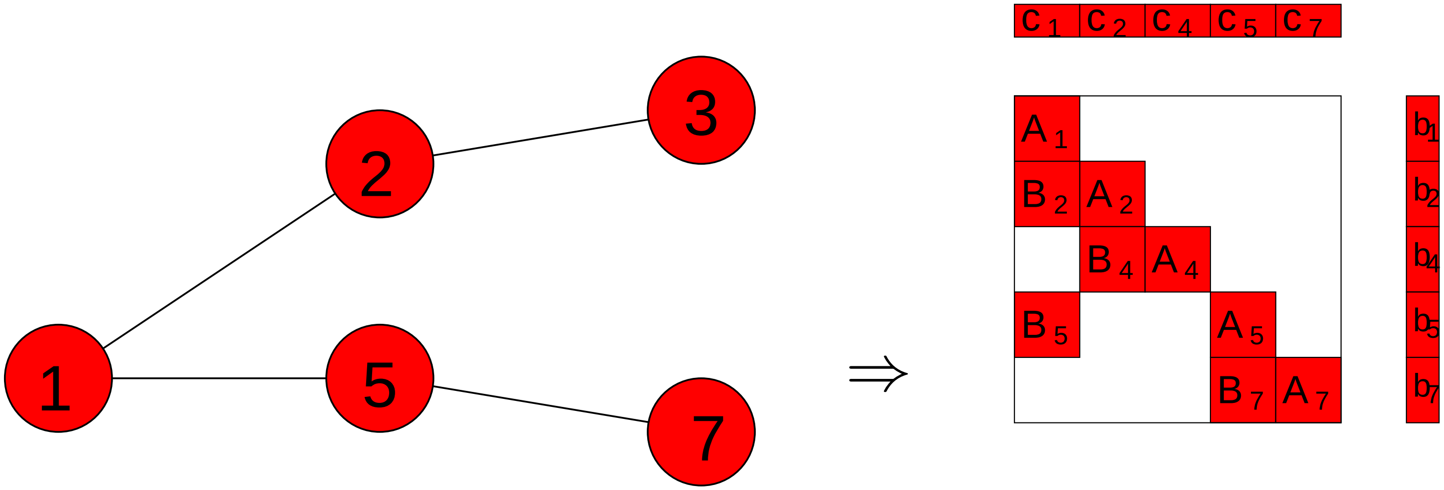
- Select sample scenarios

Towards Structured Warmstarts: Stochastic Programming



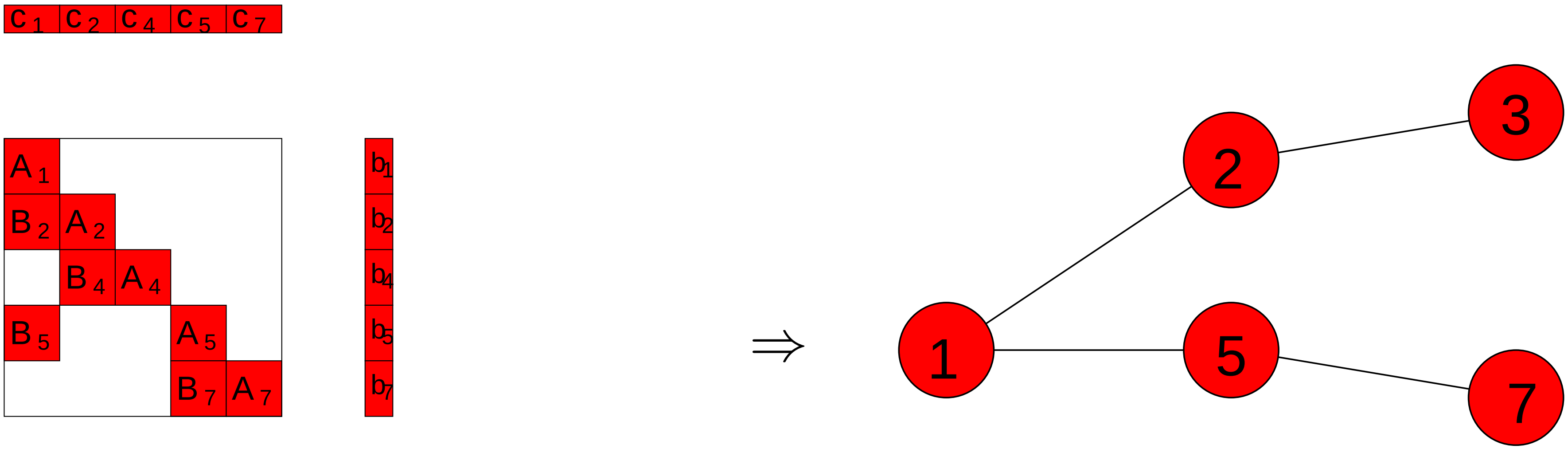
- Select sample scenarios
- Aggregate Scenarios/Reduce Problem

Towards Structured Warmstarts: Stochastic Programming



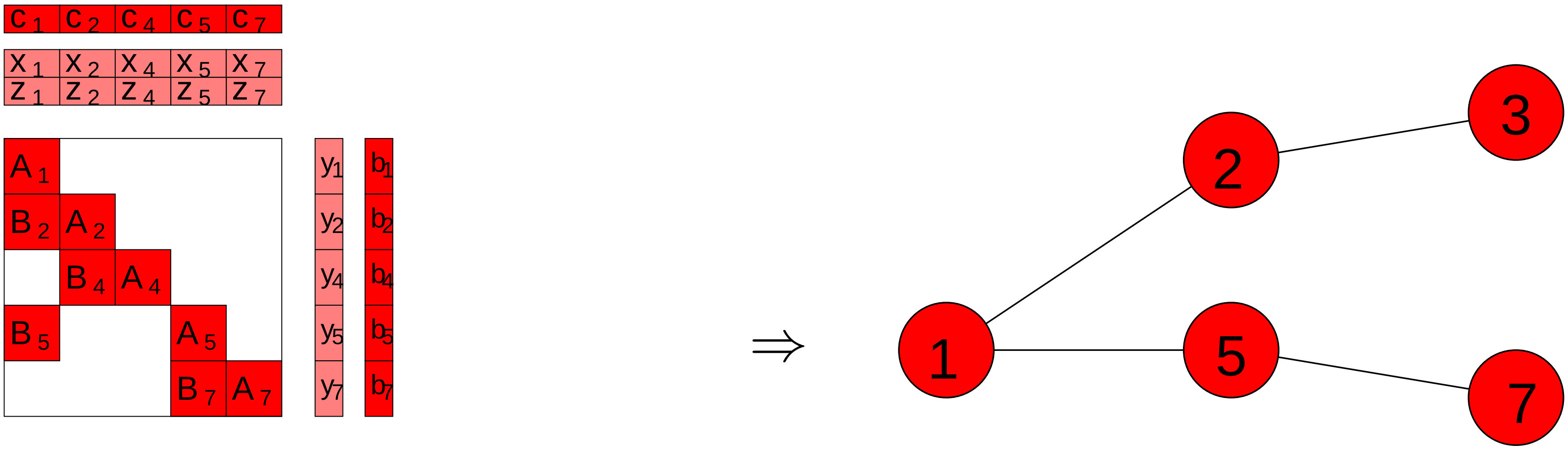
- Select sample scenarios
- Aggregate Scenarios/Reduce Problem

Towards Structured Warmstarts: Stochastic Programming



To solve the problem by warmstarting, reverse the process

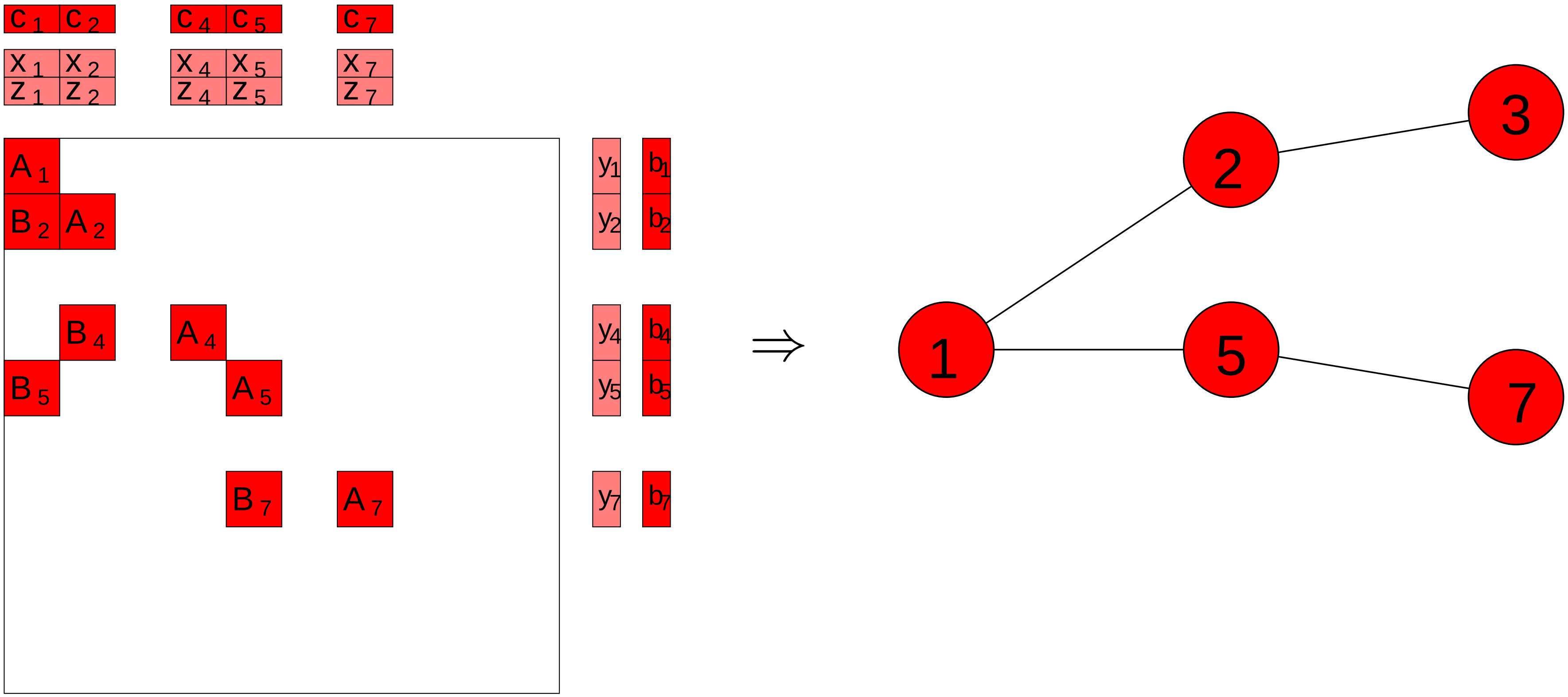
Towards Structured Warmstarts: Stochastic Programming



To solve the problem by warmstarting, reverse the process

- Solve reduced problem (to low accuracy)

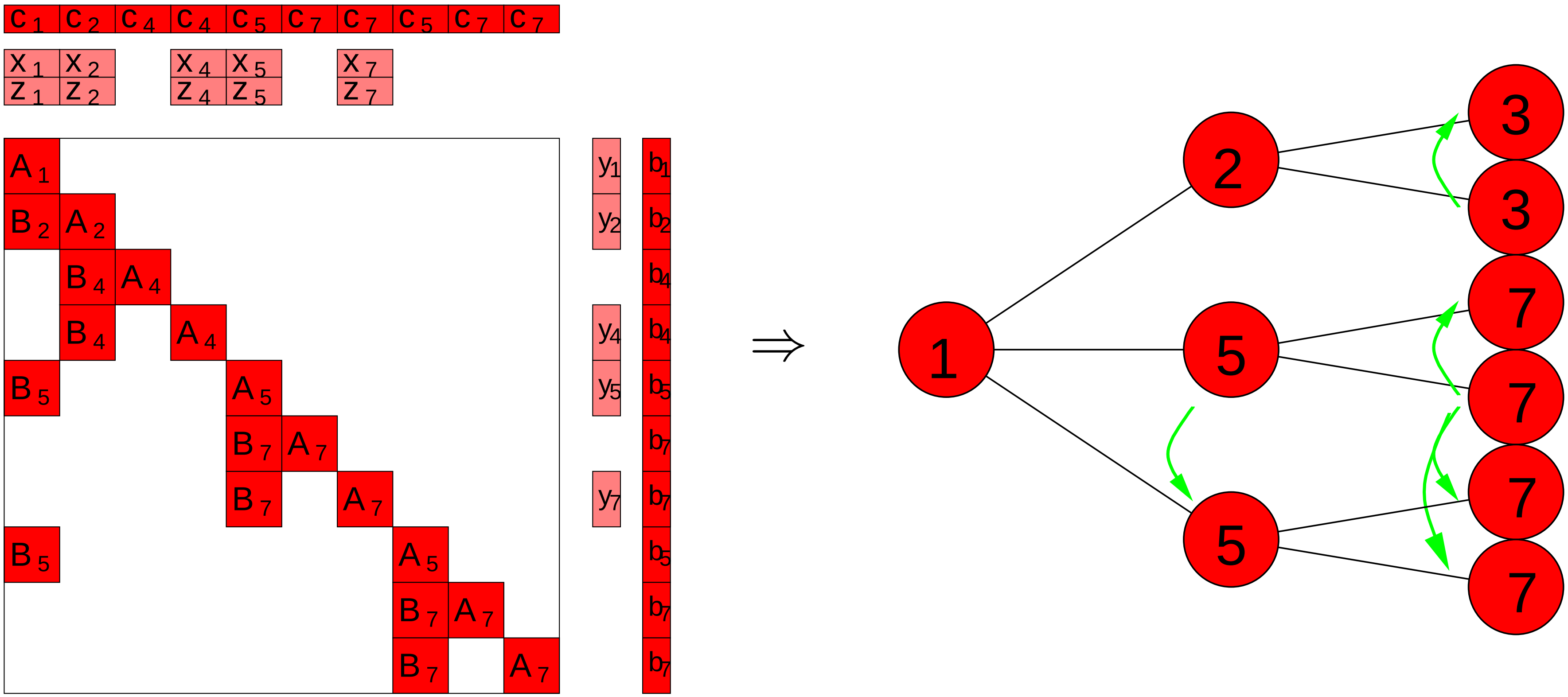
Towards Structured Warmstarts: Stochastic Programming



To solve the problem by warmstarting, reverse the process

- Solve reduced problem (to low accuracy)
- Expand the problem to original size

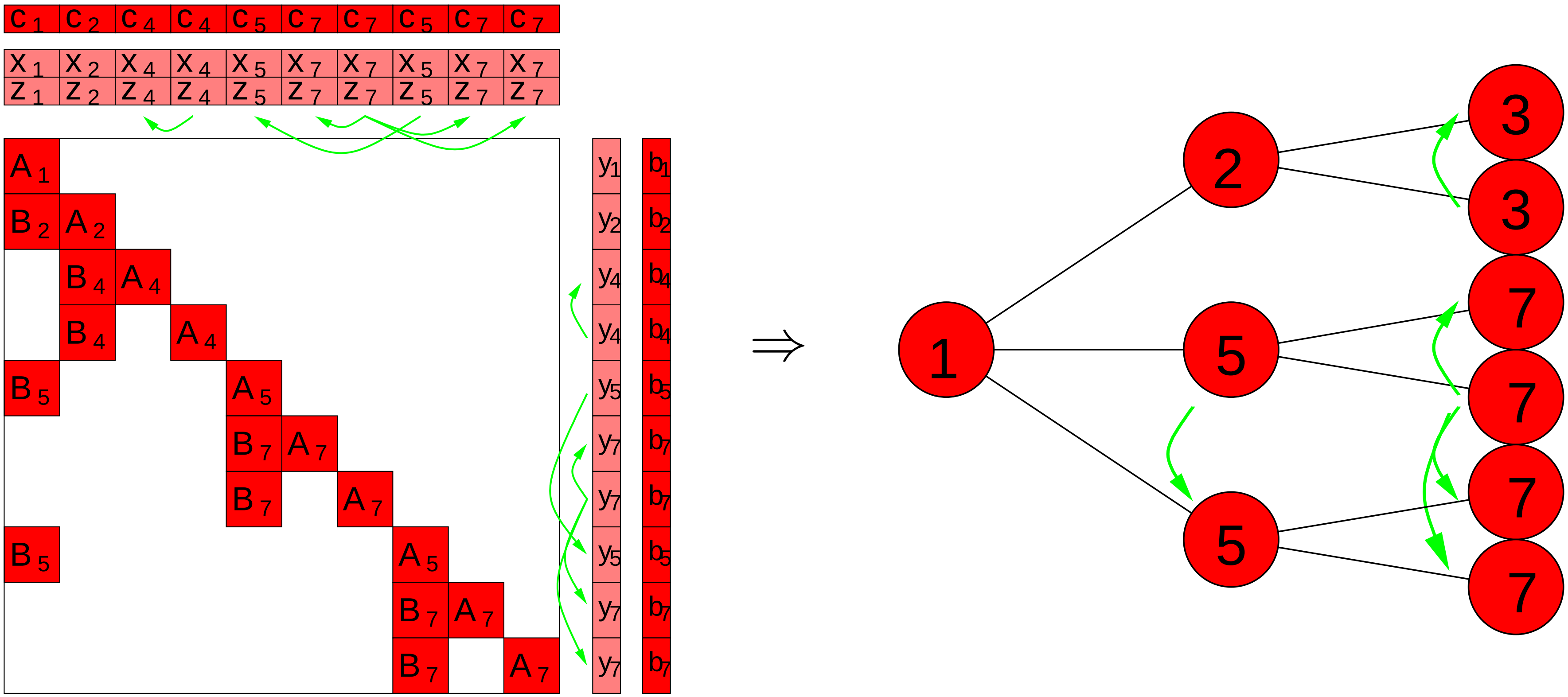
Towards Structured Warmstarts: Stochastic Programming



To solve the problem by warmstarting, reverse the process

- Solve reduced problem (to low accuracy)
- Expand the problem to original size (by duplicating scenarios)

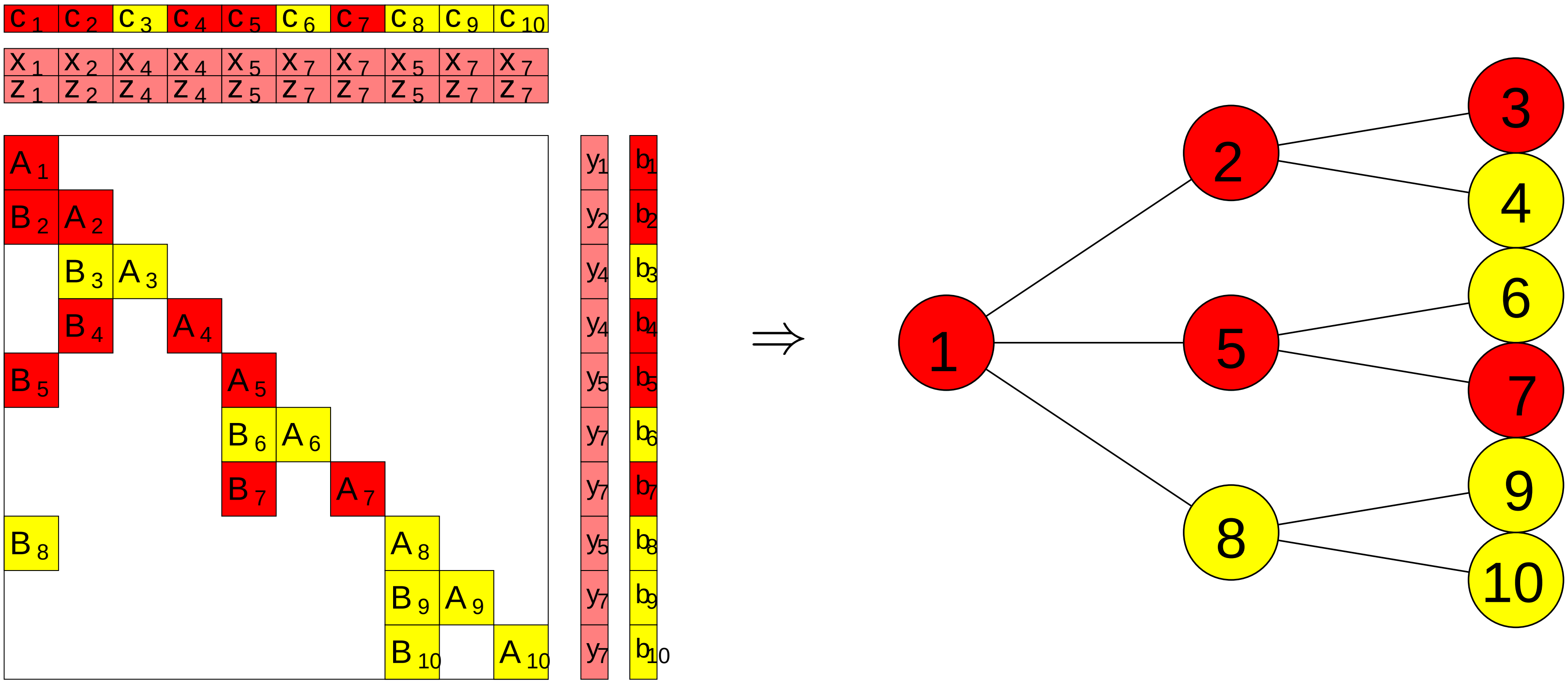
Towards Structured Warmstarts: Stochastic Programming



To solve the problem by warmstarting, reverse the process

- Solve reduced problem (to low accuracy)
- Expand the problem to original size (by duplicating scenarios)
- Expand solution to primal/dual feasible solution to expanded problem

Towards Structured Warmstarts: Stochastic Programming

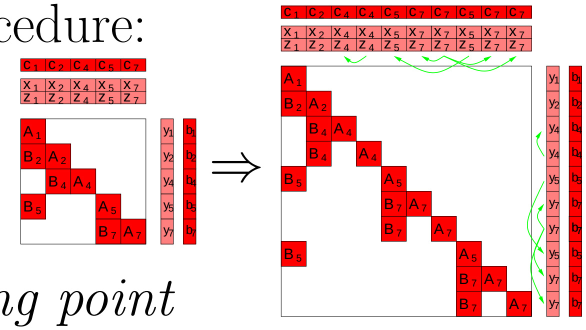


To solve the problem by warmstarting, reverse the process

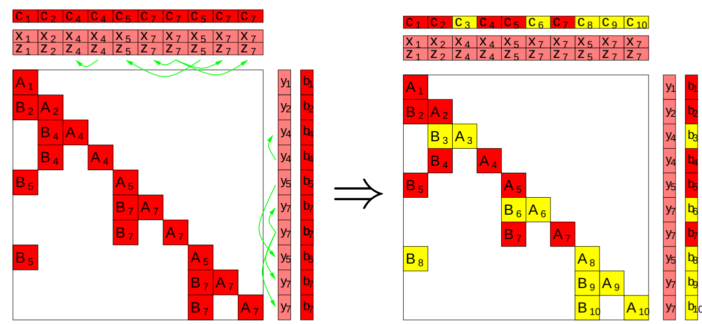
- Solve reduced problem (to low accuracy)
- Expand the problem to original size (by duplicating scenarios)
- Expand solution to primal/dual feasible solution to expanded problem
- Use this to warmstart full problem

Towards Structured Warmstarts: Stochastic Programming

The proposed warmstart procedure is a **two** stage procedure:



- Reduced Problem ⇒ Expanded Problem
  - *Can construct primal/dual feasible starting point*
  - *Although point is not central (We are duplicating constraints!)*



- Expanded Problem ⇒ Full Problem
  - *Can bound changes in problem data (differences in scenarios)*

Preliminary results

Two sources of test problems:

- Collection of standard SMPS problems.
- Capacity assignment problem with uncertain demand:

$$\begin{array}{ll} \min_{x\geq 0} & E_d[f(x,d)] \\ \text{s.t.} & \sum_{l\in\mathcal{A}} c_l x_l \leq M \end{array} \qquad f(x,d) = \min_{z_p\geq 0} \sum_{k\in\mathcal{D}} (d_k - \sum_{p\in P_k} z_p)$$

$$\text{s.t.} \sum_{k\in\mathcal{D}} \sum_{p\in\mathcal{P}_k : l\in p} z_p \leq x_l \quad \forall l \in \mathcal{A}$$

$$\sum_{p\in\mathcal{P}_k} z_p \leq d_k \qquad \forall k \in \mathcal{D}$$

Setup:

- 2 scenarios in the reduced tree
- Reduced problem optimality tolerance: 5.0×10^{-1}
- Complete problem optimality tolerance: 5.0×10^{-8}

Numerical results: Standard SP test problems

Problem data			Cold start		Warm start	
Name	Stgs	Scens	Iters	Seconds	Iters	Seconds
fxm2-16	2	16	22	1.2	13	1.0
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	41	95.4	22	53.2
storm125	2	125	73	107.3	36	69.1
storm1000	2	1000	107	1498.3	45	831.5

Numerical results: Capacity Assignment Problems:

Problem data			Cold start		Warm start	
Name	Stgs	Scens	Iters	Seconds	Iters	Seconds
mnx-200	2	200	13	12.9	7	7.3
mnx-800	2	800	17	58.8	10	39.5
mnx-1600	2	1600	19	131.1	10	68.8
jlg-200	2	200	45	164.9	17	39.5
jlg-800	2	800	27	353.4	10	152.9
jlg-1600	2	1600	32	855.3	13	360.6
mgntA-100	2	100	28	260.0	14	156.2
mgntA-200	2	200	50	877.1	35	690.6
mgntA-400	2	400	40	1470.3	14	572.5
mgntB-100	2	100	23	511.1	14	318.0
mgntB-200	2	200	25	909.4	8	332.4
mgntB-400	2	400	29	2154.5	7	538.1

Conclusions

- IPM **can** be warmstarted
- IPM warmstarts can save 50%-60% of iterations (on all problem sizes)
- We have reviewed different warmstarting techniques
- Unblocking strategies have demonstrated benefits
- Combinations leads to very efficient warmstart

Future Work:

- Use structure of the problem to construct good warmstarting point.
- Multi-Step scheme
- Complexity of such a scheme
- Carry over to other structures (PDE constrained optimization)
- Integrate into structured modelling language