Convex Quasi-Relaxations for Global Optimization

J. N. Hooker
Carnegie Mellon University, USA

November 2005

Goal

- Find a way to bound the optimal value of a problem that is nonconvex…
    - …but has no apparent convex relaxation.
- Many problems are "partially convex."
    - They become convex when some variables are fixed.
- Find a way to exploit this property.
    - No need to solve a true relaxation, so long as we get a valid bound.

Basic Idea

- Assume the problem becomes convex when certain variables $y_1, \ldots, y_k$ are fixed.
  - Branch on $y_1, \ldots, y_k$.
  - Problem is convex at leaf nodes.
- Compute a bound at each nonleaf node while branching on $y_1, \ldots, y_k$.
  - Form a *convex quasi-relaxation* of the problem.
  - This exists under certain conditions.
    - The problem is concave in $y_1, \ldots, y_k$ and homogeneous in the other variables.
    - It can be written with one y_j per constraint.

Motivation

- Take advantage of advanced solution methods:
- *Branch-and-bound* method…
  - for branching on y_1, …, y_k.
- *Nonlinear programming* method…
  - for solving convex quasi-relaxations.
  - for solving convex subproblems at nodes of the search tree.

General Form of Problem

Vector of functions

$$\begin{array}{ll} \min & x_0 \\ \text{subject to} & g^j(x, y_j) \le 0, \quad j \in J \\ & L(y) \\ & x \in \Re^n, \quad y_j \in Y_j \end{array}$$

Logical conditions on *y*, enforced while branching

General Form of Problem

Vector of functions

$$\begin{array}{ll} \min & x_0 \\ \text{subject to} & g^j(x, y_j) \le 0, \quad j \in J \\ & L(y) \\ & x \in \Re^n, \quad y_j \in Y_j \end{array}$$

Logical conditions on *y*, enforced while branching

Assume that when y is fixed to $\bar{y}$, we get a *convex* problem

convex functions of *x*

$$\begin{array}{ll} \min & x_0 \\ \text{subject to} & g^j(x, \bar{y}_j) \le 0, \quad j \in J \\ & x \in \Re^n \end{array}$$

If y_j is continuous, discretize it, to get approximate global solution.

Objective is defined in the constraints

$$\begin{array}{ll} \text{min} & x_0 \\ \text{subject to} & g^j(x, y_j) \le 0, \quad j \in J \\ & L(y) \\ & x \in \Re^n, \quad y_j \in Y_j \end{array}$$

We assume one y_j per constraint.

Many problems have this form. If not, constraints can in principle be put into this form by change of variable.

For example, consider

$$\begin{aligned} &x + \hat{y}_1 + \hat{y}_2 \geq b \\ &x + \hat{y}_2 + \hat{y}_3 \geq b \\ &\hat{y}_j \in \{0,1,2\} \end{aligned}$$

Use the change of variable

$$\begin{aligned} &y_1 = \hat{y}_1 + \hat{y}_2 \\ &y_2 = \hat{y}_2 + \hat{y}_3 \\ &y_j \in \{0,\ldots,4\} \end{aligned}$$

and the constraints have the desired form:

One y_j per constraint

$$\begin{aligned} &x + y_1 \geq b \\ &x + y_2 \geq b \end{aligned}$$

Logical condition

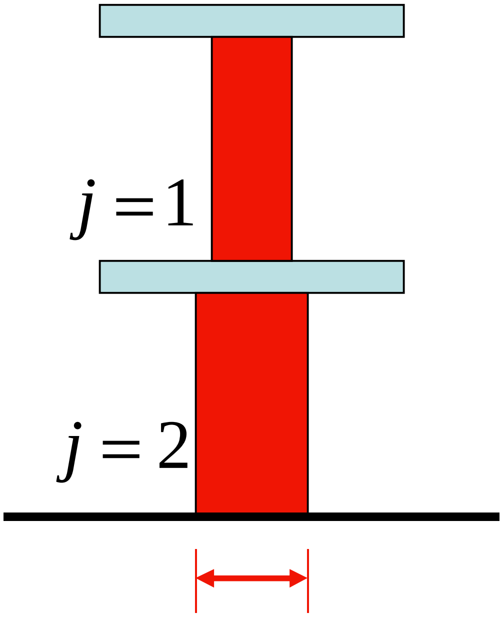
$$\begin{aligned} &|y_1 - y_2| \leq 2 \\ &y_j \in \{0,\ldots,4\} \end{aligned}$$

Structural Design Example

Choose bar thickness that minimizes cost.

This example is intended only to illustrate the idea.
A more realistic model for structural design is presented at the end of the talk.

x_j = compression of bar *j*



y_j = thickness (cross-sectional area) of bar *j*

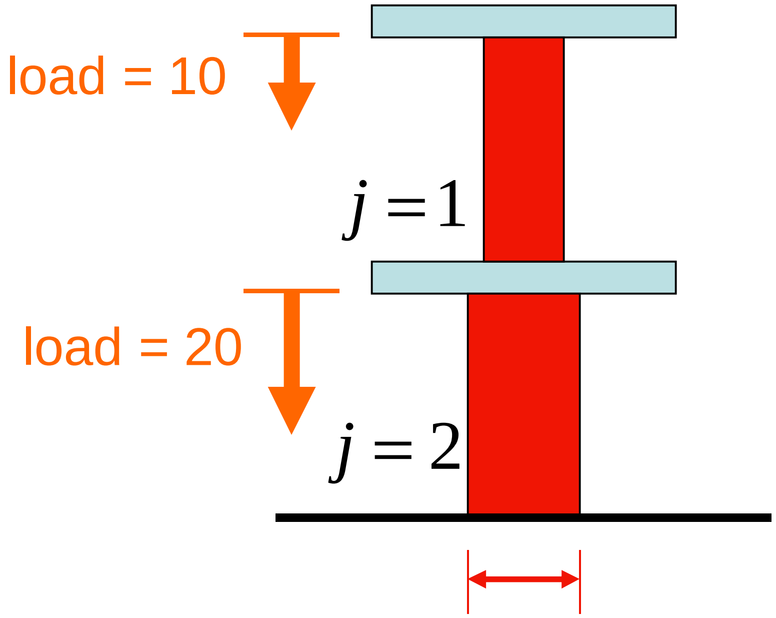
cost = $300y_1 + 300y_2 + (x_1 + x_2)^2 + x_2^2$

Structural Design Example

Choose bar thickness that minimizes cost.

This example is intended only to illustrate the algorithms.
A more realistic model for structural design is presented at the end of the talk.

x_j = compression of bar *j*



y_j = thickness (cross-sectional area) of bar *j*

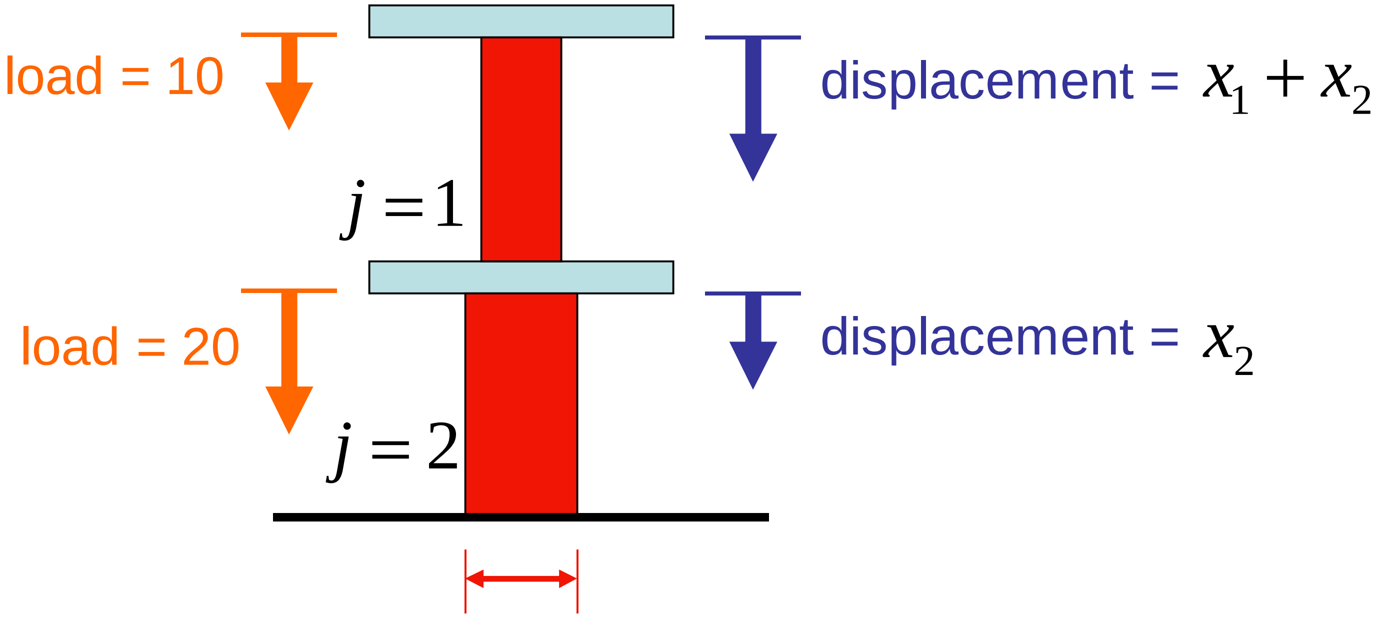
cost = $$300y_1 + 300y_2 + (x_1 + x_2)^2 + x_2^2$$

Structural Design Example

Choose bar thickness that minimizes cost.

This example is intended only to illustrate the algorithms.
A more realistic model for structural design is presented at the end of the talk.

x_j = compression of bar *j*



load = 10

load = 20

$j = 1$

$j = 2$

displacement = $x_1 + x_2$

displacement = x_2

y_j = thickness (cross-sectional area) of bar *j*

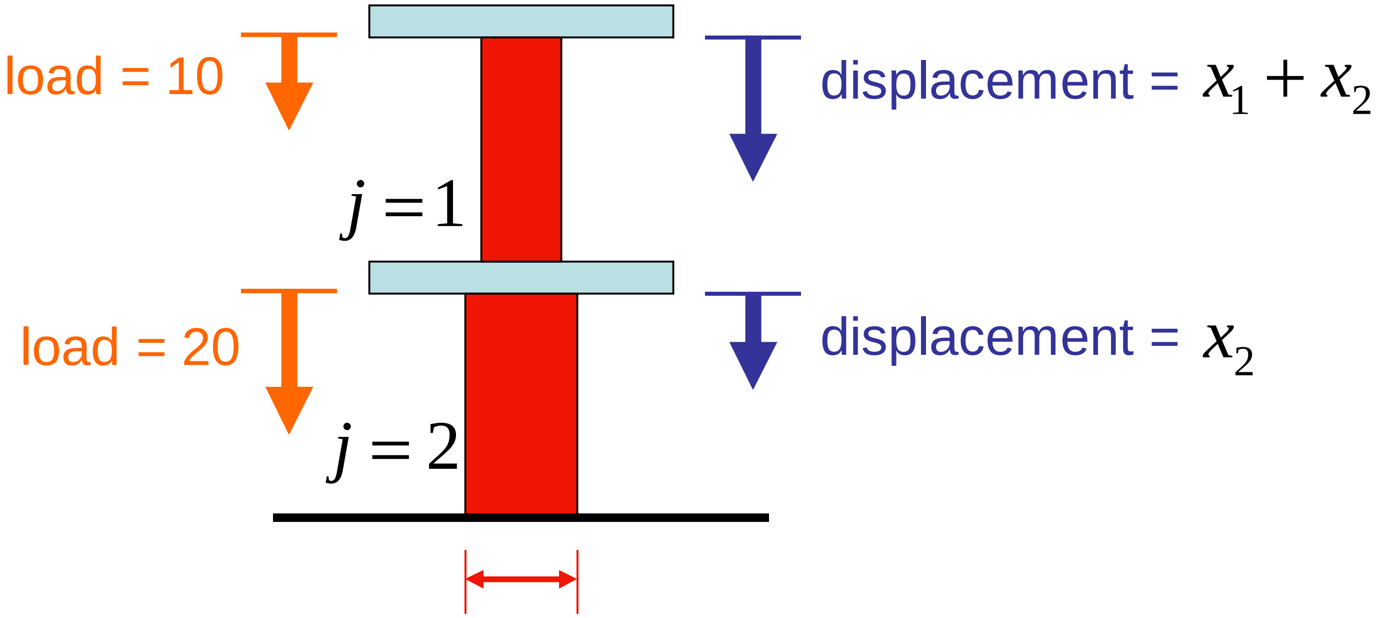
cost = $300 y_1 + 300 y_2 + (x_1 + x_2)^2 + x_2^2$

Structural Design Example

Choose bar thickness that minimizes cost.

This example is intended only to illustrate the algorithms. A more realistic model for structural design is presented at the end of the talk.

x_j = compression of bar *j*



load = 10

displacement = $x_1 + x_2$

$j = 1$

load = 20

displacement = x_2

$j = 2$

y_j = thickness (cross-sectional area) of bar *j*

cost = $300y_1 + 300y_2 + (x_1 + x_2)^2 + x_2^2$

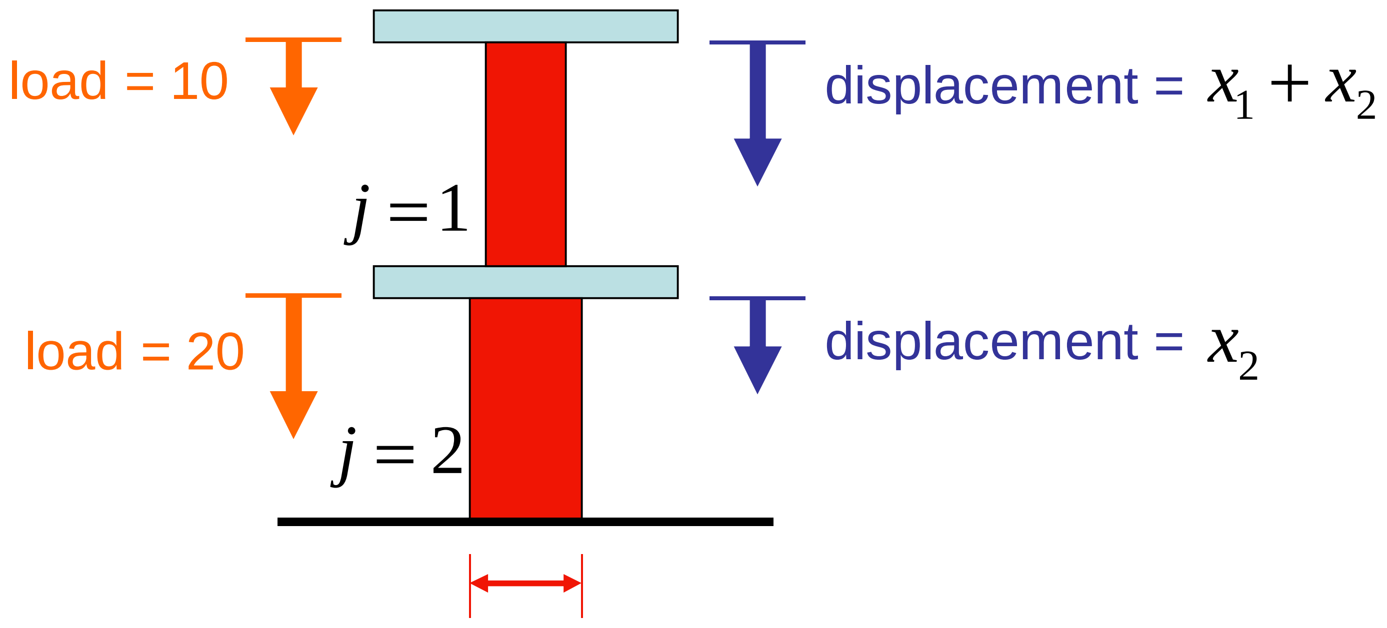
cost of steel

Structural Design Example

Choose bar thickness that minimizes cost.

This example is intended only to illustrate the algorithms.
A more realistic model for structural design is presented at the end of the talk.

x_j = compression of bar *j*



load = 10

load = 20

$j=1$

$j=2$

displacement = $x_1 + x_2$

displacement = x_2

y_j = thickness (cross-sectional area) of bar *j*

cost = $300y_1 + 300y_2 + (x_1 + x_2)^2 + x_2^2$

cost of steel

penalty for displacement

Global optimization problem:

$$\begin{aligned} \min\quad & 300y_1+300y_2+(x_1+x_2)^2+x_2^2 \\ \text{subject to}\quad & x_1y_1=10 \\ & x_2y_2=20 \\ & x_j\ge 0 \\ & y_j\in\{1,2\} \end{aligned}$$

Hooke's law

displacement

thickness

or many closely spaced values for continuous problem

Can be written in desired form:

$$\begin{aligned} \min\quad & x_0 \\ \text{s.t.}\quad & z_1+z_2+(x_1+x_2)^2+x_2^2-x_0\le 0 \\ & \begin{bmatrix} 300y_1-z_1 \\ 10-x_1y_1 \end{bmatrix} \le \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ & \begin{bmatrix} 300y_2-z_2 \\ 20-x_2y_2 \end{bmatrix} \le \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ & x_j\ge 0,\quad y_j\in\{1,2\} \end{aligned}$$

$g^1((x,z),y_1)$

$g^2((x,z),y_2)$

How to Solve It?

$$\begin{array}{ll} \min & x_0 \\ \text{subject to} & g^j(x, y_j) \le 0, \quad j \in J \\ & L(y) \\ & x \in \Re^n, \quad y_j \in Y_j \end{array}$$

Use branch and bound by branching on y_j

But continuous relaxations at nodes of the search tree are in general *nonconvex*.

So we formulate a convex *quasi-relaxation*.

Quasi-Relaxations

Given problem *P*: $\min\{f(x) \mid x \in S\}$

a problem *Q*: $\min\{f'(x) \mid x \in S'\}$

is a *quasi-relaxation* of *P* if:

for any feasible solution x of *P*, there is a feasible solution x' of *Q* with $f'(x') \leq f(x)$.

Thus one can obtain a valid lower bound by solving a quasi-relaxation.

Theorem. Suppose that each $g_i^j(x, y_j)$ is either

(a) convex or
(b) concave in *y_j* and homogeneous in *x*:

$$g_i^j(\alpha x, y_j) = \alpha g_i^j(x, y_j) \text{ for } \alpha \in [0,1]$$

Suppose also that $x^L \le x \le x^U, \quad y^L \le y \le y^U$

Then the following is a convex quasi-relaxation :

convex
concave/homogeneous

$$\min x_0$$
$$g_i^j\left(x,\ \alpha_j y^L + (1-\alpha_j) y^U\right) \le 0, \quad (i, j) \in J_1$$
$$g_i^j\left(x^{j1}, y_j^L\right) + g_i^j\left(x^{j2}, y_j^U\right) \le 0, \quad (i, j) \in J_2$$
$$\alpha_j x^L \le x^{j1} \le \alpha_j x^U, \quad j \in J$$
$$(1-\alpha_j) x^L \le x^{j2} \le (1-\alpha_j) x^U, \quad j \in J$$
$$x = x^{j1} + x^{j2}, \quad j \in J$$
$$x^{j1}, x^{j2} \in \Re^n, \quad \alpha_j \in [0,1], \quad j \in J$$

Why?

Take any feasible solution $(\bar{x}, \bar{y})$ of

$$\begin{array}{ll} \min & x_0 \\ \text{subject to} & g^j(x, y_j) \leq 0, \quad j \in J_1 \bigcup J_2 \\ & x \in \Re^n, \quad y_j \in Y_j \end{array}$$

To obtain a feasible solution of

$$\begin{array}{ll} \min & x_0 \\ \text{s.t.} & g_i^j\left(x,\ \alpha_j y^L + (1-\alpha_j) y^U\right) \leq 0, \quad (i,j) \in J_1 \\ & g_i^j\left(x^{j1}, y_j^L\right) + g_i^j\left(x^{j2}, y_j^U\right) \leq 0, \quad (i,j) \in J_2 \\ & \alpha_j x^L \leq x^{j1} \leq \alpha_j x^U, \quad j \in J \\ & (1-\alpha_j) x^L \leq x^{j2} \leq (1-\alpha_j) x^U, \quad j \in J \\ & x = x^{j1} + x^{j2}, \quad j \in J \\ & x^{j1}, x^{j2} \in \Re^n, \quad \alpha_j \in [0,1], \quad j \in J \end{array}$$

do the following:

choose $\alpha \in [0,1]$ so that $\bar{y}_j = \alpha_j y_j^L + (1-\alpha_j) y_j^U$

set $x^{j1} = \alpha_j \bar{x}, \quad x^{j2} = (1-\alpha_j) \bar{x}$

Then

concavity

$$\begin{aligned} g_i^j(\bar{x}, y_j) &= g_j^i\left(\bar{x}, \alpha_j y_j^L + (1-\alpha_j) y_j^U\right) \geq \alpha_j g_i^j(\bar{x}, y_j^L) + (1-\alpha_j) g_i^j(\bar{x}, y_j^U) \\ &= g_i^j\left(\alpha_j \bar{x}, y_j^L\right) + g_i^j\left((1-\alpha_j)\bar{x}, y_j^U\right) = g_i^j(x^{j1}, y_j^L) + g_i^j(x^{j2}, y_j^U) \end{aligned}$$

homogeneity

So we have a feasible solution of the quasi-relaxation with value that is less than or equal to (in fact equal to) that of the original problem.

convex, because $g^j(x,y)$ is convex

satisfied, by construction

satisfied, by above argument

$$\begin{aligned}
\min \quad & x_0 \\
\text{s.t.} \quad & g_i^j\left(x,\ \alpha_j y^L + (1-\alpha_j)y^U\right) \le 0, \quad (i,j) \in J_1 \\
& g_i^j\left(x^{j1}, y_j^L\right) + g_i^j\left(x^{j2}, y_j^U\right) \le 0, \quad (i,j) \in J_2 \\
& \alpha_j x^L \le x^{j1} \le \alpha_j x^U, \quad j \in J \\
& (1-\alpha_j)x^L \le x^{j2} \le (1-\alpha_j)x^U, \quad j \in J \\
& x = x^{j1} + x^{j2}, \quad j \in J \\
& x^{j1}, x^{j2} \in \Re^n, \quad \alpha_j \in [0,1], \quad j \in J
\end{aligned}$$

satisfied, by construction

convex, because $g^j(x,y)$ is convex in x

Solve continuous version of structural design example with quasi-relaxations

Original formulation:

$$
\begin{aligned}
\min \quad & x_0 \\
\text{s.t.} \quad & z_1 + z_2 + (x_1 + x_2)^2 + x_2^2 - x_0 \le 0 \\
& 300y_1 - z_1 \le 0 \\
& 10 - x_1 y_1 \le 0 \\
& 300y_2 - z_2 \le 0 \\
& 20 - x_2 y_2 \le 0 \\
& x_j \ge 0, \quad y_j \in \{0, 0.1, \ldots, 3.0\}
\end{aligned}
$$

Discretize

Put in proper form:

convex

$$
\begin{aligned}
\min \quad & x_0 \\
\text{s.t.} \quad & 300y_1 + 300y_2 + (x_1 + x_2)^2 + x_2^2 - x_0 \le 0 \\
& s_1 - x_1 y_1 \le 0 \\
& s_2 - x_2 y_2 \le 0 \\
& 5 \le x_1 \le 10, \quad 10 \le x_2 \le 20 \\
& 10 \le s_1 \le 10, \quad 20 \le s_2 \le 20 \\
& 0 \le y_j \le 3 \\
& y_j \in \{0, 0.1, \ldots, 3.0\}
\end{aligned}
$$

Concave in y_j & homogeneous in 1st argument (s_j, x_j)

The quasi-relaxation is:

$$
\begin{aligned}
\min \quad & x_0 \\
\text{s.t.} \quad & 300y_1 + 300y_2 + (x_1 + x_2)^2 + x_2^2 - x_0 \le 0 \\
& \left(s_{11} - x_{11}y_1^L\right) + \left(s_{12} - x_{12}y_1^U\right) \le 0 \\
& \left(s_{21} - x_{21}y_2^L\right) + \left(s_{22} - x_{22}y_2^U\right) \le 0 \\
& 5\alpha_1 \le x_{11} \le 10\alpha_1, \quad 5(1-\alpha_1) \le x_{12} \le 10(1-\alpha_1) \\
& 10\alpha_2 \le x_{21} \le 20\alpha_2, \quad 10(1-\alpha_2) \le x_{22} \le 20(1-\alpha_2) \\
& 10\alpha_1 \le s_{11} \le 10\alpha_1, \quad 10(1-\alpha_1) \le s_{12} \le 10(1-\alpha_1) \\
& 20\alpha_2 \le s_{21} \le 20\alpha_2, \quad 20(1-\alpha_2) \le s_{22} \le 20(1-\alpha_2) \\
& y_j = \alpha_j y_j^L + (1-\alpha_j) y_j^U, \quad \alpha_j \in [0,1]
\end{aligned}
$$

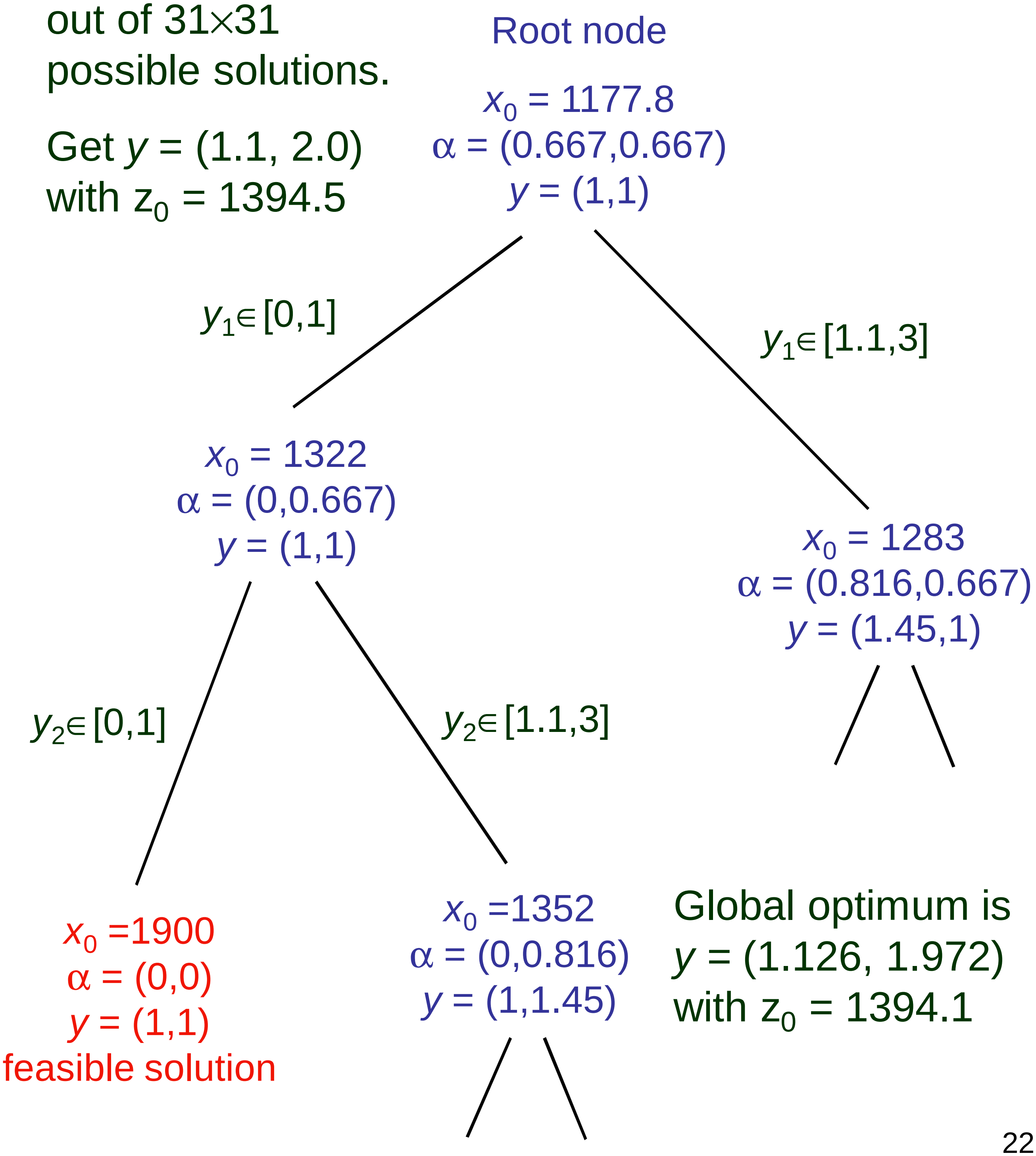
Can now re-aggregate s_j:

$$
\begin{aligned}
\min \quad & x_0 \\
\text{s.t.} \quad & 300y_1 + 300y_2 + (x_1 + x_2)^2 + x_2^2 - x_0 \le 0 \\
& 10 - x_{11}y_1^L - x_{12}y_1^U \le 0 \\
& 20 - x_{21}y_2^L - x_{22}y_2^U \le 0 \\
& 5\alpha_1 \le x_{11} \le 10\alpha_1, \quad 5(1-\alpha_1) \le x_{12} \le 10(1-\alpha_1) \\
& 10\alpha_2 \le x_{21} \le 20\alpha_2, \quad 10(1-\alpha_2) \le x_{22} \le 20(1-\alpha_2) \\
& y_j = \alpha_j y^L + (1-\alpha_j) y^U, \quad \alpha_j \in [0,1]
\end{aligned}
$$

Beginning of branch-and-bound tree

Total 63 nodes
out of 31×31
possible solutions.

Get y = (1.1, 2.0)
with z_0 = 1394.5



Root node

$x_0 = 1177.8$
$\alpha = (0.667,0.667)$
$y = (1,1)$

$y_1 \in [0,1]$

$y_1 \in [1.1,3]$

$x_0 = 1322$
$\alpha = (0,0.667)$
$y = (1,1)$

$x_0 = 1283$
$\alpha = (0.816,0.667)$
$y = (1.45,1)$

$y_2 \in [0,1]$

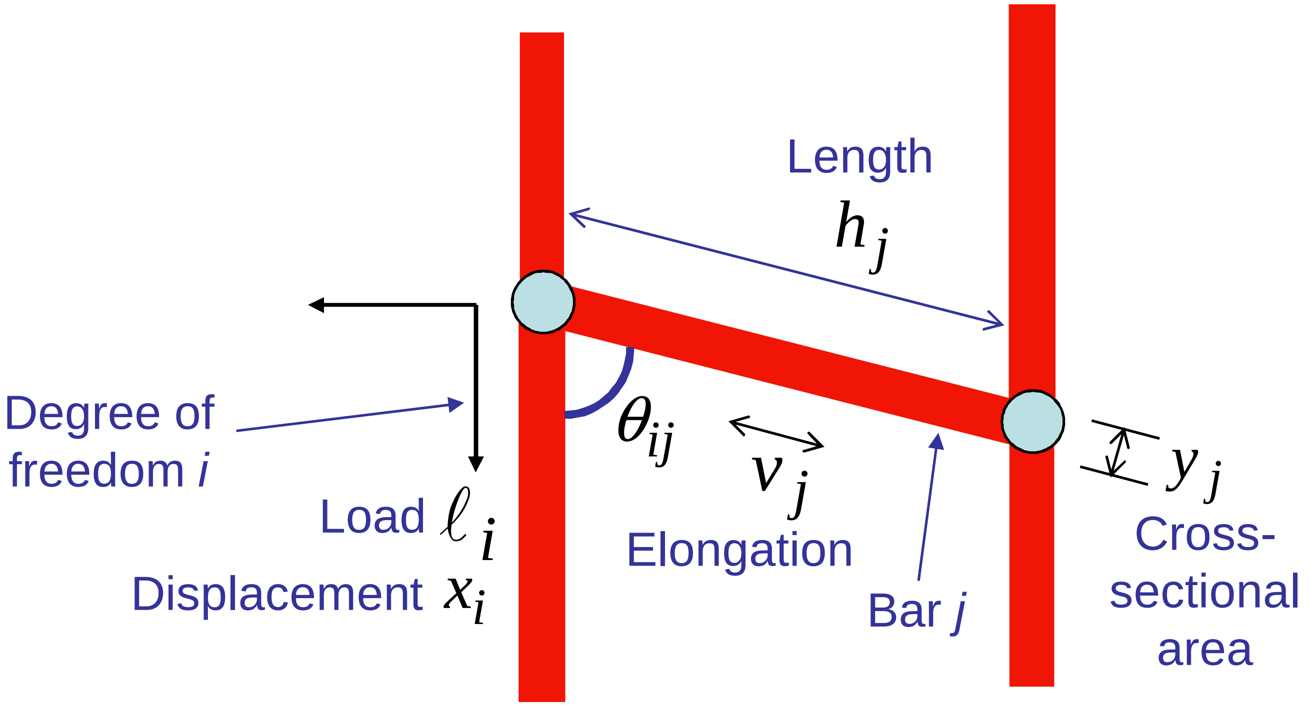
$y_2 \in [1.1,3]$

$x_0 = 1900$
$\alpha = (0,0)$
$y = (1,1)$
feasible solution

$x_0 = 1352$
$\alpha = (0,0.816)$
$y = (1,1.45)$

Global optimum is
y = (1.126, 1.972)
with z_0 = 1394.1

Realistic Structural Design Problem



$$\min \qquad \sum_j c_j h_j y_j$$

subject to

$$\frac{E_j}{h_j} y_j v_j = f_j, \quad \text{all } j$$

Hooke's law

$$\sum_j f_j \cos\theta_{ij} = \ell_i, \quad \text{all } i$$

Equilibrium

$$\sum_i x_i \cos\theta_{ij} = v_j, \quad \text{all } j$$

Compatibility

$$v_j^L \le v_j \le v_j^U, \quad \text{all } j$$

Elongation bounds

$$x_j^L \le x_j \le x_j^U, \quad \text{all } j$$

Displacement bounds

$$y_j \in Y_j$$

Solution as MILP

Ghattas and Grossmann

Discrete sizes
for bar *j*

$$
\begin{aligned}
\text{min} \quad & \sum_j c_j h_j \sum_k A_{jk} \beta_{jk} \\
\text{subject to} \quad & \frac{E_j}{h_j} A_{jk} v_{jk} = f_j, \quad \text{all } j \\
& \sum_j f_j \cos \theta_{ij} = \ell_i, \quad \text{all } i \\
& \sum_i x_i \cos \theta_{ij} = \sum_k v_{jk}, \quad \text{all } j \\
& v_j^L \beta_{jk} \le v_{jk} \le v_j^U \beta_{jk}, \quad \text{all } j, k \\
& x_j^L \le x_j \le x_j^U, \quad \text{all } j \\
& \beta_{jk} \in \{0,1\}, \ v_{jk}, f_j \in \Re
\end{aligned}
$$

Solution with convex quasi-relaxations

Bollapragada, Ghattas and Hooker

Check that the problem has the right form:

convex

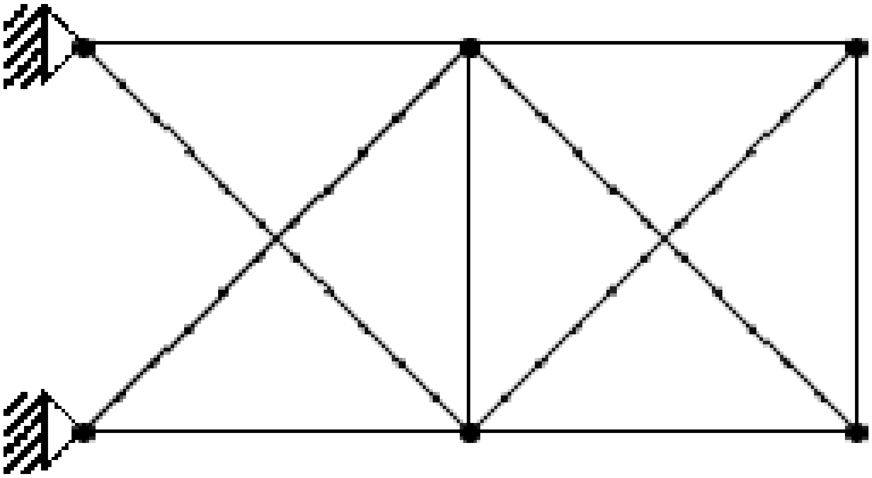
$$
\begin{array}{ll}
\min & \displaystyle\sum_j c_j h_j y_j \\
\text{subject to} & \dfrac{E_j}{h_j} y_j v_j = f_j, \quad \text{all } j \\
& \displaystyle\sum_j f_j \cos\theta_{ij} = \ell_i, \quad \text{all } i \\
& \displaystyle\sum_i x_i \cos\theta_{ij} = v_j, \quad \text{all } j \\
& v_j^L \le v_j \le v_j^U, \quad \text{all } j \\
& x_j^L \le x_j \le x_j^U, \quad \text{all } j \\
& y_j \in Y_j
\end{array}
$$

Concave (linear) in y_j and homogeneous in (v_j, f_j)

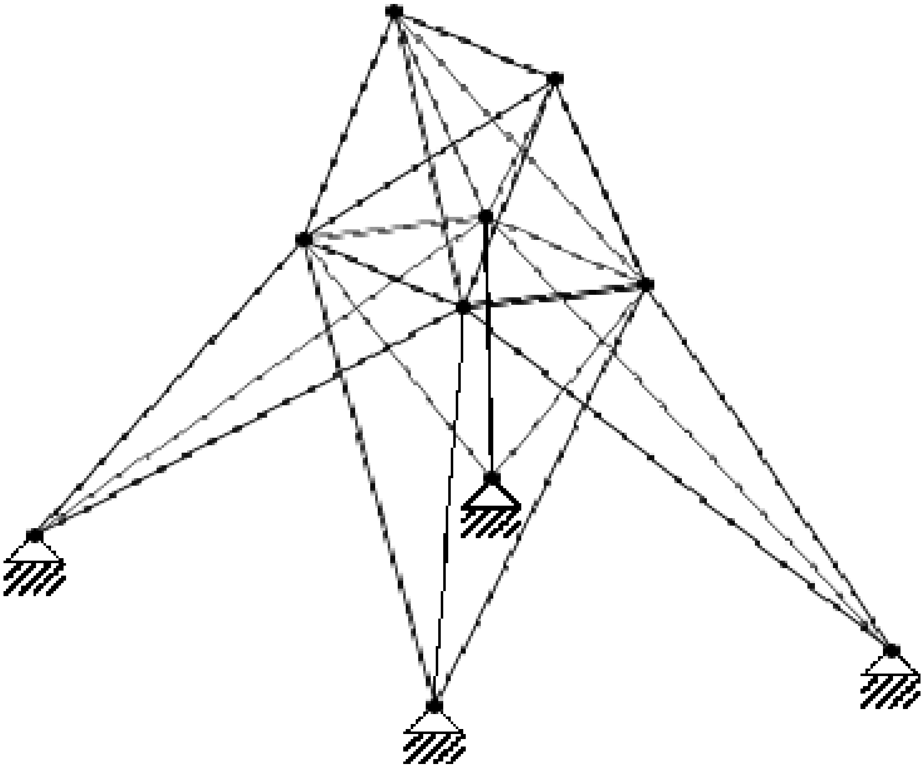
The quasi-relaxation is

$$
\begin{array}{ll}
\min & \sum_j c_j h_j y_j \\
\text{subject to} & \frac{E_j}{h_j}\left(y_j^L v_{j1} + y_j^U v_{j2}\right) = f_j, \quad \text{all } j \\
& \sum_j f_j \cos\theta_{ij} = \ell_i, \quad \text{all } i \\
& \sum_i x_i \cos\theta_{ij} = v_j, \quad \text{all } j \\
& \alpha_j v_j^L \le v_{j1} \le \alpha_j v_j^U, \quad \text{all } j \\
& (1-\alpha_j) v_j^L \le v_{j2} \le (1-\alpha_j) v_j^U, \quad \text{all } j \\
& x_j^L \le x_j \le x_j^U, \quad \text{all } j \\
& v_j = v_{j1} + v_{j2}, \quad \text{all } j \\
& y_j = \alpha_j y_j^L + (1-\alpha_j) y_j^U, \quad \text{all } j \\
& \alpha_j \in [0,1]
\end{array}
$$

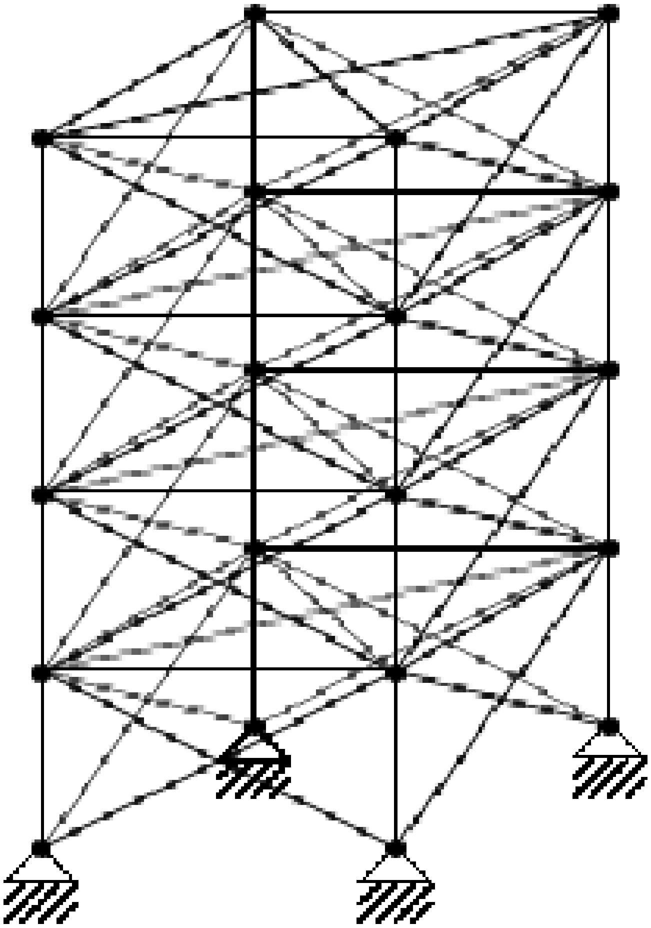
Some problem instances



10-bar cantilever truss



25-bar electrical transmission tower



72-bar building

Use symmetries to help solve problem

Computational Results (seconds)

Problem Instance		**MILP (CPLEX)**	**Quasi-relaxation**
10-bar cantilever truss 11 discrete bar sizes	1 load	1.3	0.3
	1 load, wider stress bounds	1.6	0.3
	1 load, wider stress bounds	2.6	1.2
	1 load, wider stress bounds	2.6	1.4
	2 loads	23.6	5.8
	1 load + displacement bounds	1089.4	67.5
	2 loads + displacement bounds	13743.9	1654.0
25-bar transmission tower	2 loads, 11 discrete areas	271.7	225.8
Building 11 discrete bar sizes 2 loads	72 bars	12692.7	207.9
	90 bars	*	168.9
	108 bars	*	329.4

*No solution after 20 hours (72,000 seconds)