The Branch-and-Sandwich Algorithm for Mixed-Integer Nonlinear Bilevel Problems

Polyxeni-M. Kleniati and Claire S. Adjiman

MINLP Workshop

June 2, 2014, CMU

Funding Bodies
EPSRC & LEVERHULME TRUST

OUTLINE

Outline

Bilevel Optimisation Problem

Problem Formulation

- A two-person, non-cooperative game in which the play is sequential
- The **M**ixed-**I**nteger **N**onlinear **B**ilevel **P**roblem is

$$
\begin{array}{cc}
\min\limits_{x_i,x_c,y_i,y_c} & F(x_i,x_c,y_i,y_c) \\
\text{s.t.} & G(x_i,x_c,y_i,y_c) \leq 0 \\
 & (y_i,y_c) \in \underset{y_i\in Y_I,y_c\in Y_C}{\arg\min} \{f(x_i,x_c,y_i,y_c) \text{ s.t. } g(x_i,x_c,y_i,y_c) \leq 0\} \\
 & x_i \in X_I \subset \mathbb{Z}^{n_1}, x_c \in X_C \subset \mathbb{R}^{n-n_1} \\
 & y_i \in Y_I \subset \mathbb{Z}^{m_1}, y_c \in Y_C \subset \mathbb{R}^{m-m_1}
\end{array}
\qquad \text{(MINBP)}
$$

- Subscripts i and c stand for *integer* and *continuous*, respectively
- Typical assumptions apply, such as
  - continuity of all functions and compactness of the host sets
- Assume also twice differentiability of the continuous relaxations of the functions
- No special class or convexity assumptions are made

Optimal Value Equivalent Reformulation

- Define the inner optimal value function

$$w(x_i, x_c) = \min_{y_i, y_c}\{f(x_i, x_c, y_i, y_c) \text{ s.t. } g(x_i, x_c, y_i, y_c) \leq 0, y_i \in Y_I, y_c \in Y_C\}$$

- Then problem (e.g. Dempe and Zemkoho, 2011):

$$\begin{array}{cc} \min\limits_{x_i, x_c, y_i, y_c} & F(x_i, x_c, y_i, y_c) \\ \text{s.t.} & G(x_i, x_c, y_i, y_c) \leq 0 \\ & g(x_i, x_c, y_i, y_c) \leq 0 \\ & f(x_i, x_c, y_i, y_c) \leq w(x_i, x_c) \\ & x_i \in X_I, x_c \in X_C \\ & y_i \in Y_I, y_c \in Y_C \end{array}$$

is equivalent to MINBP without any convexity assumptions whatsoever

Useful Property

A restriction of the inner problem yields a relaxation of the overall problem and vice versa (Mitsos and Barton, 2006)

Optimal Value Equivalent Reformulation

- Define the inner optimal value function

$$w(x_i, x_c) = \min_{y_i, y_c} \{ f(x_i, x_c, y_i, y_c) \text{ s.t. } g(x_i, x_c, y_i, y_c) \leq 0, y_i \in Y_I, y_c \in Y_C \}$$

- Then problem (e.g. Dempe and Zemkoho, 2011):

$$\begin{array}{cc} \min\limits_{x_i, x_c, y_i, y_c} & F(x_i, x_c, y_i, y_c) \\ \text{s.t.} & G(x_i, x_c, y_i, y_c) \leq 0 \\ & g(x_i, x_c, y_i, y_c) \leq 0 \\ & f(x_i, x_c, y_i, y_c) \leq w(x_i, x_c) \\ & x_i \in X_I, x_c \in X_C \\ & y_i \in Y_I, y_c \in Y_C \end{array}$$

is equivalent to MINBP without any convexity assumptions whatsoever

Useful Property

A restriction of the inner problem yields a relaxation of the overall problem and vice versa (Mitsos and Barton, 2006)

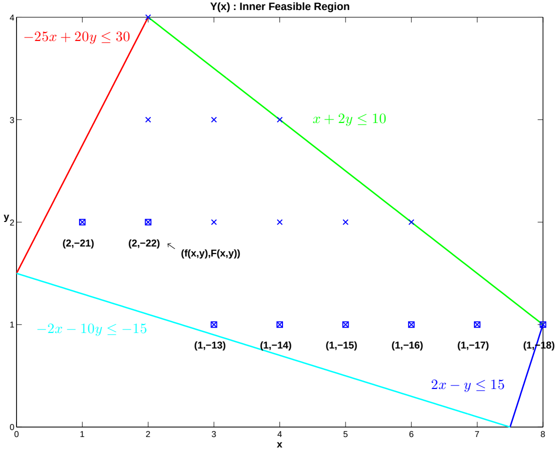
PROPOSED APPROACHES FOR DISCRETE BILEVEL PROGRAMMING

- INFLUENTIAL WORKS BY BARD ET AL., DEMPE ET AL. AND VICENTE ET AL., e.g.
  1. *The Mixed Integer Linear Bilevel Programming Problem*, Moore & Bard (1990)
  2. *An Algorithm for the Mixed-Integer Nonlinear Bilevel Program. Prob.*, Edmunds & Bard (1992)
  3. *The Discrete Linear Bilevel Programming Problem*, Vicente, Savard & Judice (1996)
  4. *Practical Bilevel Optimization*, Bard (1998)
  5. *Foundations of Bilevel Programming*, Dempe (2002)
  6. *Discrete Bilevel Programming*, Dempe, Kalashnikov & Ríos-Mercado (2005)
  7. *Bilevel programming with discrete lower level problems*, Fanghänel & Dempe (2009)

- ADVANCES ON MORE GENERAL CLASSES
  1. *Global Optimization of Mixed-Integer Bilevel Program. Prob.*, Gümüş and Floudas (2005)
  2. *Global Solution of Nonlinear Mixed-Integer Bilevel Programs*, Mitsos (2010)

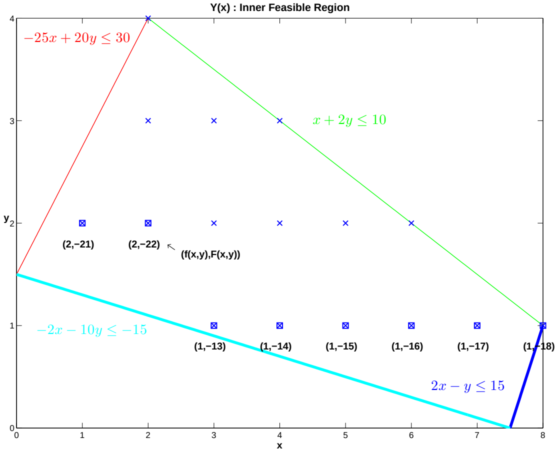
Challenges : Illustrative Example (Moore and Bard, 1990)

$$\begin{aligned} \min_{x\in[0,8]} \quad & -x - 10y \\ \text{s.t.} \quad & x \in \mathbb{Z} \\ & y \in \arg\min_{y\in[0,4]}\{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\} \end{aligned}$$



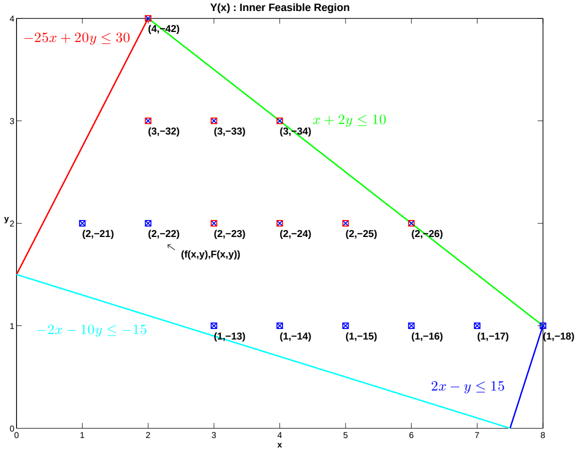
Challenges : Illustrative Example (Moore and Bard, 1990)

$$\begin{aligned} \min_{x\in[0,8]} \quad & -x-10y \\ \text{s.t.} \quad & x \in \mathbb{Z} \\ & y \in \arg\min_{y\in[0,4]}\{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\} \end{aligned}$$



Challenge 1 : Constructing a Convergent Lower Bounding Problem

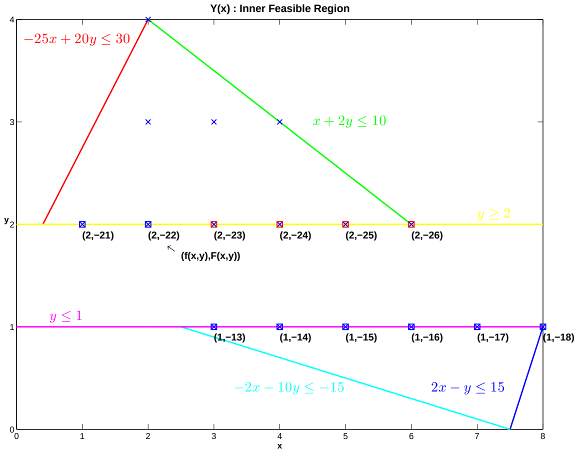
$$
\begin{array}{ll}
\min\limits_{x,y} & -x - 10y \\
\text{s.t.} & y \in Y(x) \\
 & x \in [0,8], y \in [0,4] \\
 & x, y \in \mathbb{Z}
\end{array}
$$



Challenge 2 : Partitioning the Inner Region

$$\min_{x\in[0,8]} \quad -x-10y \qquad \text{s.t.} \quad x\in\mathbb{Z}, \quad y\in \arg\min_{y\in[0,1]}\{y \text{ s.t. } y\in Y(x), y\in\mathbb{Z}\}$$

$$\min_{x\in[0,8]} \quad -x-10y \qquad \text{s.t.} \quad x\in\mathbb{Z}, \quad y\in \arg\min_{y\in[2,4]}\{y \text{ s.t. } y\in Y(x), y\in\mathbb{Z}\}$$



OUTLINE

Branch-and-Sandwich Algorithm[1,2]

Connection to earlier work

- A deterministic global optimisation algorithm for bilevel programs with
  1. twice continuously differentiable functions
  2. continuous decision variables
  3. nonconvex inner problem satisfying regularity
- The Branch-and-Sandwich algorithm was proved to be ε-convergent based on
  - exhaustiveness and the general convergence theory (Horst and Tuy, 1996)
- Branch-and-Sandwich was tested on 33 small problems with promising numerical results

[1] Kleniati, P. M. and Adjiman, C. S., 2011, Proceedings of the 21st European Symposium on Computer-Aided Process Engineering, Computer-Aided Chemical Engineering, 29, 602 – 606.

[2] ______ , 2014, Parts I & II, JOGO, DOI 10.1007/s10898-013-0121-7; DOI 10.1007/s10898-013-0120-8

We extend the Branch-and-Sandwich Algorithm

To tackle Mixed-Integer Nonlinear Bilevel Problems

Convergent lower bounding problem

- Employ the optimal value reformulation and replace the inner optimal value function $w(x_i, x_c)$ by a constant upper bound
  - the relaxed problem has one constraint (*cut*) only
  - the right hand side of the cut is updated when appropriate

Partitioning the Inner Region

- Branching scheme that allows branching on the inner variables
  - consider all inner subregions where (inner) global optima may lie

Summary of Features

- Generation of two sets of upper and lower bounds for the inner and outer objective values
  - the resulting bounding problems are MINLPs
- Tree management with auxiliary lists of nodes as well as inner and outer fathoming rules
  - exploration of two decision spaces using a single branch-and-bound tree

Outline

Inner Problem Bounding Scheme

Inner lower bound – Consider no branching yet

- The inner bounding scheme is based on finding valid lower and upper bounds on $w(x_i, x_c)$ for all values of the outer vector variable (x_i, x_c)
- The inner lower bounding problem is

$$
\begin{aligned}
f^{\mathrm{L}} = \min_{x_i, x_c, y_i, y_c} \quad & f(x_i, x_c, y_i, y_c) \\
\text{s.t.} \quad & g(x_i, x_c, y_i, y_c) \leq 0 \\
& x_i \in X_I, x_c \in X_C, y_i \in Y_I, y_c \in Y_C
\end{aligned}
$$

- A convex relaxation can be derived for problem class considered
  - use of techniques suitable for MINLPs, e.g. Adjiman et al. (2000), Tawarmalani and Sahinidis (2004)

$$
\begin{aligned}
\underline{f} = \min_{x_i, x_c, y_i, y_c} \quad & \breve{f}(x_i, x_c, y_i, y_c) \\
\text{s.t.} \quad & \breve{g}(x_i, x_c, y_i, y_c) \leq 0 \\
& x_i \in X_I, x_c \in X_C, y_i \in Y_I, y_c \in Y_C
\end{aligned}
$$

Inner Problem Bounding Scheme

Inner upper bound – Consider no branching yet

- The inner bounding scheme is based on finding valid lower and upper bounds on $w(x_i, x_c)$ for all values of the outer vector variable (x_i, x_c)
- The inner upper bounding problem is the *robust counterpart* of the inner problem

$$
\begin{aligned}
f^{\mathrm{U}} = \quad & \min_{y_i, y_c, t} && t \\
& \text{s.t.} && f(x_i, x_c, y_i, y_c) \le t \quad \forall (x_i, x_c) \in X_I \times X_C \\
& && g(x_i, x_c, y_i, y_c) \le 0 \quad \forall (x_i, x_c) \in X_I \times X_C \\
& && y_i \in Y_I, y_c \in Y_C
\end{aligned}
$$

- An upper bound can be derived by using SIP techniques, e.g. Bhattacharjee et al. (2005), & interval arithmetic extended to mixed-integer problems, e.g. Apt and Zoeteweij (2007), Berger and Granvilliers (2009)

$$
\begin{aligned}
\bar{f} = \quad & \min_{y_i, y_c, t} && t \\
& \text{s.t.} && \bar{f}(X_I, X_C, y_i, y_c) \le t \\
& && \bar{g}(X_I, X_C, y_i, y_c) \le 0 \\
& && y_i \in Y_I, y_c \in Y_C
\end{aligned}
$$

Inner Problem Bounding Scheme (cont.)

Motivation : we must have valid inner bounds over the current subdomain

Figure: Inner lower bound

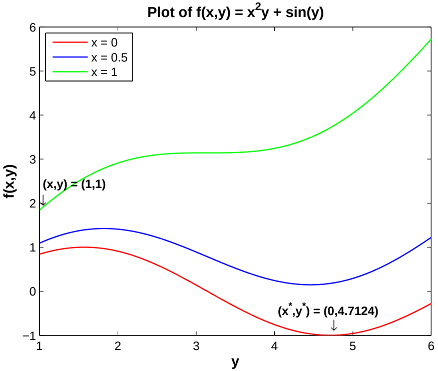
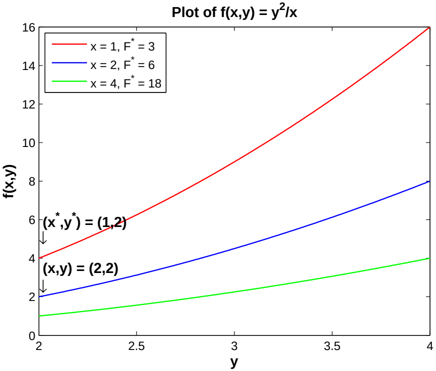


Figure: Inner upper bound



Outer Problem Bounding Scheme

Consider no branching yet

- The proposed lower bounding problem is

$$\begin{aligned} \min_{x_i, x_c, y_i, y_c} \quad & F(x_i, x_c, y_i, y_c) \\ \text{s.t.} \quad & G(x_i, x_c, y_i, y_c) \leq 0 \\ & g(x_i, x_c, y_i, y_c) \leq 0 \\ & f(x_i, x_c, y_i, y_c) \leq \bar{f} \\ & x_i \in X_I, x_c \in X_C, y_i \in Y_I, y_c \in Y_C \end{aligned}$$

  - to tighten, add the inner KKT conditions with respect to the continuous inner variables y_c
    - ★ based on regularity being satisfied for all the parameter values
  - any feasible solution in the (MINBP) is feasible in the proposed relaxation
- For $(x_i, x_c) = (\bar{x}_i, \bar{x}_c)$, the upper bounding problem is (Mitsos et al., 2008)

$$\begin{aligned} \min_{y_i, y_c} \quad & F(\bar{x}_i, \bar{x}_c, y_i, y_c) \\ \text{s.t.} \quad & G(\bar{x}_i, \bar{x}_c, y_i, y_c) \leq 0 \\ & g(\bar{x}_i, \bar{x}_c, y_i, y_c) \leq 0 \\ & f(\bar{x}_i, \bar{x}_c, y_i, y_c) \leq w(\bar{x}_i, \bar{x}_c) + \varepsilon_f \\ & y_i \in Y_I, y_c \in Y_C \end{aligned}$$

  - any feasible solution in the upper bounding problem is feasible in the (MINBP)
  - set $F^{\text{UB}} := \min\{\bar{F}, \infty\}$ to express the incumbent

Outer Problem Bounding Scheme

Consider no branching yet

- The proposed lower bounding problem is

$$\begin{aligned} \min_{x_i, x_c, y_i, y_c} \quad & F(x_i, x_c, y_i, y_c) \\ \text{s.t.} \quad & G(x_i, x_c, y_i, y_c) \leq 0 \\ & g(x_i, x_c, y_i, y_c) \leq 0 \\ & f(x_i, x_c, y_i, y_c) \leq \bar{f} \\ & x_i \in X_I, x_c \in X_C, y_i \in Y_I, y_c \in Y_C \end{aligned}$$

  - to tighten, add the inner KKT conditions with respect to the continuous inner variables y_c
    - based on regularity being satisfied for all the parameter values
  - any feasible solution in the (MINBP) is feasible in the proposed relaxation
- For $(x_i, x_c) = (\bar{x}_i, \bar{x}_c)$, the upper bounding problem is (Mitsos et al., 2008)

$$\begin{aligned} \min_{y_i, y_c} \quad & F(\bar{x}_i, \bar{x}_c, y_i, y_c) \\ \text{s.t.} \quad & G(\bar{x}_i, \bar{x}_c, y_i, y_c) \leq 0 \\ & g(\bar{x}_i, \bar{x}_c, y_i, y_c) \leq 0 \\ & f(\bar{x}_i, \bar{x}_c, y_i, y_c) \leq w(\bar{x}_i, \bar{x}_c) + \varepsilon_f \\ & y_i \in Y_I, y_c \in Y_C \end{aligned}$$

  - any feasible solution in the upper bounding problem is feasible in the (MINBP)
  - set $F^{\text{UB}} := \min\{\bar{F}, \infty\}$ to express the incumbent

Outline

List Management

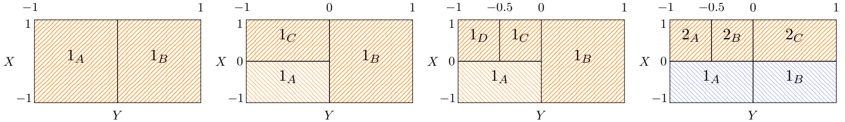
Auxiliary lists

$\mathcal{L}$: is the classical list of 'open' nodes, corresponding to the outer problem

$\mathcal{L}^{\mathrm{I}}$: is the list of exclusively inner 'open' nodes

$\mathcal{L}_{\mathcal{X}_p}$: is the list of (outer & inner) nodes that cover the whole Y for a subdomain $\mathcal{X}_p$ of X, $1 \leq p \leq n_{\mathrm{p}}$

- Number n_{p} equals the number of partition sets $\mathcal{X}_p \subseteq X$ s.t. :
  - for each subdomain $\mathcal{X}_p$, the *"whole"* Y is maintained
- Two lists $\mathcal{L}_{\mathcal{X}_{p_1}}, \mathcal{L}_{\mathcal{X}_{p_2}}$ are called independent if $\mathcal{L}_{\mathcal{X}_{p_1}} \cap \mathcal{L}_{\mathcal{X}_{p_2}} = \emptyset$
  - two lists with common nodes, e.g. $\mathcal{S}^1_{\mathcal{X}_p}$ and $\mathcal{S}^2_{\mathcal{X}_p}$, are sublists of $\mathcal{L}_{\mathcal{X}_p}$



(a) $\mathcal{L}_{[-1,1]}$ (b) $\mathcal{L}_{[-1,1]}$ (c) $\mathcal{L}_{[-1,1]}$ (d) $\mathcal{L}_{[-1,0]}, \mathcal{L}_{[0,1]}$

- For each p, best inner upper bound lowest over Y, but largest over $\mathcal{X}_p$:

$$f^{\mathrm{UB}}_{\mathcal{X}_p} = \max\{\min_{j \in \mathcal{S}^1_{\mathcal{X}_p}} \{\bar{f}^{(j)}\}, \ldots, \min_{j \in \mathcal{S}^s_{\mathcal{X}_p}} \{\bar{f}^{(j)}\}\}.$$

MODIFIED BOUNDING PROBLEMS

ALLOW BRANCHING & CONSIDER A NODE $k \in \mathcal{L}_{\mathcal{X}_p}$

OUTER UPPER BOUND

$$k' := \arg\min_{j \in \mathcal{L}_{\mathcal{X}_p}} w^{(j)}(\bar{x}_i, \bar{x}_c)$$

$$\begin{array}{rcl} \bar{F}^{(k')} = & \displaystyle\min_{y_i, y_c} & F(\bar{x}_i, \bar{x}_c, y_i, y_c) \\ & \text{s.t.} & G(\bar{x}_i, \bar{x}_c, y_i, y_c) \leq 0 \\ & & g(\bar{x}_i, \bar{x}_c, y_i, y_c) \leq 0 \\ & & f(\bar{x}_i, \bar{x}_c, y_i, y_c) \leq w^{(k')}(\bar{x}_i, \bar{x}_c) + \varepsilon_f \\ & & y_i \in Y_I^{(k')}, y_c \in Y_C^{(k')} \end{array}$$

INNER LOWER BOUND

$$\begin{array}{rcl} \underline{f}^{(k)} = & \displaystyle\min_{x_i, x_c, y_i, y_c} & \breve{f}^{(k)}(x_i, x_c, y_i, y_c) \\ & \text{s.t.} & \breve{g}^{(k)}(x_i, x_c, y_i, y_c) \leq 0 \\ & & x_i \in X_I^{(k)}, x_c \in X_C^{(k)} \\ & & y_i \in Y_I^{(k)}, y_c \in Y_C^{(k)} \end{array}$$

OUTER LOWER BOUND

$$\begin{array}{rcl} \underline{F}^{(k)} = & \displaystyle\min_{x_i, x_c, y_i, y_c} & F(x_i, x_c, y_i, y_c) \\ & \text{s.t.} & G(x_i, x_c, y_i, y_c) \leq 0 \\ & & g(x_i, x_c, y_i, y_c) \leq 0 \\ & & f(x_i, x_c, y_i, y_c) \leq f^{\mathrm{UB}}_{\mathcal{X}_p} \\ & & x_i \in X_I^{(k)}, x_c \in X_C^{(k)} \\ & & y_i \in Y_I^{(k)}, y_c \in Y_C^{(k)} \end{array}$$

INNER UPPER BOUND

$$\begin{array}{rcl} \bar{f}^{(k)} = & \displaystyle\min_{y_i, y_c, t} & t \\ & \text{s.t.} & \bar{f}(X_I^{(k)}, X_C^{(k)}, y_i, y_c) \leq t \\ & & \bar{g}(X_I^{(k)}, X_C^{(k)}, y_i, y_c) \leq 0 \\ & & y_i \in Y_I^{(k)}, y_c \in Y_C^{(k)} \end{array}$$

Node Fathoming Rules

Node $k \in \mathcal{L} \cap \mathcal{L}_{\mathcal{X}_p}$

Inner fathoming rules

If
1. $\underline{f}^{(k)} = \infty$ or
2. $\underline{f}^{(k)} > f^{\mathrm{UB}}_{\mathcal{X}_p}$

then fathom, i.e. delete from $\mathcal{L}$ (or $\mathcal{L}^{\mathrm{I}}$) and $\mathcal{L}_{\mathcal{X}_p}$.

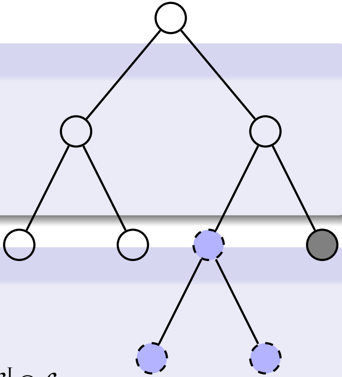
Outer fathoming rules

If
1. $\underline{F}^{(k)} = \infty$ or
2. $\underline{F}^{(k)} \geq F^{\mathrm{UB}} - \varepsilon_F$

then Outer Fathom, i.e. move from $\mathcal{L}$ to $\mathcal{L}^{\mathrm{I}}$. Hence, $k \in \mathcal{L}^{\mathrm{I}} \cap \mathcal{L}_{\mathcal{X}_p}$.

Node Fathoming Rules

Node $k \in \mathcal{L} \cap \mathcal{L}_{\mathcal{X}_p}$



Inner fathoming rules

If
1. $\underline{f}^{(k)} = \infty$ or
2. $\underline{f}^{(k)} > f^{\mathrm{UB}}_{\mathcal{X}_p}$

then fathom, i.e. delete from $\mathcal{L}$ (or $\mathcal{L}^{\mathrm{I}}$) and $\mathcal{L}_{\mathcal{X}_p}$.

Outer fathoming rules

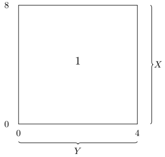
If
1. $\underline{F}^{(k)} = \infty$ or
2. $\underline{F}^{(k)} \geq F^{\mathrm{UB}} - \varepsilon_F$

then Outer Fathom, i.e. move from $\mathcal{L}$ to $\mathcal{L}^{\mathrm{I}}$. Hence, $k \in \mathcal{L}^{\mathrm{I}} \cap \mathcal{L}_{\mathcal{X}_p}$.

OUTLINE

Illustrative Example Revisited : Extended-Tree Version

$\min_{x \in [0,8]} \{-x - 10y \text{ s.t. } x \in \mathbb{Z},\ y \in \arg\min_{y \in [0,4]}\{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\}\}$ WITH $F^* = -22$ AT $(x^*, y^*) = (2, 2)$

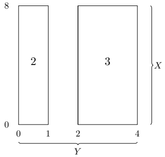


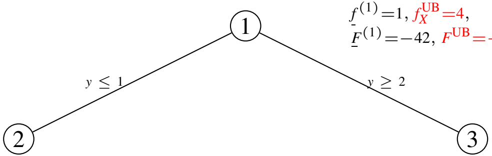
1

$\underline{f}^{(1)}=1,\ \bar{f}^{(1)}=4,$
$\underline{F}^{(1)}=-42,\ \bar{x}^{(1)}=2,$
$w(\bar{x})=2,\ \bar{F}^{(1)}=-22$

ILLUSTRATIVE EXAMPLE REVISITED : EXTENDED-TREE VERSION

$$\min_{x\in[0,8]} \{-x - 10y \text{ s.t. } x \in \mathbb{Z},\ y \in \arg\min_{y\in[0,4]}\{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\}\} \text{ WITH } F^* = -22 \text{ AT } (x^*, y^*) = (2, 2)$$





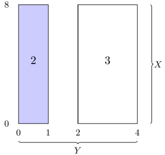
$\underline{f}^{(1)}=1$, $f_X^{\text{UB}}=4$,
$\underline{F}^{(1)}=-42$, $F^{\text{UB}}=-22$

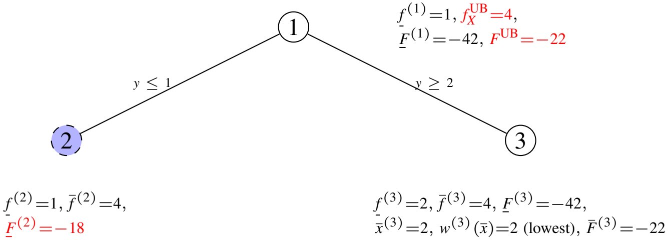
$\underline{f}^{(2)}=1$, $\bar{f}^{(2)}=4$,
$\underline{F}^{(2)}=-18$

$\underline{f}^{(3)}=2$, $\bar{f}^{(3)}=4$, $\underline{F}^{(3)}=-42$,
$\bar{x}^{(3)}=2$, $w^{(3)}(\bar{x})=2$ (lowest), $\bar{F}^{(3)}=-22$

ILLUSTRATIVE EXAMPLE REVISITED : EXTENDED-TREE VERSION

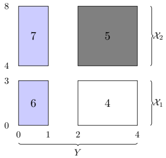
$\min\limits_{x\in[0,8]} \{-x - 10y \text{ s.t. } x \in \mathbb{Z},\ y \in \arg\min\limits_{y\in[0,4]}\{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\}\}$ WITH $F^* = -22$ AT $(x^*, y^*) = (2, 2)$

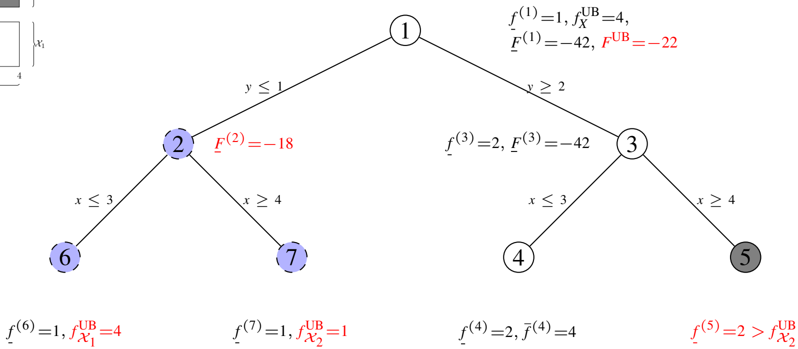




ILLUSTRATIVE EXAMPLE REVISITED : EXTENDED-TREE VERSION

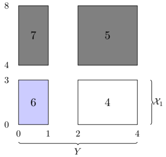
$\min\limits_{x\in[0,8]} \{-x - 10y \text{ s.t. } x \in \mathbb{Z},\ y \in \arg\min\limits_{y\in[0,4]}\{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\}\}$ WITH $F^* = -22$ AT $(x^*, y^*) = (2, 2)$

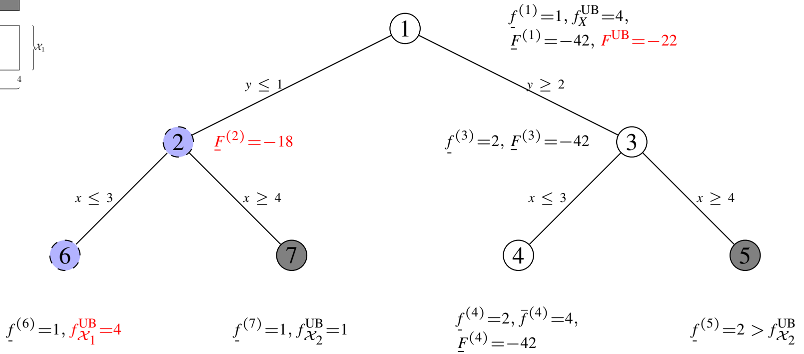




ILLUSTRATIVE EXAMPLE REVISITED : EXTENDED-TREE VERSION

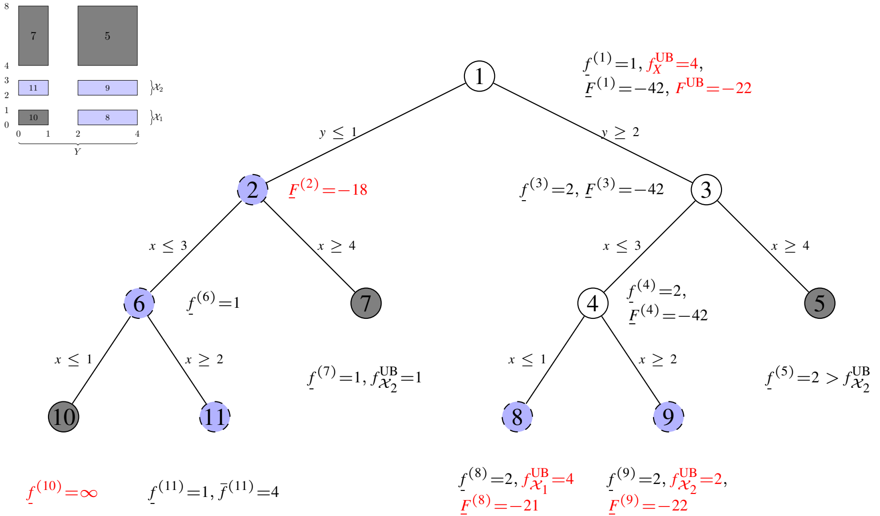
$\min_{x \in [0,8]} \{-x - 10y \text{ s.t. } x \in \mathbb{Z},\ y \in \arg\min_{y \in [0,4]}\{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\}\}$ WITH $F^* = -22$ AT $(x^*, y^*) = (2, 2)$





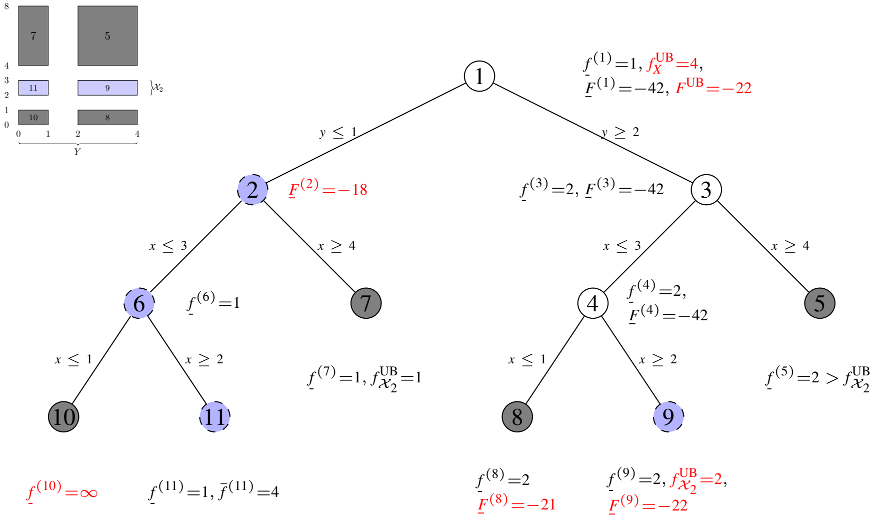
ILLUSTRATIVE EXAMPLE REVISITED : EXTENDED-TREE VERSION

$\min_{x \in [0,8]} \{-x - 10y \text{ s.t. } x \in \mathbb{Z},\ y \in \arg\min_{y \in [0,4]}\{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\}\}$ WITH $F^* = -22$ AT $(x^*, y^*) = (2, 2)$



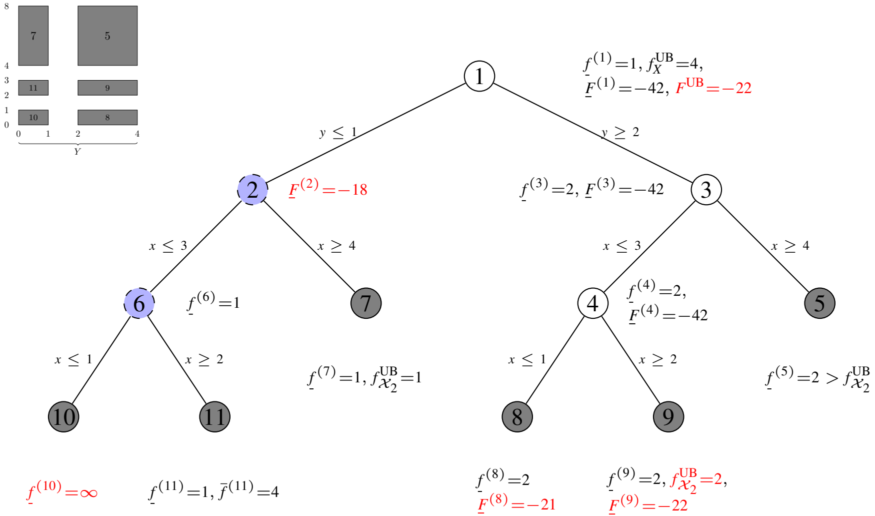
ILLUSTRATIVE EXAMPLE REVISITED : EXTENDED-TREE VERSION

$\min\limits_{x\in[0,8]} \{-x - 10y \text{ s.t. } x \in \mathbb{Z},\ y \in \arg\min\limits_{y\in[0,4]}\{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\}\}$ WITH $F^* = -22$ AT $(x^*, y^*) = (2, 2)$



Illustrative Example Revisited : Extended-Tree Version

$$\min_{x\in[0,8]} \{-x - 10y \text{ s.t. } x \in \mathbb{Z},\ y \in \arg\min_{y\in[0,4]}\{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\}\} \text{ WITH } F^* = -22 \text{ AT } (x^*, y^*) = (2, 2)$$



PRELIMINARY NUMERICAL RESULTS WITH $\varepsilon_f = 10^{-5}$ AND $\varepsilon_F = 10^{-3}$

			Number of variables					
No.	**Source**	**Problem**	$\mathbf{x}_i$	$\mathbf{x_c}$	$\mathbf{y}_i$	$\mathbf{y_c}$	F^{UB}	**Nodes**
1	**Moore & Bard (1990)**	**Example 1**	1	0	1	0	-22	7
2	**Moore & Bard (1990)**	**Example 2**	1	0	1	0	5	13
3	**Edmunds & Bard (1992)**	**Equation 3**	0	1	1	0	$\frac{4}{9}$	1
4	**Sahin & Ciric (1998)**	**Example 4**	0	2	2	0	-400	1
5	**Dempe (2002)**	**Equation 8.11**	0	2	2	0	-10.4	3
6	**Mitsos (2010)**	**am_1_0_0_1_01**	0	1	1	0	-1	1
7	**Mitsos (2010)**	**am_1_1_1_0_01**	1	1	0	1	0.5	11
8	**Mitsos (2010)**	**am_1_1_1_1_01**	1	1	1	1	-1	13
9	**Mitsos (2010)**	**am_1_1_1_1_02**	1	1	1	1	0.209	1
10	**Mitsos (2010)**	**am_3_3_3_3_01**	3	3	3	3	-2.5	1

Outline

Concluding Remarks

- Branch-and-Sandwich is a deterministic global optimisation algorithm
  - that can be applied to mixed-integer nonlinear bilevel problems
- Key features :
  1. encompasses implicitly two branch-and-bound trees
  2. introduces simple bounding problems, always obtained from the bounding problems of the parent node
  3. allows branch-and-bound with respect to x and y, but at the same time it keeps track of the partitioning of Y for successively refined subdomains of X
- Performance is linked to the tightness of the inner upper bounds $f^{\mathrm{UB}}_{\mathcal{X}_p}$
- Numerical results appear promising
- Implementation & computational experience to investigate
  - alternative choices in the way each step of the proposed algorithm is performed
  - different branching strategies

ACKNOWLEDGMENTS

The Leverhulme Trust
www.leverhulme.ac.uk

EPSRC
Pioneering research and skills

PROBLEM FORMULATION : CONTINUOUS CASE

- The optimistic bilevel problem is a LEADER-follower game
- The leader *(outer)* problem is:

$$\min_{x,y} F(x,y) \text{ s.t. } G(x,y) \leq 0,\ (x,y) \in X \times Y,\ y \in \mathcal{Y}(x) \tag{BPP}$$

- $\mathcal{Y}(x)$ is the global optimal solution set of the follower *(inner)* problem:

$$\mathcal{Y}(x) = \arg\min_{y \in Y} f(x,y) \mid g(x,y) \leq 0$$

- Common assumptions should apply, such as continuity of all functions and compactness of X and Y
- Assume also twice differentiability of all functions
- For the inner problem, assume constraint qualifications
- No convexity assumption is made

Constraint Qualification for the Inner Problem

Recall the inner problem $\min_{y \in Y}\{f(x,y) \mid g(x,y) \leq 0\}$

- Assume that a constraint qualification holds for all values of x
- Regularity ensures that the KKT conditions can be employed and are necessary
- If we replace $y \in Y$ by the corresponding bound constraints

$$y^{\mathrm{L}} \leq y \leq y^{\mathrm{U}},$$

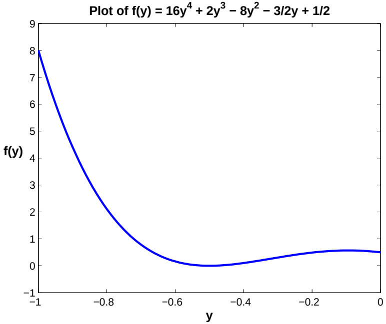
  the KKT conditions of the inner problem define the set below:

$$\Omega_{\mathrm{KKT}} = \left\{ (x,y,\mu,\lambda,\nu) \,\middle|\, \begin{array}{c} \nabla_y f(x,y) + \mu \nabla_y g(x,y) - \lambda + \nu = 0, \\ \mu^{\mathrm{T}} g(x,y) = 0,\ \mu \geq 0, \\ \lambda^{\mathrm{T}}(y^{\mathrm{L}} - y) = 0,\ \lambda \geq 0, \\ \nu^{\mathrm{T}}(y - y^{\mathrm{U}}) = 0,\ \nu \geq 0. \end{array} \right\}.$$

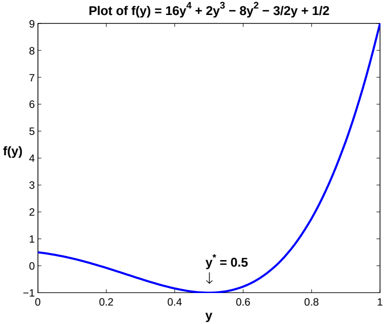
- Ω_{KKT} contains all points satisfying the KKT conditions of the inner problem
- If the inner problem is convex with a unique optimal solution for all values of x
  - the KKT conditions are also sufficient

Motivating Example (Mitsos and Barton, 2010)

$$\min_{y\in[-1,1]} y \text{ s.t. } y \in \arg\min_{y\in[-1,1]} 16y^4 + 2y^3 - 8y^2 - 3/2y + 1/2$$



(e) $-1 \leq y \leq 0$



(f) $0 \leq y \leq 1$

Inner Problem Bounding Scheme

Consider no branching yet

- The auxiliary relaxed inner problem is:

$$f^{\mathrm{L}} = \min_{x \in X, y \in Y} f(x,y) \text{ s.t. } g(x,y) \leq 0$$

- The auxiliary restricted inner problem is:

$$f^{\mathrm{U}} = \max_{x \in X} \min_{y \in Y} f(x,y) \text{ s.t. } g(x,y) \leq 0$$

Inner Problem Bounding Scheme (cont.)

Consider no branching yet

- The auxiliary relaxed inner problem is:

$$\underline{f} = \min_{x \in X, y \in Y} \breve{f}_{x,y}(x,y) \text{ s.t. } \breve{g}_{x,y}(x,y) \leq 0$$

  - Relaxation using convex underestimators $\breve{f}_{x,y}(x,y)$ and $\breve{g}_{x,y}(x,y)$ (e.g. Floudas, 2000, Tawarmalani and Sahinidis, 2002)

- The auxiliary restricted inner problem is:

$$\begin{aligned} \bar{f} = \max_{x_0, x, y, \mu, \lambda, \nu} \quad & x_0, \\ \text{s.t.} \quad & x_0 - f(x,y) \leq 0, \\ & g(x,y) \leq 0, \\ & (x,y) \in X \times Y, \\ & (x,y,\mu,\lambda,\nu) \in \Omega_{\text{KKT}}. \end{aligned}$$

  - Relaxation using the KKT-approach (Still, 2004, Stein and Still, 2002)

Outer Problem Bounding Scheme

Consider no branching yet

- The proposed lower bounding problem is:

$$\begin{aligned} \underline{F} = \min_{x,y,\mu,\lambda,\nu} \quad & F(x,y), \\ \text{s.t.} \quad & G(x,y) \leq 0, \\ & g(x,y) \leq 0, \\ & f(x,y) \leq \bar{f}, \\ & (x,y) \in X \times Y, \\ & (x,y,\mu,\lambda,\nu) \in \Omega_{\mathrm{KKT}}, \end{aligned}$$

  - any feasible solution in the BPP is feasible in the proposed relaxation
  - need to solve to global optimality
- For $x = \bar{x}$, the upper bounding problem is (Mitsos et al., 2008):

$$\bar{F} = \min_{y \in Y} F(\bar{x}, y) \text{ s.t. } G(\bar{x}, y) \leq 0, \; g(\bar{x}, y) \leq 0, \; f(\bar{x}, y) \leq \underline{w}(\bar{x}) + \varepsilon_f$$

- In this work,

$$\underline{w}(\bar{x}) = \min_{y \in Y} \breve{f}_y(\bar{x}, y) \text{ s.t. } \breve{g}_y(\bar{x}, y) \leq 0$$

- Any feasible solution $\bar{y}$ in the restricted problem is feasible in the BPP:

$$f(\bar{x}, \bar{y}) - \varepsilon_f \leq \underline{w}(\bar{x}) \leq w(\bar{x}) \leq f(\bar{x}, \bar{y}) + \varepsilon_f$$

Outer Upper Bounding Problem

Requires partitioning of the inner space Y

- Convexifying the inner problem for fixed x requires some form of refinement of Y
  - in order to compute tighter and tighter approximations of the inner problem over refined subregions of Y
- Subdivision of Y is usually applied to semi-infinite programs, but no branching with respect to y
  - the whole Y is always considered in subproblems
  - e.g. Bhattacharjee et al. (2005a;b), Floudas and Stein (2007), Mitsos et al. (2008a)
- We use partitioning of Y
  - no distinction between the inner and outer decision spaces during branching
  - possible to consider only some subregions of Y and eliminate others via fathoming
  - all Y subregions where global optima may lie are considered

After Fathoming : a useful preliminary theoretical result

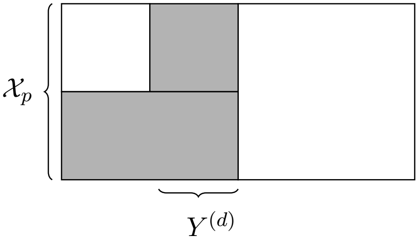
- Every independent list $\mathcal{L}_{\mathcal{X}_p}, p \in \{1, \ldots, n_\mathrm{p}\}$, still contains all promising subregions of Y where global optimal solutions may lie for any $x \in \mathcal{X}_p$
- Define the set of fathomed Y domains for $\mathcal{X}_p$ as follows:

$$\mathcal{F}_{\mathcal{X}_p} := \{\bigcup_d Y^{(d)} \mid Y^{(d)} \subset Y \text{ deleted for all } x \in \mathcal{X}_p\}.$$

- Then, we prove by contradiction that

$$\boxed{\mathcal{Y}(x) \cap \mathcal{F}_{\mathcal{X}_p} = \emptyset \ \forall x \in \mathcal{X}_p}$$

- The sets $\mathcal{F}_{\mathcal{X}_p}, p = 1, \ldots, n_\mathrm{p}$, are *infeasible* in the BPP



The Branch-and-Sandwich algorithm is ε-convergent

At termination, an ε-optimal solution of the bilevel problem is computed

Convergent best inner upper bound
$$\lim_{q\to\infty} f^{\mathrm{UB}}_{\mathcal{X}_{p_q}} = \min_{j\in\mathcal{S}_{\bar{x}}}\{\bar{f}^{(j)}\}$$
$$= \min_{j\in\mathcal{S}_{\bar{x}}} \min_{y\in Y^{(j)}} \{f(\bar{x},y) \text{ s.t. } g(\bar{x},y)\le 0\}$$
$$= w(\bar{x})$$

ε_f-finite Convergent inner B-&-B scheme

Certain-in-the-limit fathoming by outer infeasibility rule
$$\left.\begin{array}{c} f(\bar{x},\bar{y}) \le f^{\mathrm{UB}}_{\mathcal{X}_{p_q}} \\ |f^{\mathrm{UB}}_{\mathcal{X}_{p_q}} - w(\bar{x})| \le \varepsilon_f \; \forall\, q \ge q' \end{array}\right\} \Rightarrow \quad f(\bar{x},\bar{y}) \le w(\bar{x}) + \varepsilon_f$$
$$\bar{y} \notin \mathcal{Y}_{\varepsilon_f}(\bar{x}) \quad \Rightarrow \quad f(\bar{x},\bar{y}) > w(\bar{x}) + \varepsilon_f$$

Consistent bounding scheme: $\underline{F}^{(k_q)}$ and $\bar{F}^{(k_q)}$ are identical in the limit

Bound-improving selection operation

The Branch-and-Sandwich algorithm is ε-convergent by Horst and Tuy (1996)

PRELIMINARY NUMERICAL RESULTS WITH $\varepsilon_f = 10^{-5}$ AND $\varepsilon_F = 10^{-3}$

FOR ALL PROBLEMS, EXCEPT NO. 20 WHERE $\varepsilon_F = 10^{-1}$

No.	NC Inner	n	m	r	F^{UB}	Nodes
1	Yes	0	1	0	-1	1
2	No	0	1	0	1	1
3	No	0	1	0	∞	1
4	Yes	0	1	1	-1	3
5	Yes	0	1	0	1	1
6	Yes	0	1	0	0.5	11
7	Yes	0	1	0	-1	3
8	No	0	1	0	∞	1
9	No	1	1	0	0	1
10	Yes	1	1	0	-1	3
11	Yes	1	1	0	0.5	11
12	Yes	1	1	0	-0.8	1
13	Yes	1	1	0	0	11
14	Yes	1	1	0	-1	27
15	Yes	1	1	0	-1	23
16	Yes	1	1	0	0.25	15
17	Yes	1	1	0	0	13

No.	NC Inner	n	m	r	F^{UB}	Nodes
18	Yes	1	1	0	-2	19
19	Yes	1	1	0	0.1875	47
20	Yes	1	1	0	-0.25	49
21	Yes	1	1	0	-0.258	27
22	Yes	1	1	0	0.3125	39
23	Yes	1	1	0	0.2095	31
24	Yes	1	1	1	0.2095	31
25	Yes	1	1	0	-1.755	11
26	Yes	1	1	0	0	1
27	No	1	1	3	17	1
28	No	1	1	3	22.5	1
29	Yes	1	2	2	0.193616	3
30	No	2	2	3	1.75	1
31	No	2	3	3	29.2	1
32	Yes	2	3	0	-2.35	1
33	Yes	5	5	1	-10	3

BRANCH-AND-BOUND TREE FOR PROBLEM NO. 11