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

1. INTRODUCTION

2. PROPOSED METHOD

3. BOUNDING PROBLEMS: INITIAL NODE

4. BRANCHING & BOUNDING ON SUBDOMAINS

5. NUMERICAL RESULTS

6. CONCLUSIONS
OUTLINE

1 INTRODUCTION

2 PROPOSED METHOD

3 BOUNDING PROBLEMS: INITIAL NODE

4 BRANCHING & BOUNDING ON SUBDOMAINS

5 NUMERICAL RESULTS

6 CONCLUSIONS
A two-person, non-cooperative game in which the play is sequential

The Mixed-Integer Nonlinear Bilevel Problem is

\[
\begin{align*}
\min_{x_i, x_c, y_i, y_c} & \quad F(x_i, x_c, y_i, y_c) \\
s.t. & \quad G(x_i, x_c, y_i, y_c) \leq 0 \\
& \quad (y_i, y_c) \in \arg\min_{y_i \in Y_I, y_c \in Y_C} \{f(x_i, x_c, y_i, y_c) \text{ s.t. } g(x_i, x_c, y_i, y_c) \leq 0\} \\
& \quad x_i \in X_I \subset \mathbb{Z}^n, x_c \in X_C \subset \mathbb{R}^{n-n_1} \\
& \quad y_i \in Y_I \subset \mathbb{Z}^m, y_c \in Y_C \subset \mathbb{R}^{m-m_1}
\end{align*}
\]

Subscripts \(i\) and \(c\) stand for \textit{integer} and \textit{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
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):

\[
\begin{align*}
\min_{x_i, x_c, y_i, y_c} & \quad F(x_i, x_c, y_i, y_c) \\
\text{s.t.} & \quad G(x_i, x_c, y_i, y_c) \leq 0 \\
& \quad g(x_i, x_c, y_i, y_c) \leq 0 \\
& \quad f(x_i, x_c, y_i, y_c) \leq w(x_i, x_c) \\
& \quad x_i \in X_I, x_c \in X_C \\
& \quad y_i \in Y_I, y_c \in Y_C 
\end{align*}
\]

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):

\[
\begin{align*}
\min_{x_i, x_c, y_i, y_c} & \quad F(x_i, x_c, y_i, y_c) \\
\text{s.t.} & \quad G(x_i, x_c, y_i, y_c) \leq 0 \\
& \quad g(x_i, x_c, y_i, y_c) \leq 0 \\
& \quad f(x_i, x_c, y_i, y_c) \leq w(x_i, x_c) \\
& \quad x_i \in X_I, x_c \in X_C \\
& \quad y_i \in Y_I, y_c \in Y_C
\end{align*}
\]

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.**

- **Advances on more general classes**
CHALLENGES: ILLUSTRATIVE EXAMPLE (MOORE AND BARD, 1990)

\[
\begin{align*}
\text{min} & \quad -x - 10y \\
\text{s.t.} & \quad \begin{array}{c}
x \in [0,8] \\
x \in \mathbb{Z} \\
y \in \arg \min \{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\} \quad y \in [0,4]
\end{array}
\end{align*}
\]
**Challenges: Illustrative Example (Moore and Bard, 1990)**

\[
\begin{align*}
\text{min} & \quad -x - 10y \\
\text{s.t.} & \quad x \in \mathbb{Z} \\
& \quad y \in \arg \min \{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z} \} \\
& \quad y \in [0, 4]
\end{align*}
\]
### Challenge 1: Constructing a Convergent Lower Bounding Problem

\[
\begin{align*}
\text{min} & \quad -x - 10y \\
\text{s.t.} & \quad y \in Y(x) \\
& \quad x \in [0, 8], y \in [0, 4] \\
& \quad x, y \in \mathbb{Z}
\end{align*}
\]

Y(x) : Inner Feasible Region
**Challenge 2: Partitioning the Inner Region**

\[
\begin{align*}
\text{min} & \quad -x - 10y \\
\text{s.t.} & \quad x \in \mathbb{Z} \\
& \quad y \in \arg \min \{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\} \\
& \quad y \in [0,1]
\end{align*}
\]

\[
\begin{align*}
\text{min} & \quad -x - 10y \\
\text{s.t.} & \quad x \in \mathbb{Z} \\
& \quad y \in \arg \min \{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\} \\
& \quad y \in [2,4]
\end{align*}
\]
OUTLINE

1 INTRODUCTION

2 PROPOSED METHOD

3 BOUNDING PROBLEMS: INITIAL NODE

4 BRANCHING & BOUNDING ON SUBDOMAINS

5 NUMERICAL RESULTS

6 CONCLUSIONS
Branch-and-Sandwich Algorithm\textsuperscript{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 $\varepsilon$-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


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

1. INTRODUCTION

2. PROPOSED METHOD

3. BOUNDING PROBLEMS: INITIAL NODE

4. BRANCHING & BOUNDING ON SUBDOMAINS

5. NUMERICAL RESULTS

6. CONCLUSIONS
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

$$f^L = \min_{x_i, x_c, y_i, y_c} f(x_i, x_c, y_i, y_c)$$

subject to

$$g(x_i, x_c, y_i, y_c) \leq 0$$

$$x_i \in X_I, x_c \in X_C, y_i \in Y_I, y_c \in Y_C$$

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)

$$\bar{f} = \min_{x_i, x_c, y_i, y_c} \bar{f}(x_i, x_c, y_i, y_c)$$

subject to

$$\bar{g}(x_i, x_c, y_i, y_c) \leq 0$$

$$x_i \in X_I, x_c \in X_C, y_i \in Y_I, y_c \in Y_C$$
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:

\[
\begin{align*}
    f^U &= \min_{y_i, y_c, t} \quad t \\
    \text{s.t.} & \quad f(x_i, x_c, y_i, y_c) \leq t \quad \forall (x_i, x_c) \in X_I \times X_C \\
    & \quad g(x_i, x_c, y_i, y_c) \leq 0 \quad \forall (x_i, x_c) \in X_I \times X_C \\
    & \quad y_i \in Y_I, \ y_c \in Y_C
\end{align*}
\]

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):

\[
\begin{align*}
    \bar{f} &= \min_{y_i, y_c, t} \quad t \\
    \text{s.t.} & \quad \bar{f}(X_I, X_C, y_i, y_c) \leq t \\
    & \quad \bar{g}(X_I, X_C, y_i, y_c) \leq 0 \\
    & \quad y_i \in Y_I, \ y_c \in Y_C
\end{align*}
\]
**INNER PROBLEM BOUNDING SCHEME (CONT.)**

**Motivation:** We must have valid inner bounds over the current subdomain.

---

**Figure:** Inner lower bound

![Plot of $f(x,y) = x^2 y + \sin(y)$](image)

- $(x^*, y^*) = (0, 4.7124)$
- $(x, y) = (1, 1)$

**Figure:** Inner upper bound

![Plot of $f(x,y) = \frac{y^2}{x}$](image)

- $x = 1, F^* = 3$
- $x = 2, F^* = 6$
- $x = 4, F^* = 18$
The proposed lower bounding problem is

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

▶ 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)

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

▶ any feasible solution in the upper bounding problem is feasible in the (MINBP)
▶ set \( F^{UB} := \min\{\bar{F}, \infty\} \) to express the incumbent
**Outer Problem Bounding Scheme**

**Consider no branching yet**

- The proposed lower bounding problem is

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

- 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)

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

- Any feasible solution in the upper bounding problem is feasible in the (MINBP).

- Set $F^{UB} := \min\{\bar{F}, \infty\}$ to express the incumbent.
OUTLINE

1 INTRODUCTION

2 PROPOSED METHOD

3 BOUNDING PROBLEMS: INITIAL NODE

4 BRANCHING & BOUNDING ON SUBDOMAINS

5 NUMERICAL RESULTS

6 CONCLUSIONS
**List Management**

**Auxiliary lists**

- $\mathcal{L}$: is the classical list of ‘open’ nodes, corresponding to the outer problem
- $\mathcal{L}^1$: is the list of exclusively inner ‘open’ nodes
- $\mathcal{L}_{x_p}$: is the list of (outer & inner) nodes that cover the whole $Y$ for a subdomain $x_p$ of $X$, $1 \leq p \leq n_p$

- Number $n_p$ equals the number of partition sets $x_p \subseteq X$ s.t.:
  - for each subdomain $x_p$, the “whole” $Y$ is maintained
- Two lists $\mathcal{L}_{x_{p1}}, \mathcal{L}_{x_{p2}}$ are called independent if $\mathcal{L}_{x_{p1}} \cap \mathcal{L}_{x_{p2}} = \emptyset$
  - two lists with common nodes, e.g. $S_{x_p}^1$ and $S_{x_p}^2$, are sublists of $\mathcal{L}_{x_p}$

- For each $p$, best inner upper bound lowest over $Y$, but largest over $x_p$:

\[
    f_{x_p}^{UB} = \max \left\{ \min_{j \in S_{x_p}^1} \tilde{f}(j), \ldots, \min_{j \in S_{x_p}^s} \tilde{f}(j) \right\}.
\]
**MODIFIED BOUNDING PROBLEMS**

**ALLOW BRANCHING & CONSIDER A NODE** \( k \in \mathcal{L}\mathcal{X}_p \)

### Outer Upper Bound

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

\[
\bar{F}(k') = \min_{y_i, y_c} F(\bar{x}_i, \bar{x}_c, y_i, y_c)
\]

\[
\text{s.t.} \quad G(\bar{x}_i, \bar{x}_c, y_i, y_c) \leq 0
\]

\[
g(\bar{x}_i, \bar{x}_c, y_i, y_c) \leq 0
\]

\[
f(\bar{x}_i, \bar{x}_c, y_i, y_c) \leq w^{(k')}(\bar{x}_i, \bar{x}_c) + \varepsilon f
\]

\[
y_i \in Y_I^{(k')}, y_c \in Y_C^{(k')}
\]

### Outer Lower Bound

\[
F^{(k)} = \min_{x_i, x_c, y_i, y_c} F(x_i, x_c, y_i, y_c)
\]

\[
\text{s.t.} \quad G(x_i, x_c, y_i, y_c) \leq 0
\]

\[
g(x_i, x_c, y_i, y_c) \leq 0
\]

\[
f(x_i, x_c, y_i, y_c) \leq f^{(k)}(x_i, x_c, y_i, y_c)
\]

\[
y_i \in Y_I^{(k)}, y_c \in Y_C^{(k)}
\]

### Inner Lower Bound

\[
f^{(k)} = \min_{x_i, x_c, y_i, y_c} \bar{f}(x_i, x_c, y_i, y_c)
\]

\[
\bar{f}(x_i, x_c, y_i, y_c) = \min_{y_i, y_c, t} \bar{f}(x_i, x_c, y_i, y_c)
\]

\[
\text{s.t.} \quad \bar{f}(x_i, x_c, y_i, y_c) \leq 0
\]

\[
x_i \in X_I^{(k)}, x_c \in X_C^{(k)}
\]

\[
y_i \in Y_I^{(k)}, y_c \in Y_C^{(k)}
\]

### Inner Upper Bound

\[
\bar{f}(x_i, x_c, y_i, y_c) = \min_{x_i, x_c, y_i, y_c} \bar{f}(x_i, x_c, y_i, y_c)
\]

\[
\text{s.t.} \quad \bar{f}(x_i, x_c, y_i, y_c) \leq t
\]

\[
\bar{g}(x_i, x_c, y_i, y_c) \leq 0
\]

\[
x_i \in X_I^{(k)}, x_c \in X_C^{(k)}
\]

\[
y_i \in Y_I^{(k)}, y_c \in Y_C^{(k)}
\]
NODE FATHOMING RULES

Node $k \in \mathcal{L} \cap \mathcal{L}_{X_{p}}$

**INNER FATHOMING RULES**

1. If $f^{(k)} = \infty$ or
2. $f^{(k)} > f^{\text{UB}}_{X_{p}}$

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

**OUTER FATHOMING RULES**

1. If $F^{(k)} = \infty$ or
2. $F^{(k)} \geq F^{\text{UB}} - \varepsilon_{F}$

then **OUTER FATHOM**, i.e. move from $\mathcal{L}$ to $\mathcal{L}^{I}$. Hence, $k \in \mathcal{L}^{I} \cap \mathcal{L}_{X_{p}}$. 
**Node Fathoming Rules**

Node \( k \in \mathcal{L} \cap \mathcal{L}_{X_p} \)

**Inner Fathoming Rules**

**If**

1. \( f^{(k)} = \infty \) or
2. \( f^{(k)} > f^\text{UB} \)

**Then** fathom, i.e. delete from \( \mathcal{L} \) (or \( \mathcal{L}^I \)) and \( \mathcal{L}_{X_p} \).

**Outer Fathoming Rules**

**If**

1. \( F^{(k)} = \infty \) or
2. \( F^{(k)} \geq F^\text{UB} - \varepsilon_F \)

**Then** Outer Fathom, i.e. move from \( \mathcal{L} \) to \( \mathcal{L}^I \). Hence, \( k \in \mathcal{L}^I \cap \mathcal{L}_{X_p} \).
OUTLINE

1 INTRODUCTION

2 PROPOSED METHOD

3 BOUNDING PROBLEMS: INITIAL NODE

4 BRANCHING & BOUNDING ON SUBDOMAINS

5 NUMERICAL RESULTS

6 CONCLUSIONS
ILLUSTRATIVE EXAMPLE REVISITED: EXTENDED-TREE VERSION

\[
\begin{align*}
\min_{x \in [0,8]} \{ -x - 10y & \text{ s.t. } x \in \mathbb{Z}, \ y \in \arg\min \{ y \text{ s.t. } y \in Y(x), y \in \mathbb{Z} \} \}\quad \text{with } F^* = -22 \text{ at } (x^*, y^*) = (2, 2).
\end{align*}
\]

\[
\begin{align*}
& f^{(1)}(1) = 1, \bar{f}^{(1)}(1) = 4, \\
& \bar{F}^{(1)}(1) = -42, \bar{x}^{(1)} = 2, \\
& w(\bar{x}) = 2, \bar{F}^{(1)} = -22.
\end{align*}
\]
**ILLUSTRATIVE EXAMPLE REVISITED : EXTENDED-TREE VERSION**

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

\[
\begin{align*}
\text{min} & \quad \{ -x - 10y \text{ s.t. } x \in \mathbb{Z}, \ y \in \arg \min \{ y \text{ s.t. } y \in Y(x), y \in \mathbb{Z} \} \} \text{ with } F^* = -22 \text{ at } (x^*, y^*) = (2, 2) \\
& \quad \text{min} \quad \{ -x - 10y \text{ s.t. } x \in \mathbb{Z}, \ y \in \arg \min \{ y \text{ s.t. } y \in Y(x), y \in \mathbb{Z} \} \} \text{ with } F^* = -22 \text{ at } (x^*, y^*) = (2, 2)
\end{align*}
\]
**ILLUSTRATIVE EXAMPLE REVISITED: EXTENDED-TREE VERSION**

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

\[
\begin{align*}
\text{with } F^* &= -22 \\
(2, 2) &\text{ (lowest)}
\end{align*}
\]

\[
\begin{align*}
f^{(1)} &= 1, \quad F^*_{UB} = 4, \quad F^{(1)} = -42, \quad F^*_{UB} = -22 \\
f^{(2)} &= 1, \quad F^*_{UB} = 4, \quad F^{(2)} = -18
\end{align*}
\]

\[
\begin{align*}
f^{(3)} &= 2, \quad F^{(3)} = -42, \quad F^*_{UB} = -22 \\
\bar{x}^{(3)} &= 2, \quad w^{(3)}(\bar{x}) = 2 (\text{lowest}), \quad F^{(3)} = -22
\end{align*}
\]
Illustrative Example Revisited: Extended-Tree Version

$$\min_{x \in [0,8]} \{ -x - 10y \text{ s.t. } x \in \mathbb{Z}, \ y \in \arg \min \{ y \text{ s.t. } y \in Y(x), y \in \mathbb{Z} \} \} \text{ with } F^* = -22 \text{ at } (x^*, y^*) = (2, 2)$$
ILLUSTRATIVE EXAMPLE REVISITED: EXTENDED-TREE VERSION

\[
\min_{x \in [0,8]} \{-x - 10y\} \text{ s.t. } x \in \mathbb{Z}, \ y \in \arg \min \{y \text{ s.t. } y \in Y(x), y \in \mathbb{Z}\} \text{ with } F^* = -22 \text{ at } (x^*, y^*) = (2, 2)
\]
ILLUSTRATIVE EXAMPLE REVISITED : EXTENDED-TREE VERSION

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

\[
\text{with } \bar{f}(10) = \infty, \bar{f}(11) = 4, \bar{f}^\text{UB}(11) = 4
\]

\[
f^\text{UB}(8) = 4, f^\text{UB}(9) = 2
\]

\[
F^\text{UB}(1) = -22
\]

\[
F^\text{UB}(2) = -42
\]

\[
F^\text{UB}(3) = -22
\]

\[
F^\text{UB}(4) = -42
\]

\[
F^\text{UB}(5) = -42
\]

\[
F^\text{UB}(6) = -22
\]

\[
F^\text{UB}(7) = 1
\]

\[
F^\text{UB}(8) = 4
\]

\[
F^\text{UB}(9) = 2
\]
ILLUSTRATIVE EXAMPLE REVISITED: EXTENDED-TREE VERSION

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

\[ x \leq 3 \quad x \geq 4 \]
\[ f^{(6)} = 1 \]
\[ f^{(7)} = 1, \bar{f}^{\text{UB}}_{x_2} = 1 \]
\[ f^{(10)} = \infty \quad f^{(11)} = 1, \bar{f}^{(11)} = 4 \]

\[ y \leq 1 \quad y \geq 2 \]
\[ F^{(2)} = -18 \]
\[ f^{(3)} = 2, \bar{F}^{(3)} = -42 \]
\[ f^{(4)} = 2, \bar{F}^{(4)} = -42 \]
\[ f^{(5)} = 2 > \bar{f}^{\text{UB}}_{x_2} \]

\[ \bar{f}^{(9)} = 2, \bar{f}^{\text{UB}}_{x_2} = 2, \bar{F}^{(9)} = -22 \]

\[ f^{(8)} = 2, \bar{F}^{(8)} = -21 \]

\[ f^{(1)} = 1, \bar{f}^{\text{UB}}_x = 4, \bar{F}^{(1)} = -42, \bar{F}^{\text{UB}} = -22 \]
ILLUSTRATIVE EXAMPLE REVISITED: EXTENDED-TREE VERSION

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

![Diagram of the extended tree version with bounding problems and branching on subdomains](image)

**NUMERICAL RESULTS**

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

- \( f^{(3)} = 2, F^{(3)} = -42 \)

- \( f^{(5)} = 2 > f^\text{UB}_X \)

- \( f^{(10)} = \infty \)
- \( f^{(11)} = 1, f^{(11)} = 4 \)

- \( f^{(8)} = 2, F^{(8)} = -21 \)
- \( f^{(9)} = 2, f^\text{UB}_X = 2, F^{(9)} = -22 \)
## Preliminary numerical results with $\varepsilon_f = 10^{-5}$ and $\varepsilon_F = 10^{-3}$

<table>
<thead>
<tr>
<th>No.</th>
<th>Source</th>
<th>Problem</th>
<th>$x_i$</th>
<th>$x_c$</th>
<th>$y_i$</th>
<th>$y_c$</th>
<th>$F_{UB}$</th>
<th>Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Moore &amp; Bard (1990)</td>
<td>Example 1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>-22</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>Moore &amp; Bard (1990)</td>
<td>Example 2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>Edmunds &amp; Bard (1992)</td>
<td>Equation 3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$\frac{4}{9}$</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Sahin &amp; Ciric (1998)</td>
<td>Example 4</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>-400</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Dempe (2002)</td>
<td>Equation 8.11</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>-10.4</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>Mitsos (2010)</td>
<td>am_1.0_0_1.01</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>Mitsos (2010)</td>
<td>am_1.1_1.0_01</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>Mitsos (2010)</td>
<td>am_1.1_1.1_01</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>Mitsos (2010)</td>
<td>am_1.1_1.1_02</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.209</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Mitsos (2010)</td>
<td>am_3_3_3_3_01</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>-2.5</td>
<td>1</td>
</tr>
</tbody>
</table>
INTRODUCTION

PROPOSED METHOD

BOUNDING PROBLEMS: INITIAL NODE

BRANCHING & BOUNDING ON SUBDOMAINS

NUMERICAL RESULTS

CONCLUSIONS
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 $f_{\mathcal{X}_p}^{UB}$

- 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
Problem Formulation: Continuous Case

- The optimistic bilevel problem is a LEADER-follower game.
- The leader (outer) problem is:
  \[
  \min_{x,y} F(x, y) \quad \text{s.t.} \quad G(x, y) \leq 0, \ (x, y) \in X \times Y, \ y \in \mathcal{Y}(x) \quad \text{(BPP)}
  \]
- \(\mathcal{Y}(x)\) 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

\[
y^L \leq y \leq y^U,
\]

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

\[
\Omega_{\text{KKT}} = \left\{ (x, y, \mu, \lambda, \nu) \middle| \begin{align*}
\nabla_y f(x, y) + \mu \nabla_y g(x, y) - \lambda + \nu &= 0, \\
\mu^T g(x, y) &= 0, \quad \mu \geq 0, \\
\lambda^T (y^L - y) &= 0, \quad \lambda \geq 0, \\
\nu^T (y - y^U) &= 0, \quad \nu \geq 0.
\end{align*} \right\}
\]

- \( \Omega_{\text{KKT}} \) 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)**

\[
\min_{y \in \mathbb{R}} \quad f(y) = 16y^4 + 2y^3 - 8y^2 - 3/2y + 1/2
\]

subject to:

\[
y \in \mathbb{R}
\]

**Graphical Representation:**

- **(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:

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

- The auxiliary restricted inner problem is:

\[ f^U = \max_{x \in X} \min_{y \in Y} f(x, y) \text{ s.t. } g(x, y) \leq 0 \]
INNER PROBLEM BOUNDING SCHEME (CONT.)

CONSIDER NO BRANCHING YET

- The auxiliary relaxed inner problem is:

  \[
  f = \min_{x \in X, y \in Y} f_{x,y}(x, y) \text{ s.t. } g_{x,y}(x, y) \leq 0
  \]

  Relaxation using convex underestimators \( f_{x,y}(x, y) \) and \( g_{x,y}(x, y) \) (e.g. Floudas, 2000, Tawarmalani and Sahinidis, 2002)

- The auxiliary restricted inner problem is:

  \[
  \tilde{f} = \max_{x_0, x, y, \mu, \lambda, \nu} x_0, \text{ s.t. } \]
  \[
  x_0 - f(x, y) \leq 0, \\
  g(x, y) \leq 0, \\
  (x, y) \in X \times Y, \\
  (x, y, \mu, \lambda, \nu) \in \Omega_{\text{KKT}}.
  \]

  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:

\[
F = \min_{x, y, \mu, \lambda, \nu} F(x, y), \\
\text{s.t.} \quad G(x, y) \leq 0, \\
\quad g(x, y) \leq 0, \\
\quad f(x, y) \leq \bar{f}, \\
\quad (x, y) \in X \times Y, \\
\quad (x, y, \mu, \lambda, \nu) \in \Omega_{KKT},
\]

- 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):

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

- In this work,

\[
\bar{w}(\bar{x}) = \min_{y \in Y} \tilde{f}_y(\bar{x}, y) \quad \text{s.t.} \quad \tilde{g}_y(\bar{x}, y) \leq 0
\]

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

\[
f(\bar{x}, \bar{y}) - \varepsilon_f \leq \bar{w}(\bar{x}) \leq w(\bar{x}) \leq f(\bar{x}, \bar{y}) + \varepsilon_f
\]
OUTER UPPER BOUNDED 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 $\mathcal{L}_x$, $p \in \{1, \ldots, n_p\}$, still contains all promising subregions of $Y$ where global optimal solutions may lie for any $x \in x_p$
- Define the set of fathomed $Y$ domains for $x_p$ as follows:

$$\mathcal{F}_x := \bigcup_{d} Y^{(d)} \mid Y^{(d)} \subset Y \text{ deleted for all } x \in x_p.$$ 

- Then, we prove by contradiction that

$$\mathcal{Y}(x) \cap \mathcal{F}_x = \emptyset \ \forall x \in x_p$$

- The sets $\mathcal{F}_x$, $p = 1, \ldots, n_p$, are infeasible in the BPP

\[\text{Diagram of $x_p$ and $Y^{(d)}$} \]
The Branch-and-Sandwich algorithm is \( \varepsilon \)-convergent

At termination, an \( \varepsilon \)-optimal solution of the bilevel problem is computed

\[
\lim_{q \to \infty} f_{\bar{x}_{pq}}^{UB} = \min_{j \in S_{\bar{x}}} \{ \bar{f}(j) \}
= \min_{j \in S_{\bar{x}}} \min_{y \in Y(j)} \{ f(\bar{x}, y) \text{ s.t. } g(\bar{x}, y) \leq 0 \}
= w(\bar{x})
\]

\( \varepsilon \)-finite Convergent inner B-&-B scheme

\[
f(\bar{x}, \bar{y}) \leq f_{\bar{x}_{pq}}^{UB}
\]
\[
|f_{\bar{x}_{pq}}^{UB} - w(\bar{x})| \leq \varepsilon f \forall q \geq q'
\]
\[
\bar{y} \notin Y_{\varepsilon f}(\bar{x})
\]

\[
\Rightarrow f(\bar{x}, \bar{y}) \leq w(\bar{x}) + \varepsilon f
\]

\[
\Rightarrow f(\bar{x}, \bar{y}) > w(\bar{x}) + \varepsilon f
\]

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

The Branch-and-Sandwich algorithm is \( \varepsilon \)-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 $\varepsilon_F = 10^{-1}$

<table>
<thead>
<tr>
<th>No.</th>
<th>NC Inner</th>
<th>n</th>
<th>m</th>
<th>r</th>
<th>$f^{\text{UB}}$</th>
<th>Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>$-1$</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>$1$</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>$\infty$</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>$-1$</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Yes</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>$1$</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Yes</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>$0.5$</td>
<td>11</td>
</tr>
<tr>
<td>7</td>
<td>Yes</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>$-1$</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>No</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>$\infty$</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>No</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$0$</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$-1$</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$0.5$</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$-0.8$</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$0$</td>
<td>11</td>
</tr>
<tr>
<td>14</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$-1$</td>
<td>27</td>
</tr>
<tr>
<td>15</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$-1$</td>
<td>23</td>
</tr>
<tr>
<td>16</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$0.25$</td>
<td>15</td>
</tr>
<tr>
<td>17</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$0$</td>
<td>13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No.</th>
<th>NC Inner</th>
<th>n</th>
<th>m</th>
<th>r</th>
<th>$f^{\text{UB}}$</th>
<th>Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$-2$</td>
<td>19</td>
</tr>
<tr>
<td>19</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$0.1875$</td>
<td>47</td>
</tr>
<tr>
<td>20</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$-0.25$</td>
<td>49</td>
</tr>
<tr>
<td>21</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$-0.258$</td>
<td>27</td>
</tr>
<tr>
<td>22</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$0.3125$</td>
<td>39</td>
</tr>
<tr>
<td>23</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$0.2095$</td>
<td>31</td>
</tr>
<tr>
<td>24</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$0.2095$</td>
<td>31</td>
</tr>
<tr>
<td>25</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$-1.755$</td>
<td>11</td>
</tr>
<tr>
<td>26</td>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$0$</td>
<td>1</td>
</tr>
<tr>
<td>27</td>
<td>No</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>$17$</td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>No</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>$22.5$</td>
<td>1</td>
</tr>
<tr>
<td>29</td>
<td>Yes</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>$0.193616$</td>
<td>3</td>
</tr>
<tr>
<td>30</td>
<td>No</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>$1.75$</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>No</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>$29.2$</td>
<td>1</td>
</tr>
<tr>
<td>32</td>
<td>Yes</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>$-2.35$</td>
<td>1</td>
</tr>
<tr>
<td>33</td>
<td>Yes</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>$-10$</td>
<td>3</td>
</tr>
</tbody>
</table>
Branch-and-Bound Tree for Problem No. 11

1. \( f^{(1)} = -52.50, \bar{f}_{[0.1,1]} = 0.57, F^{(1)} = -0.50, \bar{F}^{(1)} = \infty \)

   - \( y \in [-1,0] \)
   - \( x \in [0.1,0.55] \)
   - \( f^{(2)} = -24, \bar{f}^{(2)} = 0.57, F^{(2)} = \infty \)

   - 2
     - \( x \in [0.55,1] \)
     - \( y \in [-1,-0.5] \)
     - \( f^{(6)} = -14.50, \bar{f}^{(6)} = 0.31 \)
     - \( f^{(7)} = -16, \bar{f}^{(7)} = 0.57 \)

   - 7
     - \( y \in [-0.5,0] \)
     - \( f^{(10)} = -4.60, \bar{f}^{(10)} = 0 \)
     - \( f^{(11)} = -1.70, \bar{f}^{(11)} = 0.57 \)

2. \( x \in [0.1,0.55] \)

   - 3
     - \( x \in [0.55,1] \)
     - \( y \in [0,0.5] \)
     - \( f^{(3)} = -22.60, \bar{f}_{[0.1,1]} = -0.10, F^{(3)} = 0.50, \bar{F}^{(3)} = \infty \)

   - 4
     - \( y \in [0.5,1] \)
     - \( f^{(4)} = -14.20, \bar{f}^{(4)} = -0.10, F^{(4)} = 0.50, \bar{F}^{(4)} = \) \( y \in [0,0.5] \)

3. \( f^{(5)} = -15.40, \bar{f}_{[0.55,1]} = -0.55, F^{(5)} = 0.50, \bar{F}^{(5)} = \infty \)

   - 5
     - \( y \in [0.5,1] \)
     - \( f^{(8)} = -2, \bar{f}^{(8)} = -0.55, F^{(8)} = 0.50, F^{UB} = 0.50 \)

   - 8
     - \( f^{(9)} = -4.70, \bar{f}^{(9)} = -0.55, F^{(9)} = 0.50, \bar{F}^{(9)} = 0.50 \)

   - 11
     - \( f^{(11)} = -1.70, \bar{f}^{(11)} = 0.57 \)

10

- \( f^{(6)} = -14.50, \bar{f}^{(6)} = 0.31 \)
- \( f^{(7)} = -16, \bar{f}^{(7)} = 0.57 \)