Single and Multi-Objective Bilevel Optimization - An Illustrated Travel Guide

Aalto University School of Science, 21st June 2023

CARLOS HENGGELER ANTUNES
(joint work with MARIA JOÃO ALVES)
University of Coimbra, Portugal

AGENDA

- I – Single-objective bilevel optimization - main concepts
- II – Multi-objective bilevel optimization
- III – Bilevel optimization - an application in the energy sector

Part I

Single-Objective Bilevel Optimization

BILEVEL OPTIMIZATION - STACKELBERG GAMES

- **Hierarchical structure**: a second optimization problem is embedded in the constraints of the bilevel programming problem
- Roots from game theory: **Stackelberg game**
  - Stackelberg (1934, engl. transl 1952) used by the first time a hierarchical model to describe real market situations.
  - Two players: the **leader** and the **follower**
  - **Sequential** decision process: the leader moves first and the follower moves next: The leader makes a decision and commits a strategy before the other player (follower)
  - **Full information** is available and the game is **non-cooperative**

BILEVEL PROGRAMMING PROBLEM

$$
\begin{aligned}
\max_{x} \quad & F(x,y) \\
\text{s.t.} \quad & G(x,y) \leq 0 \\
& \max_{y} \quad f(x,y) \\
& \text{s.t.} \quad g(x,y) \leq 0
\end{aligned}
$$

(BLP)

x - variables controlled by the **leader** (upper level)

y - variables controlled by the **follower** (lower level)

- A **bilevel** program is a "*mathematical program that contains an optimization problem in the constraints*" (Bracken, 1973)
- Two decision makers: the **leader** and the **follower,** who pursue different objectives in a non-cooperative manner.
- The **leader makes his decision first.** The follower **reacts** by choosing his optimal candidate on the feasible choices restricted by the leader.

EXAMPLE OF APPLICATION

- For example, in a **toll-setting problem**:
  - the owners of a highway system (*leader*) have to set tolls and they want to **maximize revenues**
  - but users of highways (follower) want to **minimize their travel costs**

The leader's decision must take into account that few users will use the highways if tolls are set too high, which may result in small revenues for the leader.

DEFINITIONS

BLP:

$$\max_{x} \quad F(x,y)$$
$$\text{s.t.} \quad G(x,y) \leq 0$$
$$\quad \max_{y} \quad f(x,y)$$
$$\quad \text{s.t.} \quad g(x,y) \leq 0$$

where $x \in R^{n1}$ and $y \in R^{n2}$

- **Constraint Region** of the BLP

$$S = \{(x,y): G(x,y) \leq 0\,,\ g(x,y) \leq 0\,\}$$

(It is assumed that S is non-empty and compact)

- Follower's **rational reaction set** to a given x

$$\Psi(x) = \arg\ \max_{y}\{f(x,y)\ : g(x,y) \leq 0\}$$

It is assumed that $\Psi(x) \neq \emptyset$ for all x taken by the leader

DEFINITIONS

BLP:

$$\max_{x} \quad F(x,y)$$
$$\text{s.t.} \quad G(x,y) \leq 0$$
$$\max_{y} \quad f(x,y)$$
$$\text{s.t.} \quad g(x,y) \leq 0$$

$\Big\}\ (x,y) \in IR$

where $x \in R^{n1}$ and $y \in R^{n2}$

- **Inducible region**

$$IR = \{(x,y): (x,y) \in S,\ y \in \Psi(x)\}$$

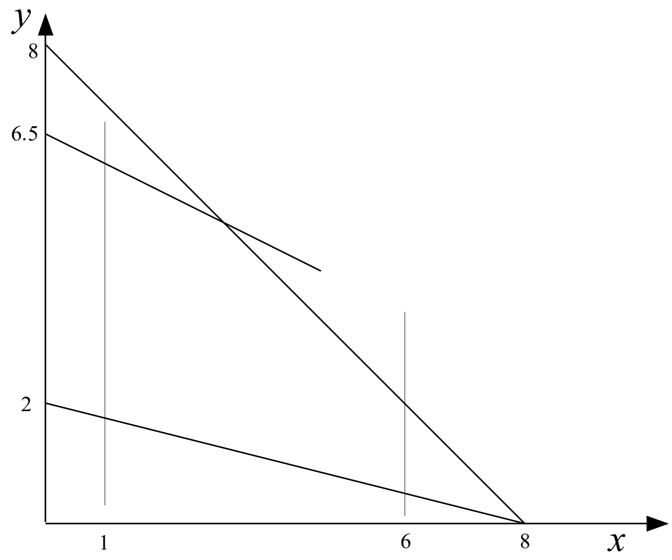
The BLP can be written as:

$$\max_{x} \quad F(x,y)$$
$$\text{subject to} \quad (x,y) \in IR$$

EXAMPLE 1

$$\max_{x} \quad F(x,y) = -x - 3y$$
$$\text{s.t.} \quad 1 \leq x \leq 6$$
$$\max_{y} \quad f(y) = y$$
$$\text{s.t.} \quad x + y \leq 8$$
$$x + 4y \geq 8$$
$$x + 2y \leq 13$$

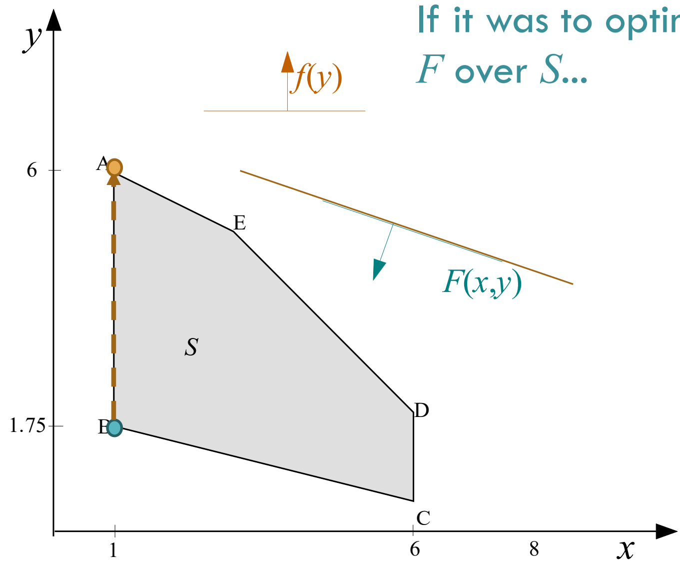
(In: Dempe, 2002)



EXAMPLE 1

$$\max_x \quad F(x,y) = -x - 3y$$
$$\text{s.t.} \quad 1 \leq x \leq 6$$
$$\max_y \quad f(y) = y$$
$$\text{s.t.} \quad x + y \leq 8$$
$$x + 4y \geq 8$$
$$x + 2y \leq 13$$

(In: Dempe, 2002)



If it was to optimize *F* over *S*...

...point B would be the optimal solution

But, the follower optimizes *f*(y) for each *x* chosen by the leader

EXAMPLE 1

$$\max_{x} \quad F(x,y) = -x - 3y$$

$$\text{s.t.} \quad 1 \le x \le 6$$

$$\max_{y} \quad f(y) = y$$

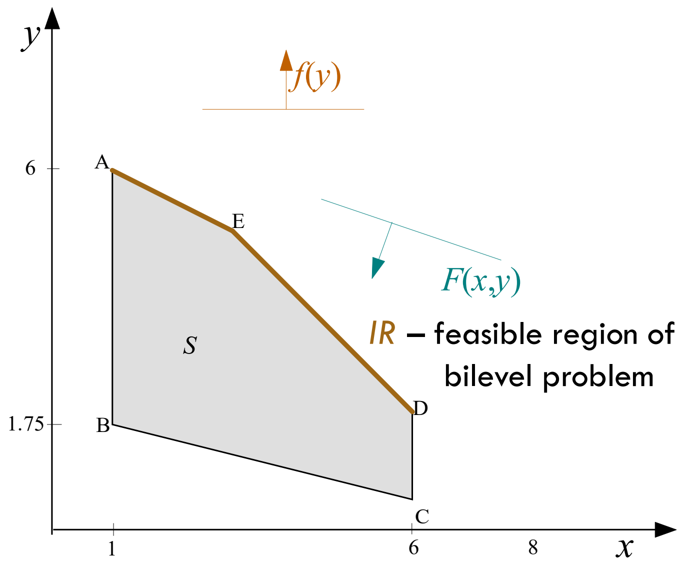
$$\text{s.t.} \quad x + y \le 8$$

$$x + 4y \ge 8$$

$$x + 2y \le 13$$

IR

(In: Dempe, 2002)



IR – feasible region of the bilevel problem

EXAMPLE 1

$$\max_{x} \quad F(x,y) = -x - 3y$$

$$\text{s.t.} \quad 1 \le x \le 6$$

$$\max_{y} \quad f(y) = y$$

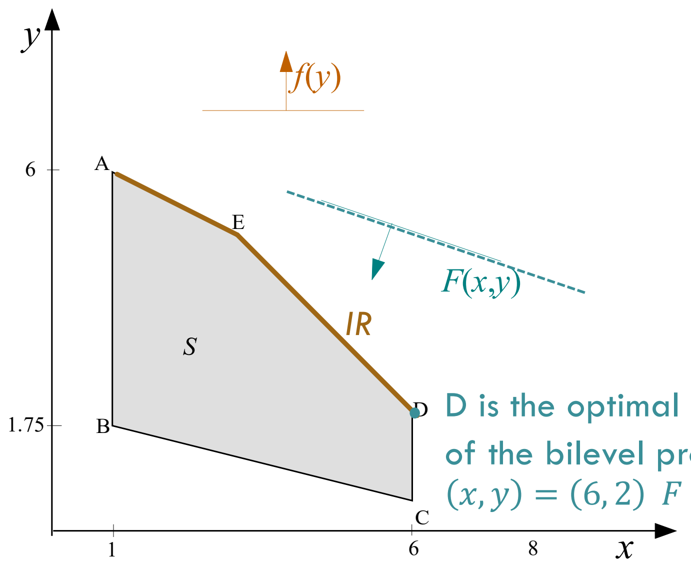
$$\text{s.t.} \quad x + y \le 8$$

$$x + 4y \ge 8$$

$$x + 2y \le 13$$

IR

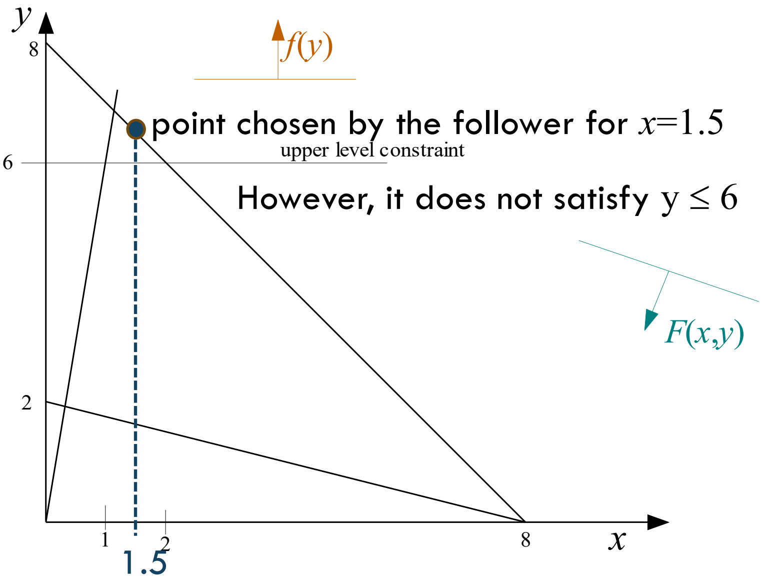
(In: Dempe, 2002)



D is the optimal solution of the bilevel problem $(x,y) = (6,2)$ $F = -12$

EXAMPLE 2

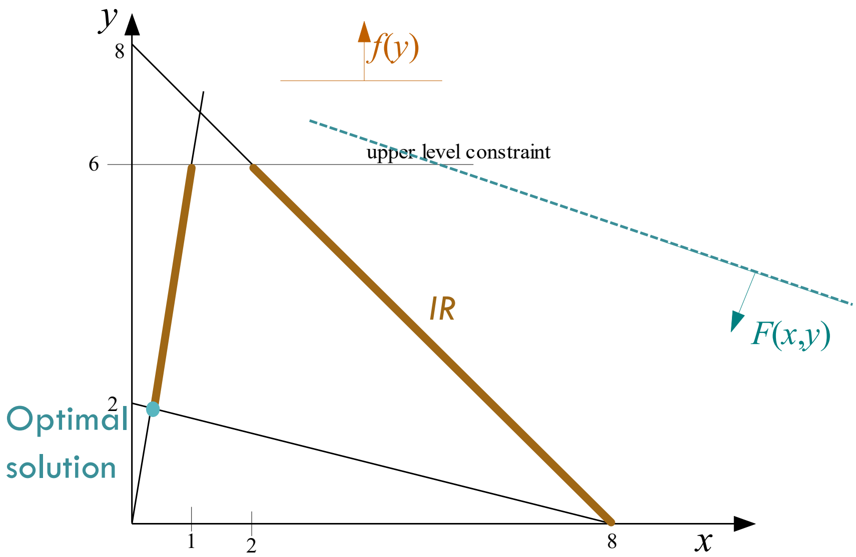
$$\max_{x} \quad F(x,y) = -x - 3y$$
$$\text{s.t.} \quad 0 \le x \le 8$$
$$y \le 6$$
$$\max_{y} \quad f(y) = y$$
$$\text{s.t.} \quad x + y \le 8$$
$$x + 4y \ge 8$$
$$-6x + y \le 0$$



(Adap. from Dempe, 2002)

EXAMPLE 2

$$\max_x \quad F(x,y) = -x - 3y$$
$$\text{s.t.} \quad 0 \le x \le 8$$
$$y \le 6$$
$$\max_y \quad f(y) = y$$
$$\text{s.t.} \quad x + y \le 8$$
$$x + 4y \ge 8$$
$$-6x + y \le 0$$



(Adap. from Dempe, 2002)

Problem with disconnected feasible set

PROPERTIES OF THE LINEAR BLP

- The linear BLP is NP-hard
- The **inducible region** IR is composed by the union of faces of S
  - **connected** faces if upper-level constraints do not depend on the lower-level optimal solutions

$$\begin{aligned}
\max_{x} \quad & F(x,y) = c_1 x + d_1 y \\
\text{s.t.} \quad & A_1 x + B_1 y \le b_1 \\
& x \ge 0 \\
& \max_{y} \quad f(x,y) = c_2 x + d_2 y \\
& \text{s.t.} \quad A_2 x + B_2 y \le b_2 \\
& \qquad y \ge 0
\end{aligned}$$

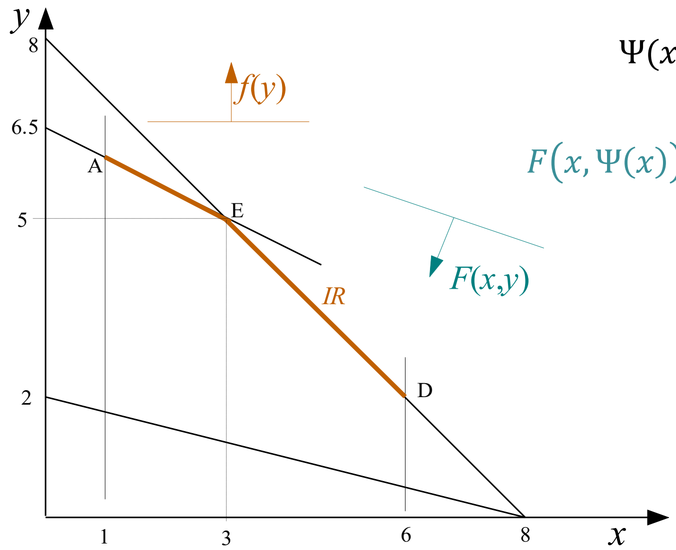
- The solution occurs at a **vertex** of IR

PROPERTIES OF THE (LINEAR) BLP

Bilevel problems are **nonconvex** and **nondifferentiable** optimization problems

$$\max_x \quad F(x,y) = -x - 3y$$
$$\text{s.t.} \quad 1 \le x \le 6$$
$$\max_y \quad f(y) = y$$
$$\text{s.t.} \quad x + y \le 8, \quad x + 4y \ge 8, \quad x + 2y \le 13$$

IR



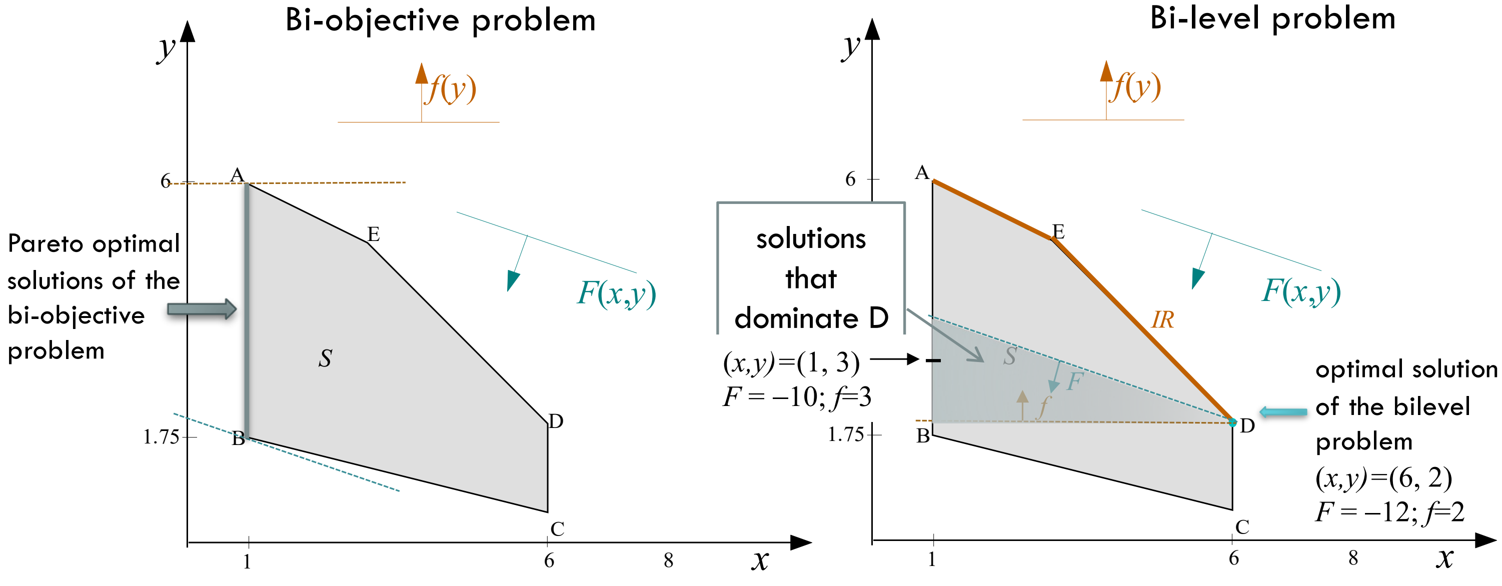
Rational reaction set to each x

$$\Psi(x) = \begin{cases} 6.5 - 0.5x & 1 \le x \le 3 \\ 8 - x & 3 \le x \le 6 \end{cases}$$

$$F\big(x, \Psi(x)\big) = \begin{cases} -19.5 + 0.5x & 1 \le x \le 3 \\ -24 + 2x & 3 \le x \le 6 \end{cases}$$

BILEVEL VS. BI-OBJECTIVE

- In general, the optimal solution of the bilevel problem is **not a Pareto optimal solution** of the bi-objective problem.



METHODS FOR THE LINEAR BLP

Main algorithmic approaches:

1. **Vertex** enumeration
2. Reformulate the L-BLP as a single-level problem using the **Karush-Kuhn-Tucker (KKT) conditions**
3. Descent algorithms and penalty approaches

2. KKT-BASED APPROACHES

L-BLP:

$$\max_{x} \quad F(x,y) = c_1 x + d_1 y$$
$$\text{s.t.} \quad A_1 x + B_1 y \leq b_1$$
$$x \geq 0$$
$$\max_{y} \quad f(y) = d_2 y$$
$$\text{s.t.} \quad A_2 x + B_2 y \leq b_2$$
$$y \geq 0$$

Mathematical program with complementarity constraints:

$$\max_{x,y} \quad F(x,y) = c_1 x + d_1 y$$
$$\text{s.t.} \quad A_1 x + B_1 y \leq b_1$$
$$A_2 x + B_2 y \leq b_2$$
$$\lambda B_2 \geq d_2$$
$$\lambda \, (b_2 - A_2 x - B_2 y) = 0$$
$$y \, (\lambda B_2 - d_2) = 0$$
$$x \geq 0, \; y \geq 0, \; \lambda \geq 0$$

Where $\lambda \in \Re^{m2}$ are the dual variables of the lower level problem

2. KKT-BASED APPROACHES

Mathematical program with complementarity constraints:

$$\max_{x,y} \quad F(x,y) = c_1 x + d_1 y$$

$$\text{s.t.} \quad A_1 x + B_1 y \leq b_1$$

$$A_2 x + B_2 y \leq b_2$$

$$\lambda B_2 \geq d_2$$

$$\lambda \, (b_2 - A_2 x - B_2 y) = 0$$

$$y \, (\lambda B_2 - d_2) = 0$$

$$x \geq 0, \; y \geq 0, \; \lambda \geq 0$$

- Use a branch and bound strategy to deal with the complementarity constraints
- Transform the program into a mixed-integer linear programming (MILP) problem and solve it using a general MILP solver (e.g., *cplex*)

OPTIMISTIC VS. PESSIMISTIC SOLUTIONS

If it cannot be ensured the uniqueness of optimal solutions for the lower-level problem, the formulation of the BLP is **ambiguous:**

$$\begin{aligned} \text{"}\max_x\text{"} \quad & F(x,y) \\ \text{s.t.} \quad & G(x,y) \leq 0 \\ & \max_y \quad f(x,y) \\ & \text{s.t.} \quad g(x,y) \leq 0 \end{aligned}$$

That is: the follower's **rational reaction set**

$$\Psi(x) = \arg\max_y \{ f(y) : g(x,y) \leq 0 \}$$

is not single valued for every x

OPTIMISTIC VS. PESSIMISTIC SOLUTIONS

If it cannot be ensured the uniqueness of optimal solutions for the lower-level problem, the formulation of the BLP is **ambiguous:**

$$
\begin{aligned}
&\underset{x}{\text{"max"}} \quad F(x,y) \\
&\text{s.t.} \quad G(x,y) \leq 0 \\
&\qquad \max_{y} \quad f(x,y) \\
&\qquad \text{s.t.} \quad g(x,y) \leq 0
\end{aligned}
$$

Two main approaches have been suggested:

- **Optimistic** approach
- **Pessimistic** approach

which consider the optimistic and the pessimistic reformulations, respectively

OPTIMISTIC VS. PESSIMISTIC SOLUTIONS

- **Optimistic** formulation: the follower chooses, among his alternative optima, the solution that is the **best** for the leader:

BLP$_o$: $$\max_{x, G(x,y)\leq 0} \varphi_o(x) := \max_y \{F(x,y) : y \in \Psi(x)\}$$

- **Pessimistic** formulation: the follower chooses, among his alternative optima, the solution that is the **worst** for the leader:

BLP$_p$: $$\max_{x, G(x,y)\leq 0} \varphi_p(x) := \min_y \{F(x,y) : y \in \Psi(x)\}$$

OPTIMISTIC VS. PESSIMISTIC SOLUTIONS

- The **optimistic** formulation is much simpler to handle and has been **mostly investigated,** mainly for its simplified version:

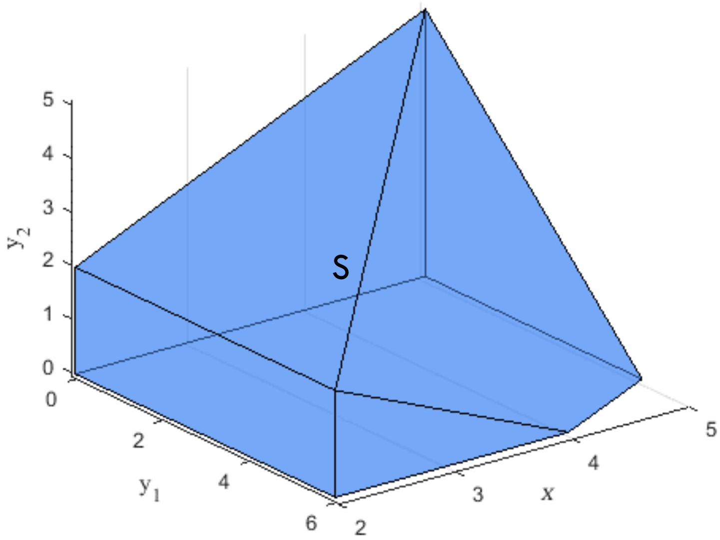
$$\begin{aligned} \max_{x,y} \quad & F(x,y) \\ \text{s.t.} \quad & G(x,y) \leq 0 \\ & \max_{y} \quad f(x,y) \\ & \text{s.t.} \quad g(x,y) \leq 0 \end{aligned}$$

upper-level optimization is taken with respect to x and y

KKT conditions → only valid in the optimistic case

OPTIMISTIC VS. PESSIMISTIC SOLUTIONS - EXAMPLE 1

$$
\begin{aligned}
&\max \quad F = x + 2y_1 - y_2 \\
&s.t. \\
&\quad 2 \leq x \leq 5 \\
&\quad \left.\begin{aligned}
\max \quad f &= y_1 + y_2 \\
s.t. \quad y_1 &\leq 6 \\
y_1 + y_2 &\leq 10 - x \\
y_2 &\leq x \\
y_1, y_2 &\geq 0
\end{aligned}\right\} IR
\end{aligned}
$$

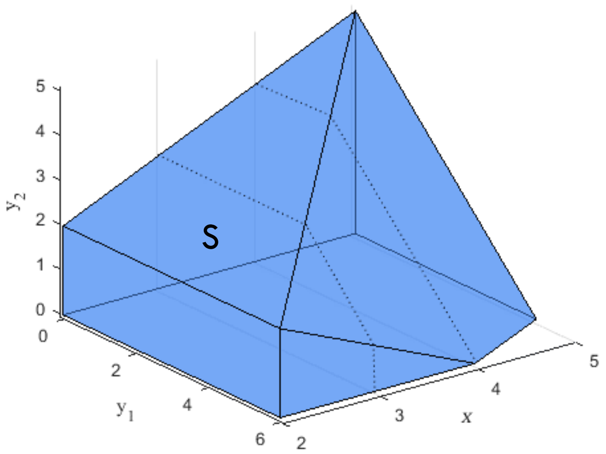


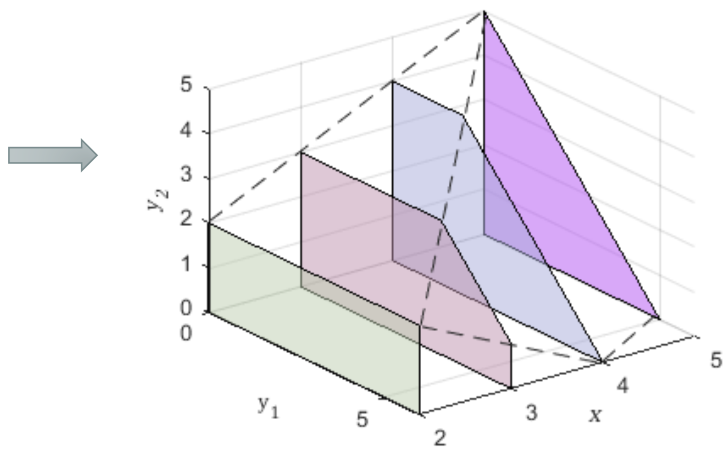
S≡ constraint region

x is the upper level variable and y_1, y_2 are the lower level variables

OPTIMISTIC VS. PESSIMISTIC SOLUTIONS - EXAMPLE 1

Let us analyze the behavior of F and f for some discrete values of x:

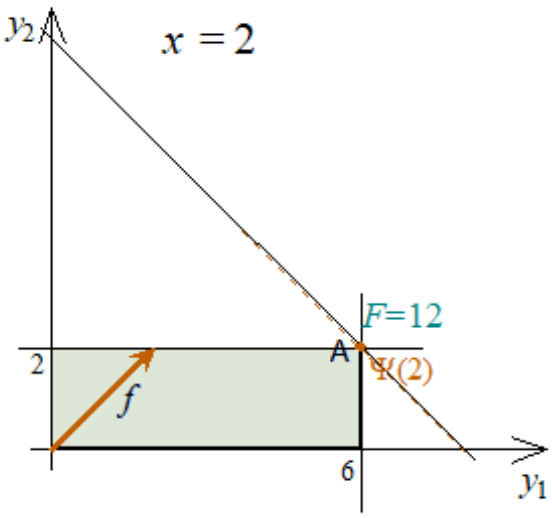


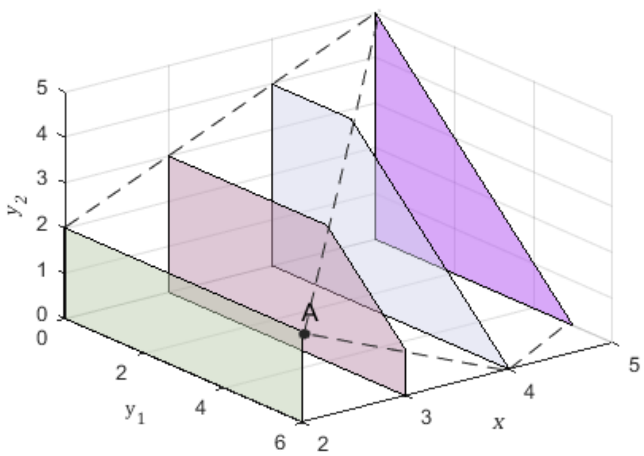


OPTIMISTIC VS. PESSIMISTIC SOLUTIONS - EXAMPLE 1

$\max \; F = x + 2y_1 - y_2$

$\max \; f = y_1 + y_2$

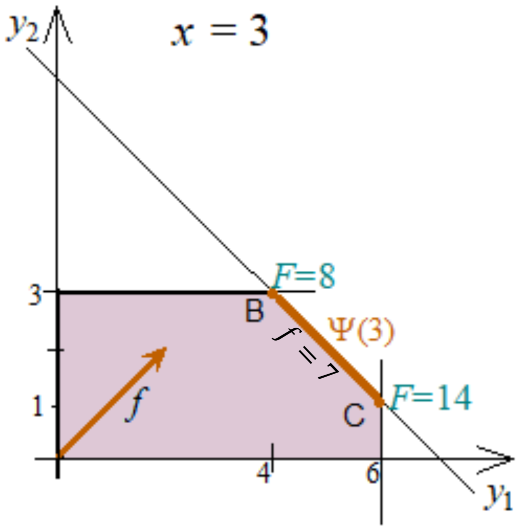


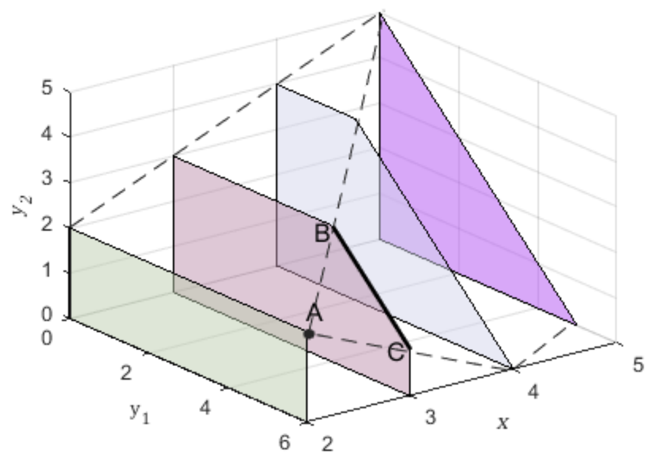


OPTIMISTIC VS. PESSIMISTIC SOLUTIONS
- EXAMPLE 1

$$\max \quad F = x + 2y_1 - y_2$$

$$\max \quad f = y_1 + y_2$$

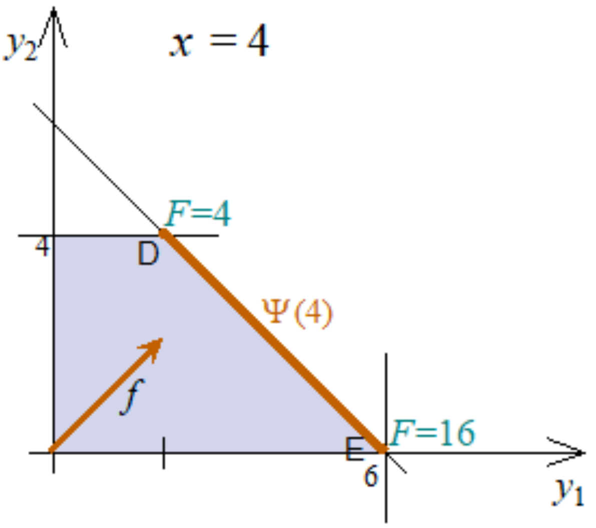


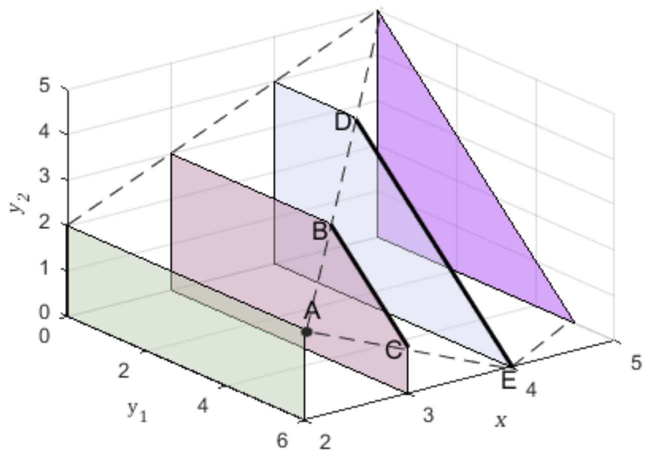


OPTIMISTIC VS. PESSIMISTIC SOLUTIONS
- EXAMPLE 1

$$\max \quad F = x + 2y_1 - y_2$$

$$\max \quad f = y_1 + y_2$$

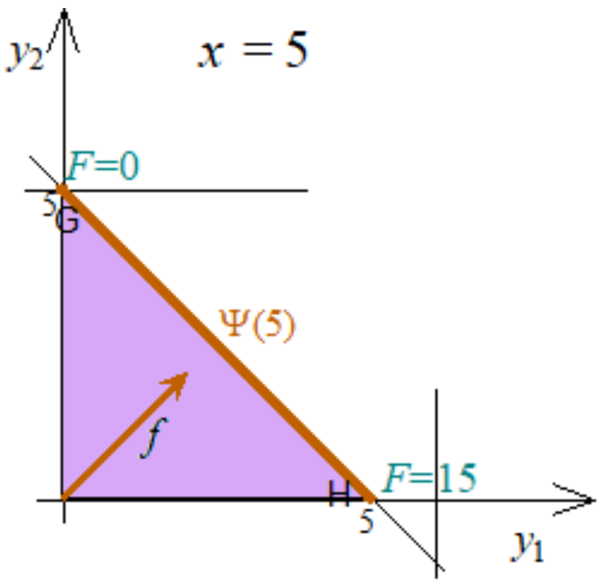


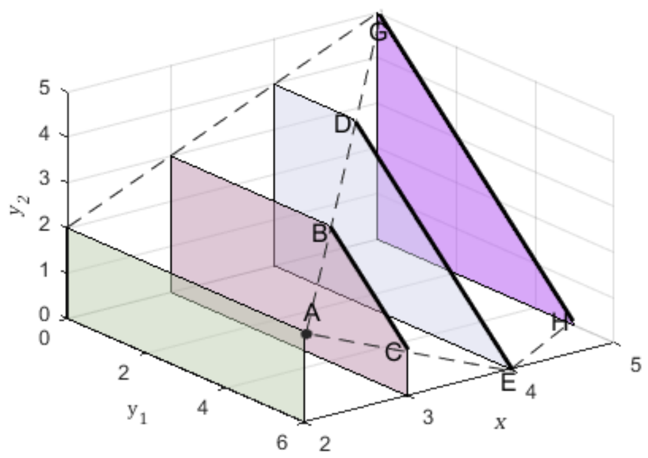


OPTIMISTIC VS. PESSIMISTIC SOLUTIONS - EXAMPLE 1

$$\max \quad F = x + 2y_1 - y_2$$

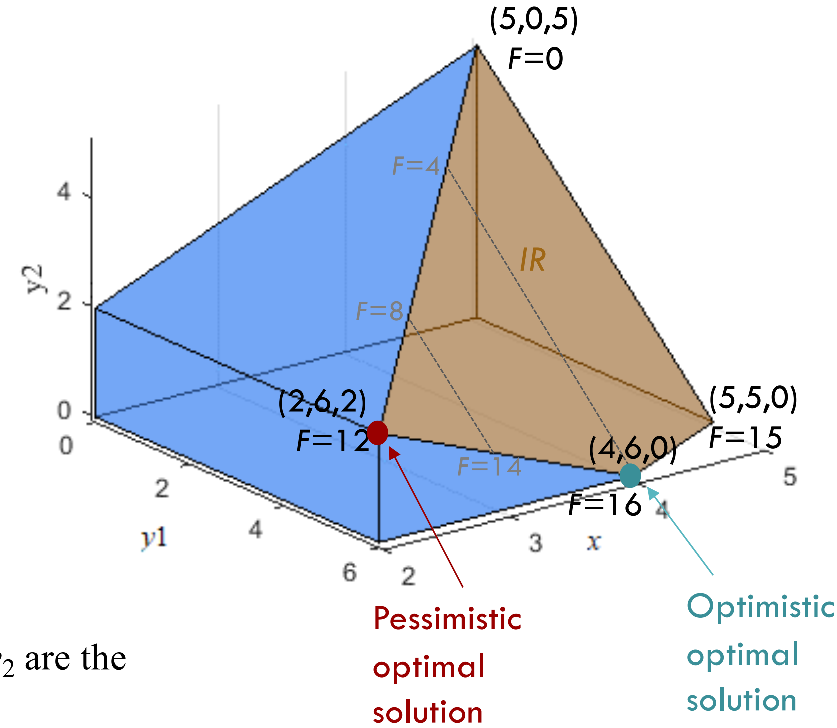
$$\max \quad f = y_1 + y_2$$





OPTIMISTIC VS. PESSIMISTIC SOLUTIONS - EXAMPLE 1

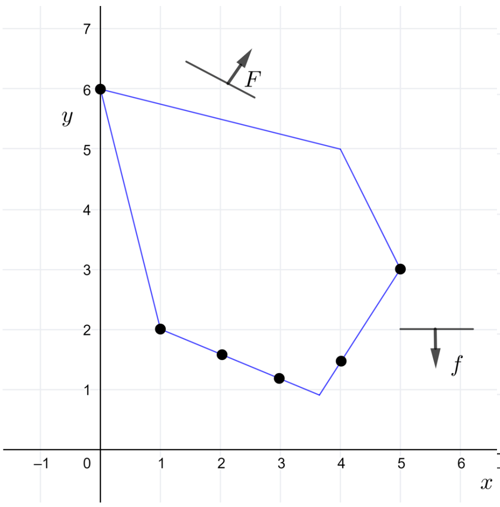
$$
\begin{aligned}
&\max \quad F = x + 2y_1 - y_2 \\
&s.t. \\
&\quad \left.\begin{aligned}
&2 \le x \le 5 \\
&\max \quad f = y_1 + y_2 \\
&s.t. \; y_1 \le 6 \\
&\quad y_1 + y_2 \le 10 - x \\
&\quad y_2 \le x \\
&\quad y_1, y_2 \ge 0
\end{aligned}\right\} IR
\end{aligned}
$$



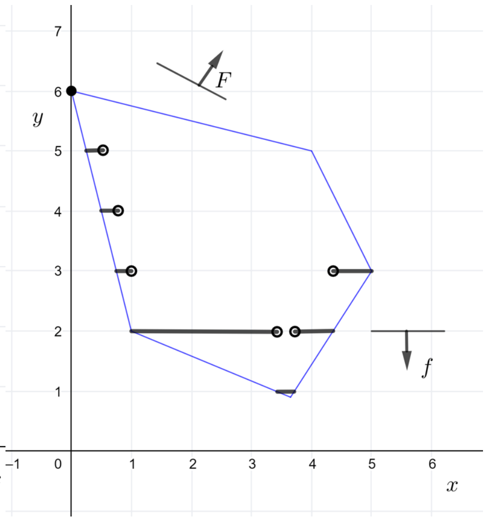
x is the upper-level variable and y_1, y_2 are the lower-level variables

BLP WITH INTEGER VARIABLES

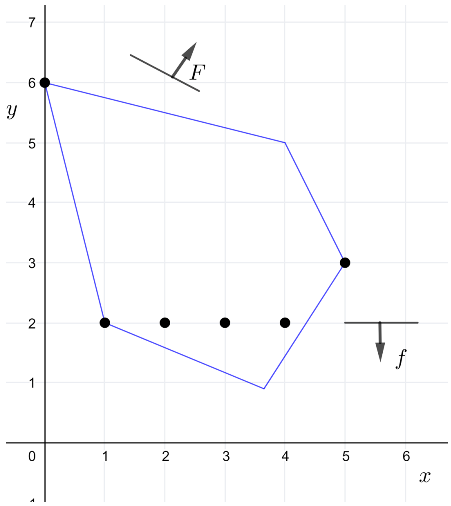
Figs. adapaed from: Saharidis, Conejo, Kozanidis (2013)



Discrete-Continuous
(DC-BLP)



Continuous-Discrete
(CD-BLP)



Discrete-Discrete
(DD-BLP)

The *IR* of **DC-BLP** is included in the *IR* of the (continuous) **BLP**

The *IR* of **DD-BLP** is included in the *IR* of **CD-BLP**

INTEGER AND MIXED-INTEGER BLP:
OBSTACLES IN ALGORITHMIC DEVELOPMENT

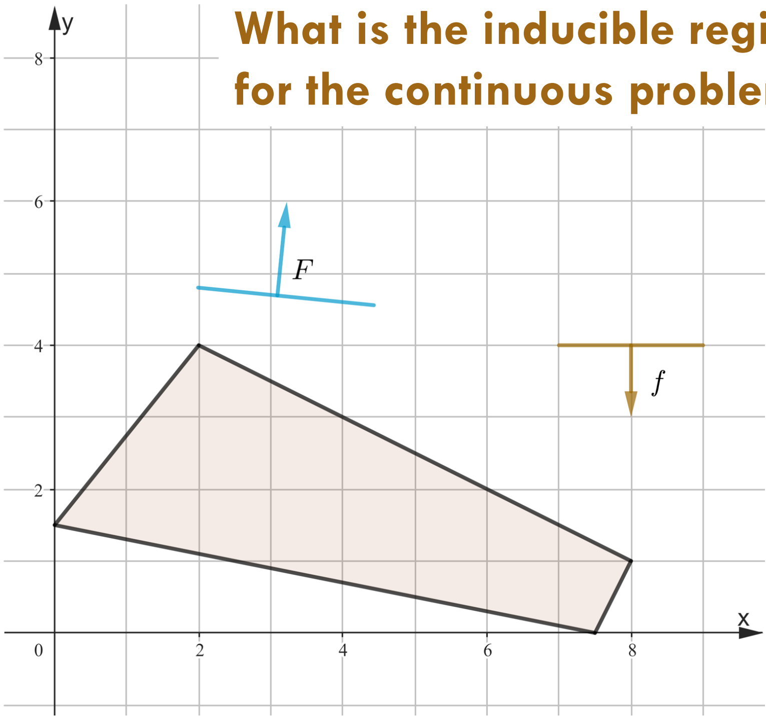
$$\underset{x}{\text{Maximize}}\ F(x, y) = x + 10y$$

where y solves

$$\max_{y} f(x, y) = -y$$

subject to
$$-25x + 20y \leq 30$$
$$x + 2y \leq 10$$
$$2x - y \leq 15$$
$$2x + 10y \geq 15$$
$$x, y \geq 0$$
~~x, y integer.~~

In: Bard and Moore (1990)



What is the inducible region for the continuous problem?

INTEGER AND MIXED-INTEGER BLP: OBSTACLES IN ALGORITHMIC DEVELOPMENT

$$\underset{x}{\text{Maximize}}\ F(x, y) = x + 10y$$

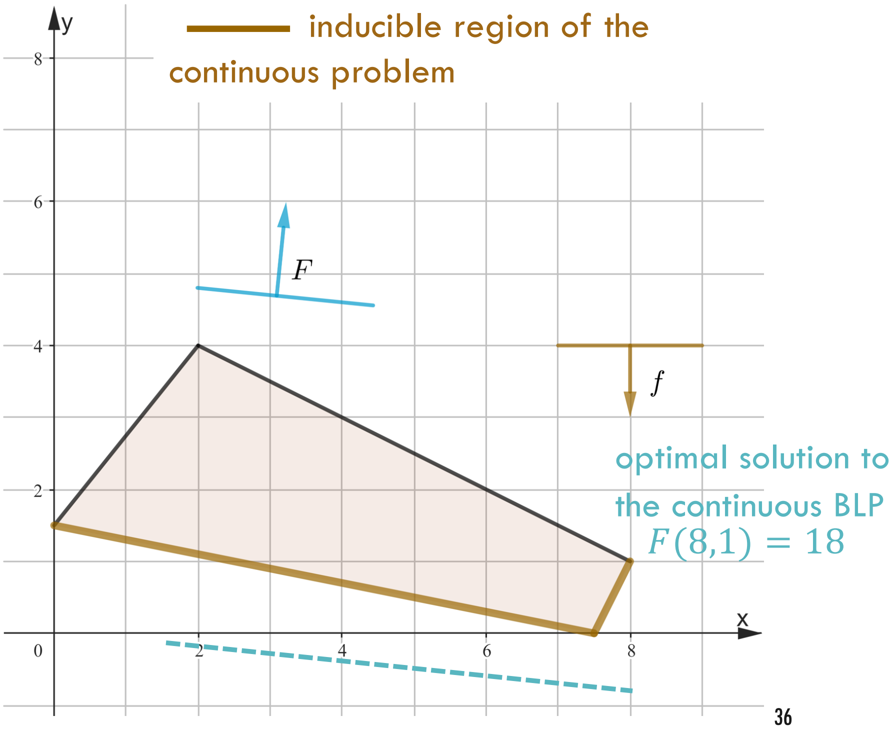
where y solves

$$\max_{y} f(x, y) = -y$$

subject to
$$-25x + 20y \leq 30$$
$$x + 2y \leq 10$$
$$2x - y \leq 15$$
$$2x + 10y \geq 15$$
$$x, y \geq 0$$

~~x, y integer.~~

In: Bard and Moore (1990)



INTEGER AND MIXED-INTEGER BLP: OBSTACLES IN ALGORITHMIC DEVELOPMENT

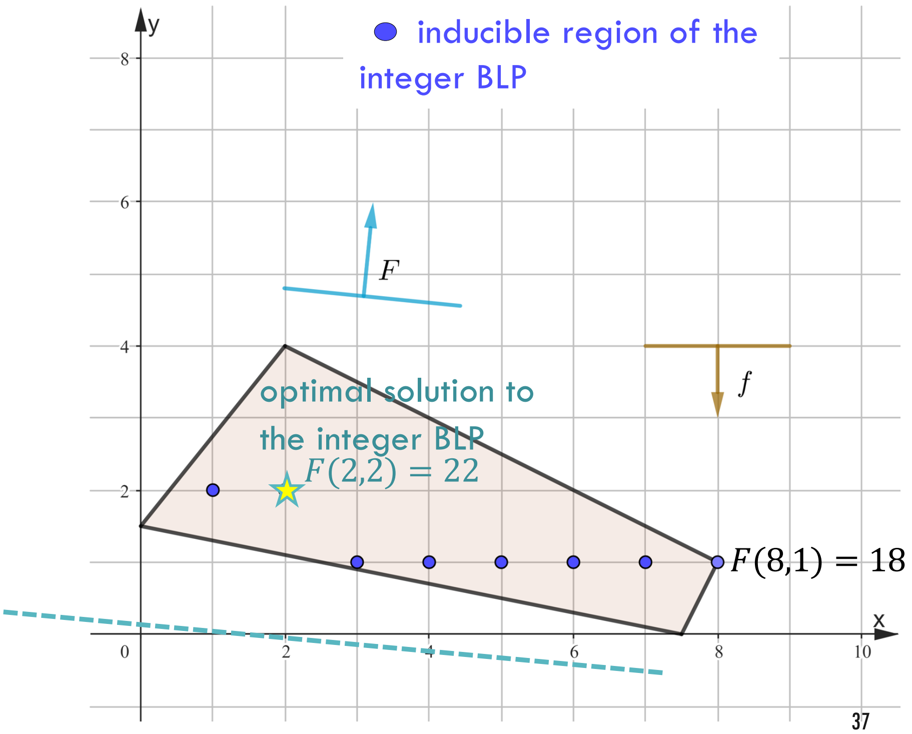
$$\underset{x}{\text{Maximize}}\ F(x, y) = x + 10y$$

where y solves

$$\max_{y} f(x, y) = -y$$

$$\begin{aligned}\text{subject to}\quad & -25x + 20y \leq 30\\ & x + 2y \leq 10\\ & 2x - y \leq 15\\ & 2x + 10y \geq 15\\ & x, y \geq 0\\ & x, y \text{ integer.}\end{aligned}$$

In: Bard and Moore (1990)

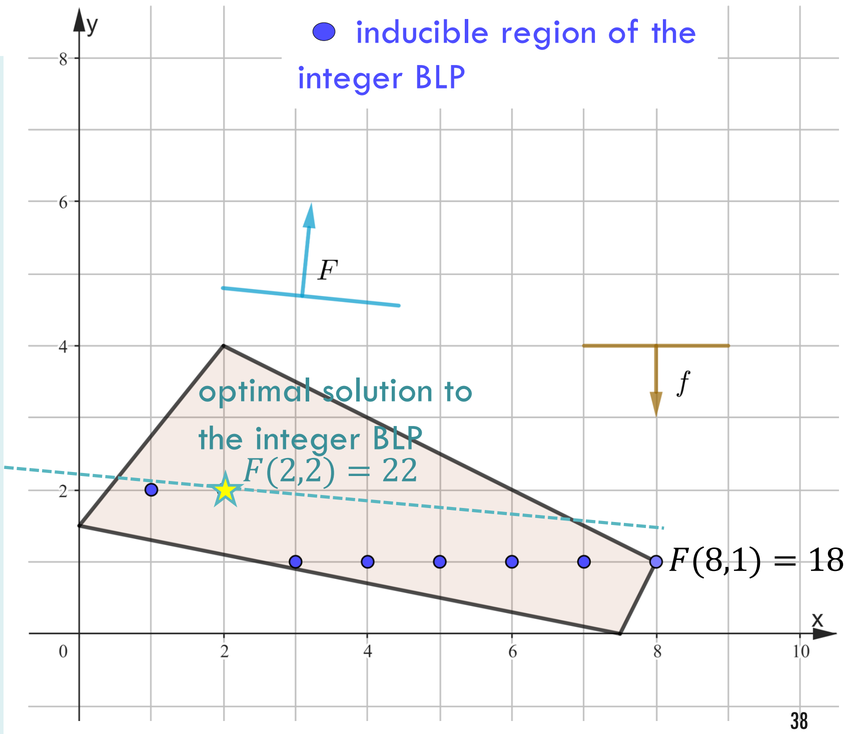


INTEGER AND MIXED-INTEGER BLP: OBSTACLES IN ALGORITHMIC DEVELOPMENT

It may happen:

- the optimal solution to the relaxed BLP (continuous BLP) is integer and it is not the optimal solution to the integer BLP

- the optimal F to the relaxed BLP (continuous BLP) does not provide a valid bound to the optimal F for the integer BLP.



Part II
Multi-objective Bilevel Optimization

MULTI-OBJECTIVE BILEVEL PROGRAMMING

$$\max_{x} \quad (F_1(x,y), \ldots, F_k(x,y))$$
$$s.t. \quad G(x,y) \leq 0$$
$$y \in \arg\max_{y} \{(f_1(x,y), \ldots, f_m(x,y)) : g(x,y) \leq 0\}$$

$x \in \mathbb{R}^{n_1}$ – variables controlled by the leader
$y \in \mathbb{R}^{n_2}$ – variables controlled by the follower

MULTI-OBJECTIVE BILEVEL PROBLEMS: CASES

- **Multiple objective functions at the upper-level**, single-objective function at the lower-level

(SVBLP)

- Single-objective function at the upper-level**, multiple objective functions at the lower-level** (also known as semi-vectorial BL problem):

$$\max_{x} \quad F(x,y)$$
$$s.t. \quad G(x,y) \leq 0$$
$$y \in \arg\max_{y} \{(f_1(x,y),\ldots,f_m(x,y)) : g(x,y) \leq 0\}$$

- **Multiple objective functions at both levels**

SEMI-VECTORIAL BILEVEL PROBLEM (SVBLP)

$$\begin{aligned}
&\max_{x} \quad F(x,y) \\
&\text{s.t.} \quad G(x) \leq 0 \\
&\qquad \max_{y} \quad (f_1(x,y), f_2(x,y), \ldots, f_m(x,y)) \\
&\qquad \text{s.t.} \quad g(x,y) \leq 0
\end{aligned}$$

(SVBLP)

- Only **efficient** solutions to the lower-level problem for each x -vector are feasible to the SVBLP.
- For each x-vector, there is no best f-value to the follower, but rather a set of lower-level efficient solutions

(For simplicity reasons, we consider that upper-level constraints $G(x) \leq 0$ do not include lower-level variables)

SEMI-VECTORIAL BILEVEL PROBLEM (SVBLP)

Let $Y(x') = \{y: g(x', y) \leq 0\}$

The set of **efficient** solutions to the lower-level problem of the SVBLP for a given x' is:

$$\Psi_{\text{Ef}}(x') = \{y' \in Y(x'): \text{ there is no } y \in Y(x') \text{ such that } f(x', y) \succ f(x', y')\}$$

(where $\succ$ denotes the dominance relation)

The **inducible region** (feasible region) of the SVBLP is:

$$IR = \left\{(x, y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}: G(x) \leq 0,\ y \in \Psi_{\text{Ef}}(x)\right\}$$

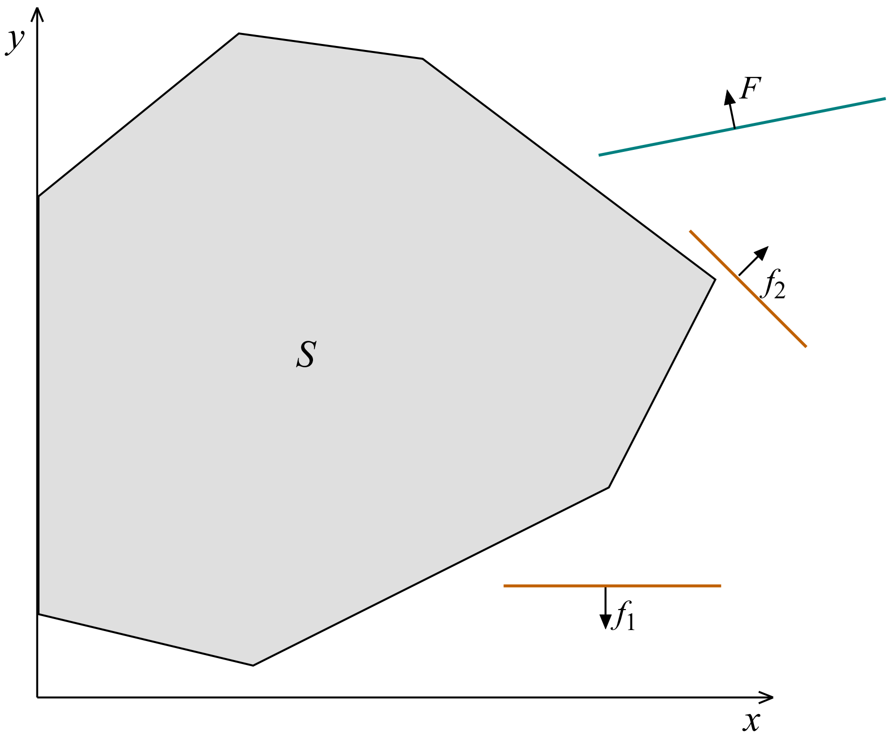
LINEAR SVBLP - EXAMPLE 1

$\max \; F(x, y) = -x + 5y$
s.t.

$\max \; f_1(y) = \; - y$
$\max \; f_2(y) = \; x + y$

s.t.

$$\left.\begin{array}{l} x - 2y \le 4 \\ 2x - y \le 24 \\ 3x + 4y \le 96 \\ x + 7y \le 126 \\ -4x + 5y \le 65 \\ x + 4y \ge 8 \\ x, y \ge 0 \end{array}\right\} S$$



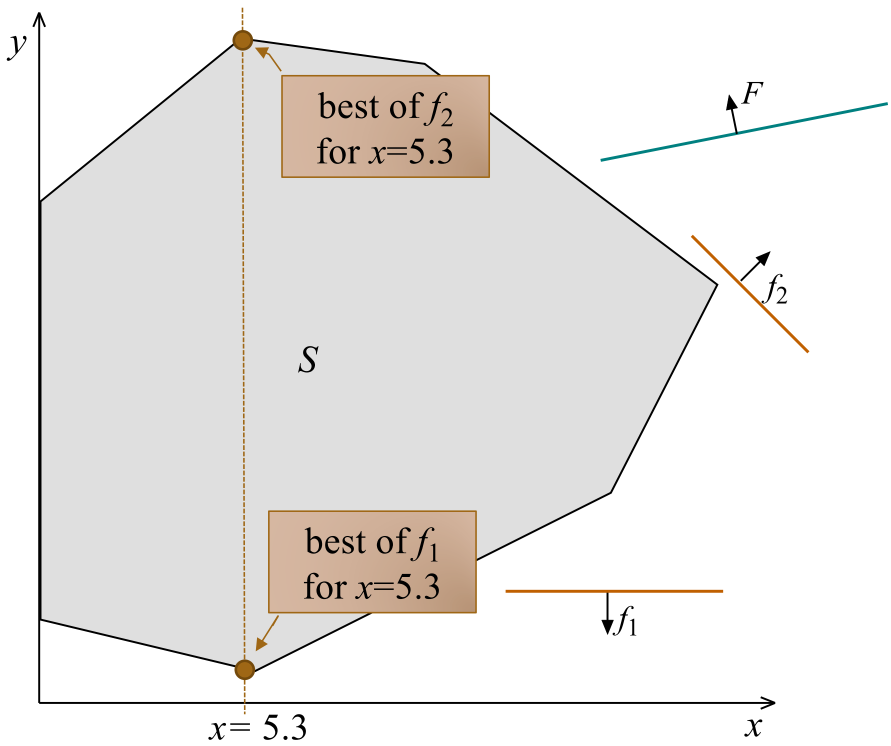
LINEAR SVBLP - EXAMPLE 1

$$\max \; F(x, y) = -x + 5y$$
s.t.
$$\max \; f_1(y) = \quad -y$$
$$\max \; f_2(y) = \; x + y$$
$$\text{s.t.} \; (x, y) \in S$$

Suppose the leader takes a decision:
e.g. $x = 5.3$

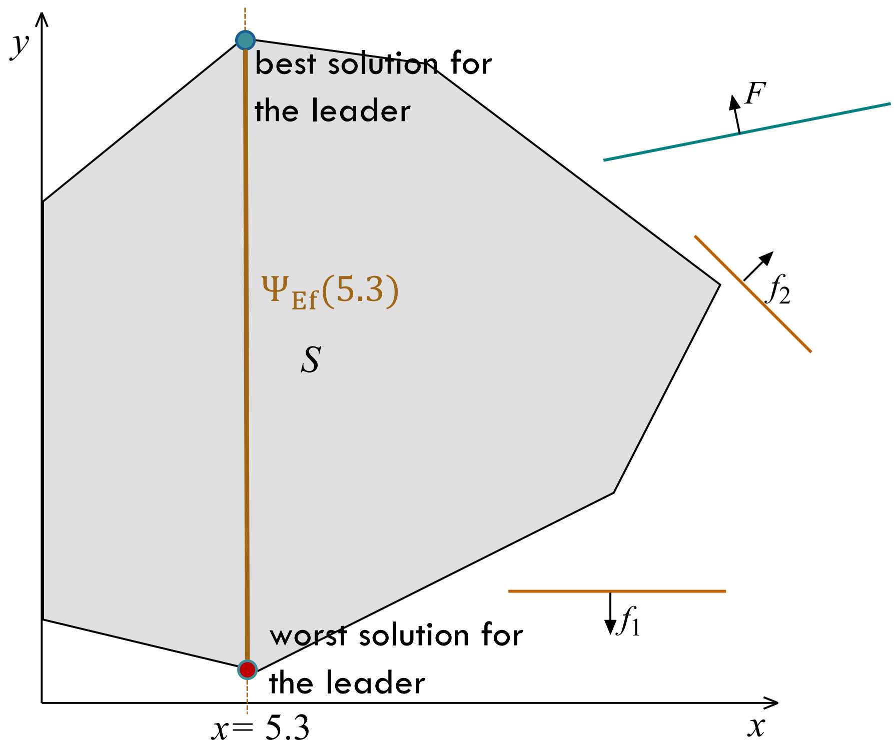
The follower's objectives become:

$$\max f_1 = \quad -y$$
$$\max f_2 = 5.3 + y$$



LINEAR SVBLP - EXAMPLE 1

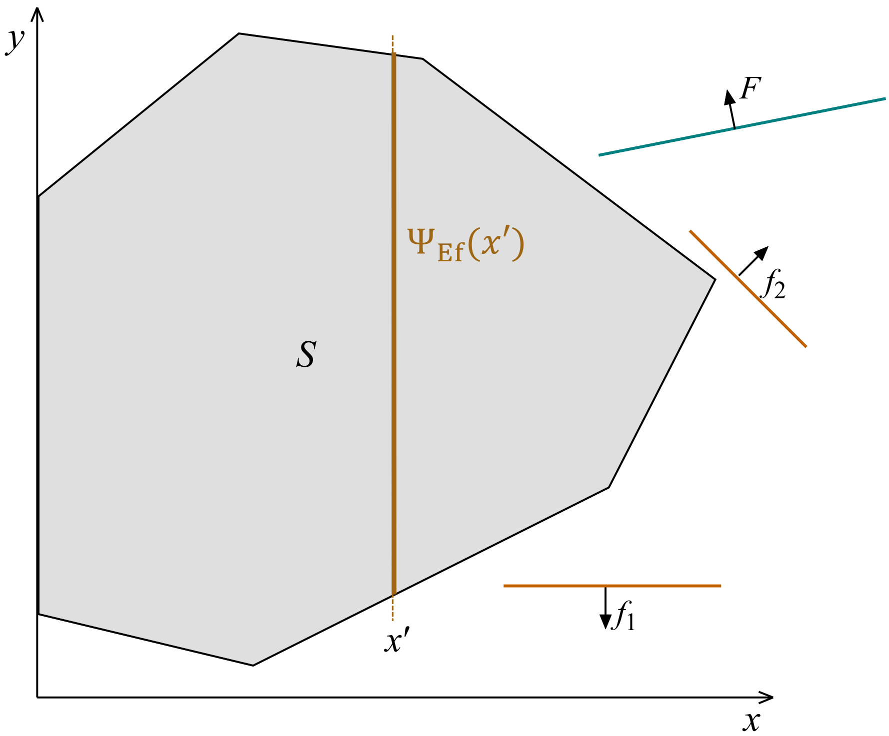
All the solutions in $\Psi_{\text{Ef}}(5.3)$ are **efficient** to the follower for the leader's decision $x = 5.3$



What will be the choice of the follower?

LINEAR SVBLP - EXAMPLE 1

For each leader's decision x' there is a set of efficient solutions $\Psi_{\mathrm{Ef}}(x')$ to the follower.



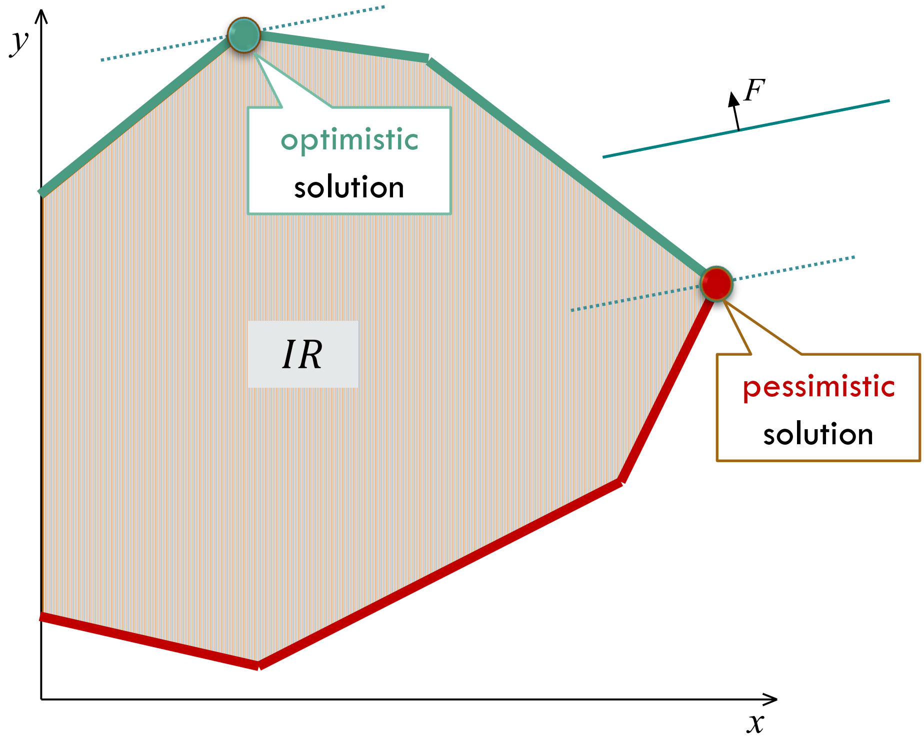
LINEAR SVBLP - EXAMPLE 1

For each leader's decision x' there is a set of efficient solutions $\Psi_{\mathrm{Ef}}(x')$ to the follower.

The leader may believe that the follower will choose a solution that most benefits the leader – **optimistic approach**

Or the leader prepares for the worst case – **pessimistic approach**

Intermediate approaches may exist.



DIFFICULTIES FOR THE LEADER IN SVBLP/MOBLP

- The follower's decision may be very difficult to anticipate.
- Almost all literature devoted to SVBLP and MOBLP has adopted the **optimistic** approach
  - The optimistic approach assumes that the follower is indifferent to all efficient solutions obtained for a given decision of the

However, this assumption is seldom realistic in many practical decision problems

OPTIMISTIC APPROACH: REALISTIC?

Ex: **a toll-setting problem**

- The owners of a highway system have to set tolls and they want to **maximize total revenue;** users of highways want to **minimize travel costs** and **minimize travel times.**
- For each leader's decision, the follower has a set of efficient solutions:
  - At one extreme, the solution that **minimizes cost** → drivers do not use highways → the leader's revenue would be 0.
  - At the other extreme, the solution that **minimizes time** → many drivers use highways → high revenues for the leader.

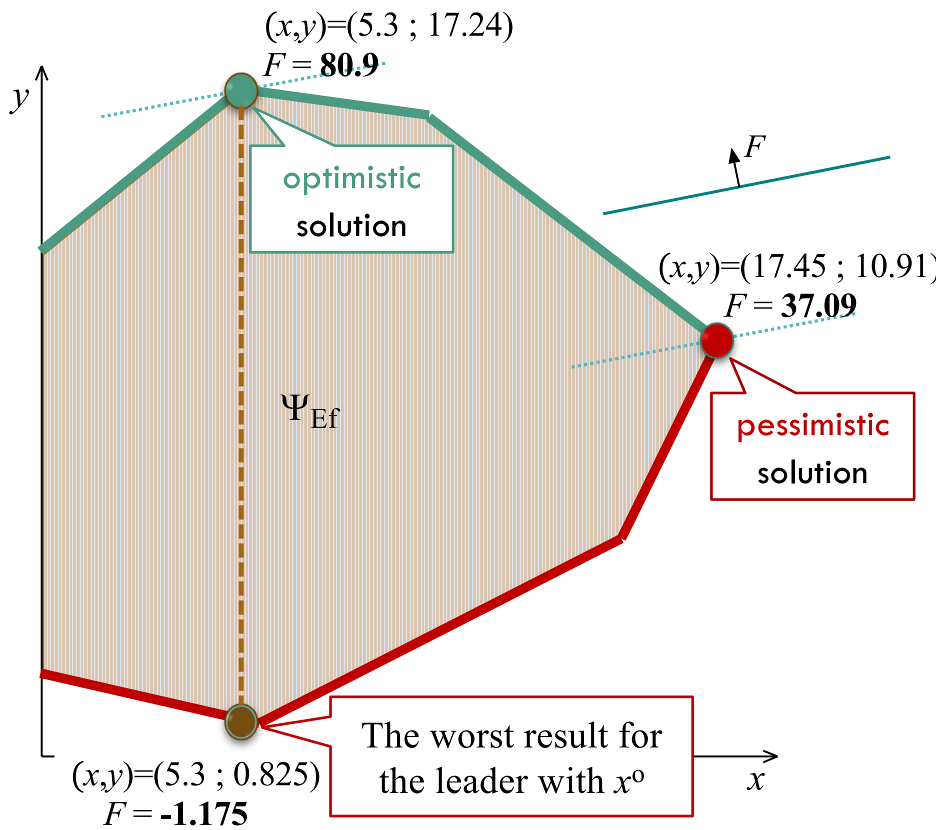
The **optimistic approach assumes that the follower is indifferent to those two opposite situations**, which would lead the leader to set very high prices for tolls.

WHAT HAPPENS IF THE OPTIMISTIC APPROACH FAILS?

If the leader takes the best decision according to the **optimistic** approach (x^o), but the follower's choice (y') is the worst for the leader.

The solution obtained (x^o, y') may be **significantly worse** than the pessimistic solution.

→ this solution will be called deceiving solution



MOST COMMON APPROACHES

- **Optimistic** approach
  - assumes that the follower accepts any efficient solution to the lower-level problem and, so, his response is always the most convenient for the leader
- **Pessimistic** approach
  - the leader is risk-averse and prepares for the worst case.

OPTIMISTIC FORMULATION IN SVBLP

- Assumes that the follower's response is always the **best** for the leader

$$\max_{x,y} \quad F(x,y) \qquad \text{(optimistic formulation)}$$
$$\text{s.t.} \quad G(x) \leq 0, y \in \Psi_{\mathrm{Ef}}(x)$$

→ the **optimistic solution,** $(x^{\mathrm{o}}, y^{\mathrm{o}})$ is a solution that optimizes the optimistic formulation

$$(x^{\mathrm{o}}, y^{\mathrm{o}}) \leftarrow \max_{x,y} \left\{F(x,y): G(x) \leq 0, y \in \Psi_{\mathrm{Ef}}(x)\right\}$$

PESSIMISTIC FORMULATION IN SVBLP

- Assumes that the follower's response is always the **worst** for the leader

$$\max_{x} \min_{y} F(x,y) \qquad \text{(pessimistic formulation)}$$
$$\text{s.t.} \quad G(x) \leq 0, y \in \Psi_{\mathrm{Ef}}(x)$$

→ the **pessimistic solution,** $(x^{\mathrm{p}}, y^{\mathrm{p}})$ is a solution that optimizes the pessimistic formulation

$$(x^{\mathrm{p}}, y^{\mathrm{p}}) \leftarrow \max_{x} \left\{ \min_{y} \{F(x,y): y \in \Psi_{\mathrm{Ef}}(x)\}: G(x) \leq 0 \right\}$$

OPTIMISTIC VS. PESSIMISTIC APPROACHES

- It is important to **acknowledge the risk the leader** takes if he adopts an **optimistic approach**
- The **pessimistic approach** is the most conservative, because it pays attention to the worst choices of the follower for the leader's interests (but it may also offer **opportunities**).
- Other types of solutions can provide useful information to the leader about the risk/opportunity he takes when making a specific decision, in particular:
  - the result of a "**failed**" **optimistic approach: deceiving solution**
  - the result of a "**successful**" **pessimistic approach: rewarding solution**

"FAILED" OPTIMISTIC APPROACH: DECEIVING SOLUTION

- **Deceiving solution, $(x^{\mathrm{d}}, y^{\mathrm{d}})$:** if the leader makes an **optimistic decision** and the follower's reaction is the **worst** for the leader.
- It gives an indication of the maximum risk the leader will take if he adopts an optimistic approach.

$$x^{\mathrm{d}} = x^{\mathrm{o}} \qquad (\textit{optimistic x-vector})$$

$$y^{\mathrm{d}} \leftarrow \min_{y}\{F(x^{\mathrm{o}}, y): y \in \Psi_{\mathrm{Ef}}(x^{\mathrm{o}})\}$$

"SUCCESSFUL" PESSIMISTIC APPROACH: REWARDING SOLUTION

- **Rewarding solution,** (x^r, y^r): if the leader makes a **pessimistic decision** and the follower's reaction is the **best** for the leader.
- It gives an indication of the opportunity the leader will have if he adopts a pessimistic approach.

$$x^r = x^p \qquad (\textit{pessimistic } x\textit{-vector})$$

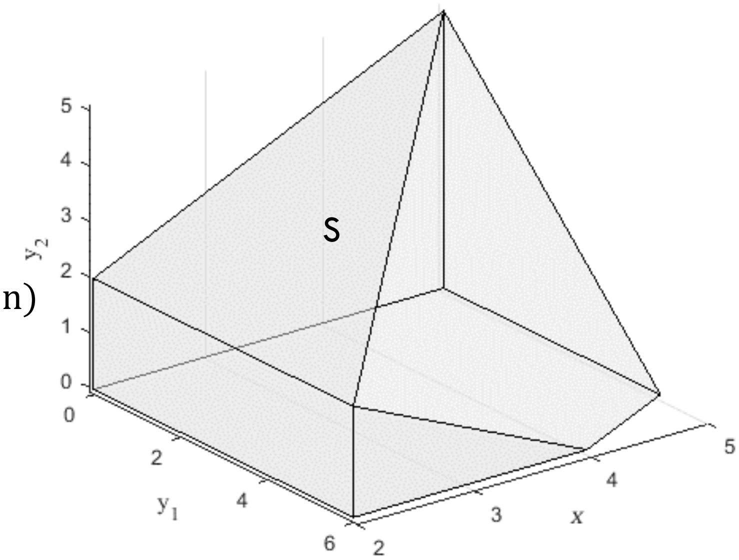
$$y^r \leftarrow \max_y \{F(x^p, y) : y \in \Psi_{Ef}(x^p)\}$$

OPTIMISTIC AND **DECEIVING**
PESSIMISTIC AND **REWARDING** SOLUTIONS

- These four solutions represent "**extreme**" outcomes that can provide the leader important insights about the ranges of possible values for his objective function.

LINEAR SVBLP - EXAMPLE 2

$$\max \quad F = x + 2y_1 - y_2$$
$$s.t.$$
$$\left.\begin{aligned} & 2 \le x \le 5 \\ & \max \quad f_1 = y_1 + 2y_2 \\ & \max \quad f_2 = y_1 - y_2 \\ & s.t. \; y_1 \le 6 \\ & \quad y_1 + y_2 \le 10 - x \\ & \quad y_2 \le x \\ & \quad y_1, y_2 \ge 0 \end{aligned}\right\} IR \text{ (induced region)}$$

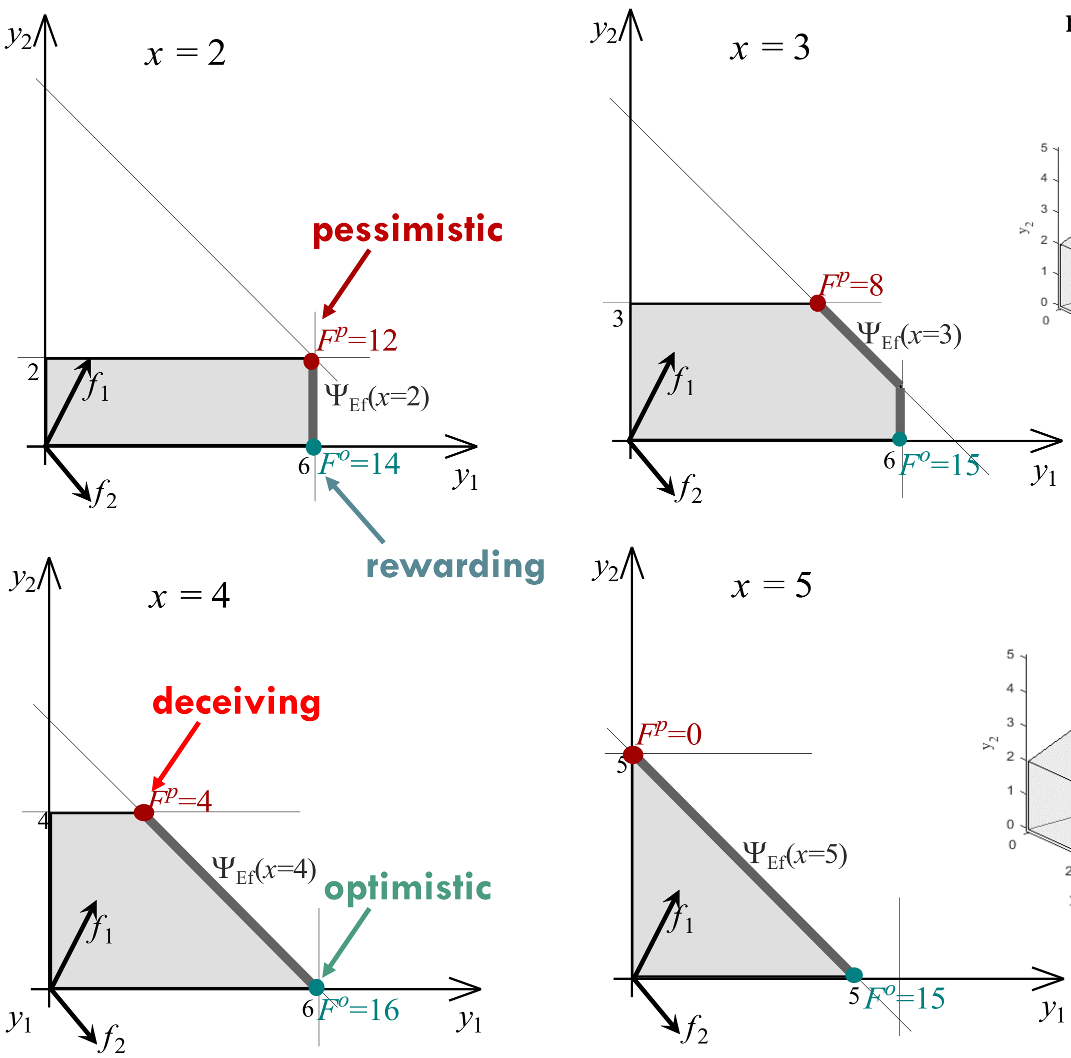


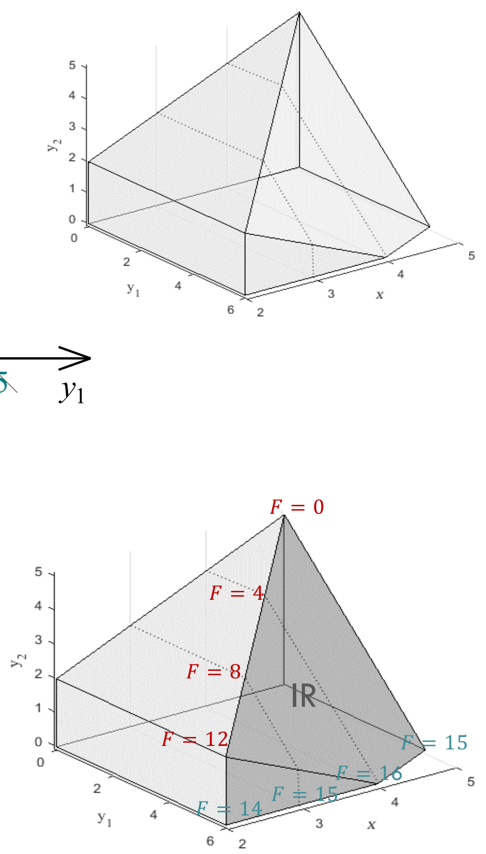
S≡ constraint region

x is the upper level variable

y_1, y_2 are the lower level variables

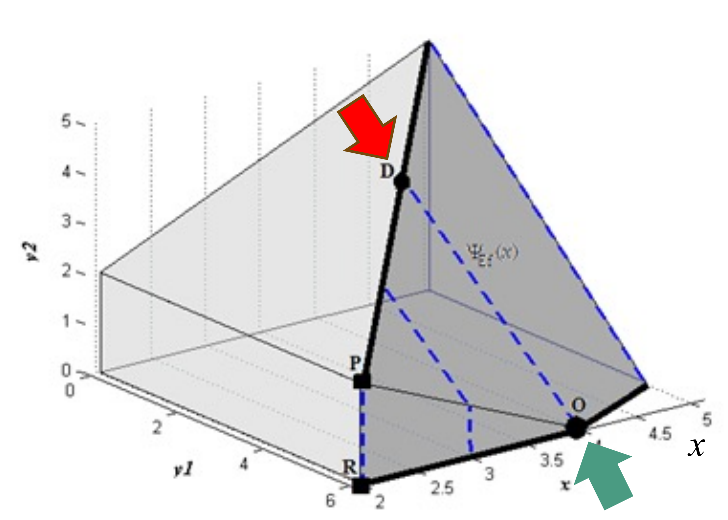
$$\max \quad F = x + 2y_1 - y_2$$





LINEAR SVBLP - EXAMPLE 2

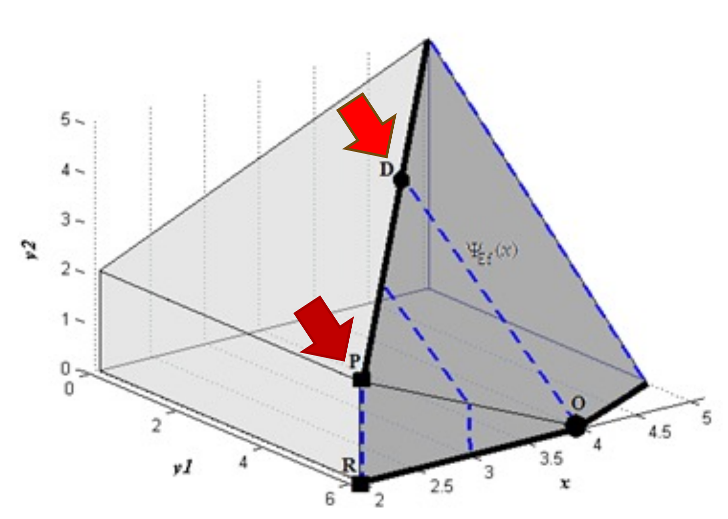
- The leader may take a **high risk** if he adopts an optimistic approach
  - because there is a **large difference** in *F* between the **optimistic** and the **deceiving** solutions



F (max)	Optimistic approach	Pessimistic approach
optimistic / rewarding solutions	**$F^O = 16$**	$F^R = 14$
deceiving / pessimistic solutions	**$F^D = 4$**	$F^P = 12$

LINEAR SVBLP - EXAMPLE 2

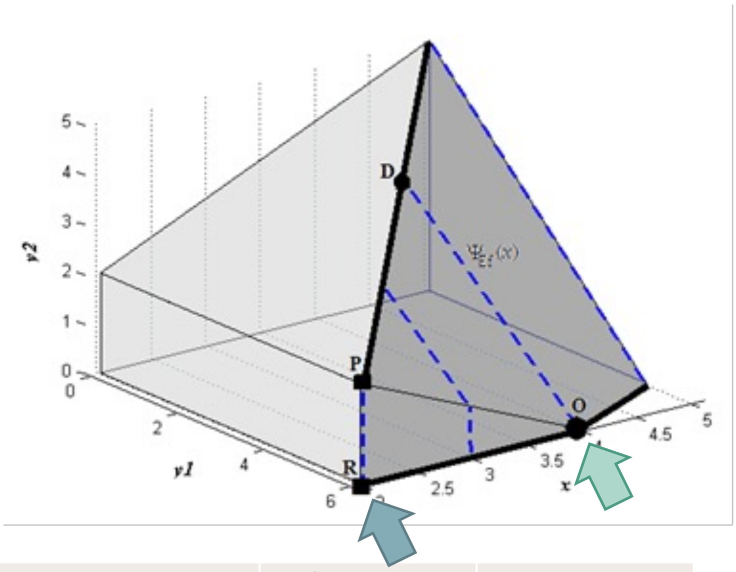
- the **deceiving** solution is **significantly worse** than the **pessimistic** one.



***F* (max)**	Optimistic approach	Pessimistic approach
optimistic / rewarding solutions	$F^O = 16$	$F^R = 14$
deceiving / **pessimistic** solutions	$\mathbf{F^D = 4}$	$\mathbf{F^P = 12}$

LINEAR SVBLP - EXAMPLE 2

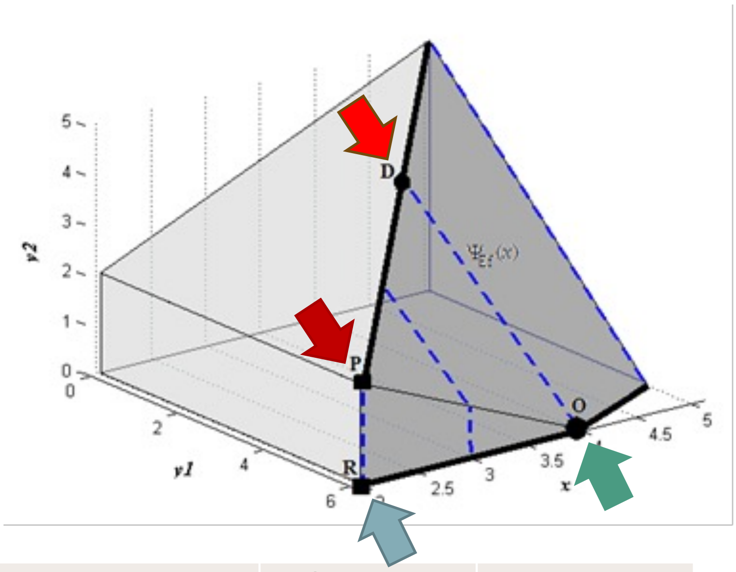
- and the value of *F* in the **rewarding** solution is **close** to the **optimistic** one



***F* (max)**	Optimistic approach	Pessimistic approach
optimistic / **rewarding** solutions	$F^O = 16$	$F^R = 14$
deceiving / pessimistic solutions	$F^D = 4$	$F^P = 12$

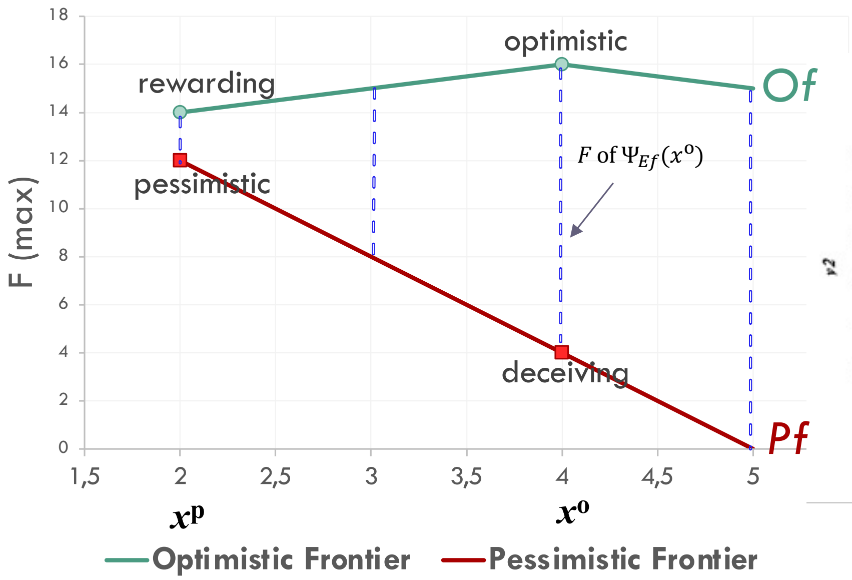
LINEAR SVBLP - EXAMPLE 2

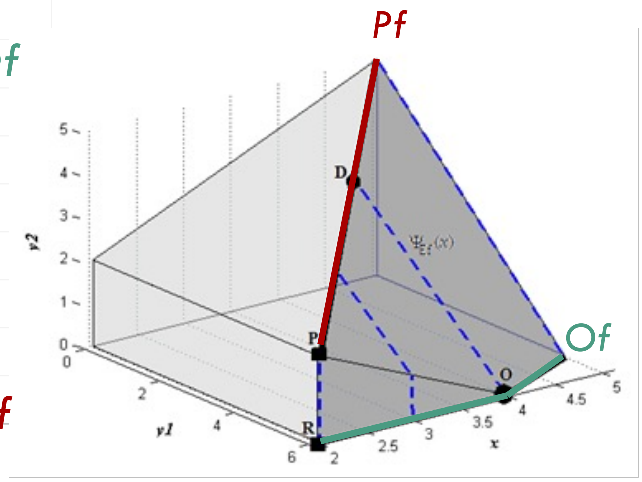
- So, in this problem, a **pessimistic approach could be more advisable** than an optimistic one, since a high risk is associated with the optimistic approach.



F **(max)**	Optimistic approach	Pessimistic approach
optimistic / **rewarding** solutions	**$F^O = 16$**	**$F^R = 14$**
deceiving / **pessimistic** solutions	**$F^D = 4$**	**$F^P = 12$**

OPTIMISTIC AND **P**ESSIMISTIC FRONTIERS
- EXAMPLE 2

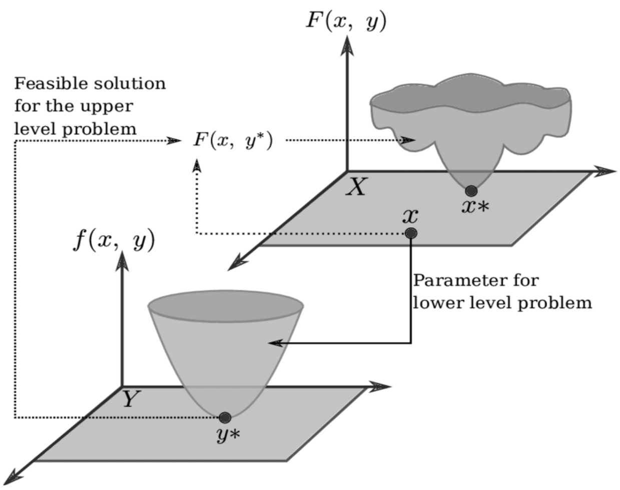




META-HEURISTIC APPROACHES

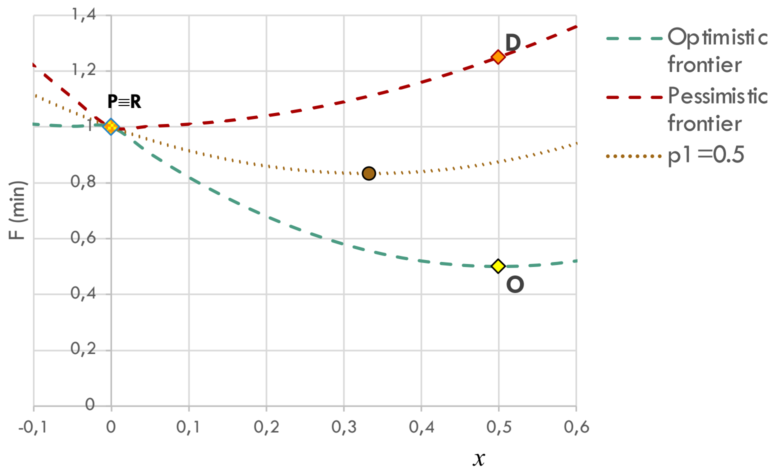
In particular involving population-based meta-heuristics:

- Meta-heuristics (EA, PSO, DE, etc.) at both levels
- Hybrid algorithms with meta-heuristics at the upper level and NLP/MILP/etc. *solver* to solve the lower-level problem for each instantiation of the upper-level variables x.

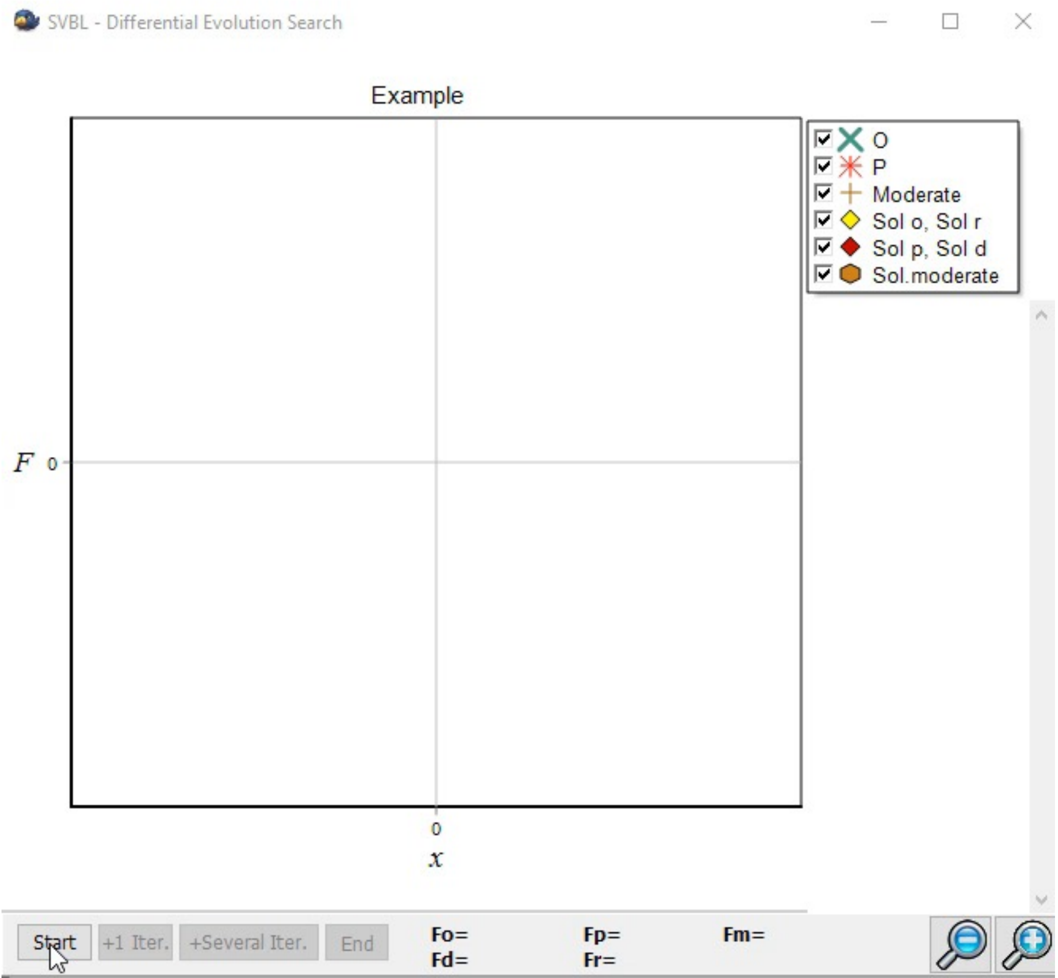


COMPUTING SOLUTIONS FOR A NONLINEAR BLP

- Approximating the *4 extreme* solutions and a *moderate solution* using **Differential Evolution**



moderate solution: optimizes the expected value of the leader's objective function, considering that p1 is the probability of the follower to choose the optimistic solution for a given *x*, and (1-p1) is the probability to choose the pessimistic solution.


SVBL - Differential Evolution Search
Example
O
P
Moderate
Sol o, Sol r
Sol p, Sol d
Sol.moderate
F
0
0
x
Start
+1 Iter.
+Several Iter.
End
Fo=
Fd=
Fp=
Fr=
Fm=

p1=0.5

FINAL REMARKS FOR SVBL

Multiple objective functions at the lower level pose additional difficulties for the leader to anticipate the follower's reaction

➔ different types of solutions can provide broader information to the leader

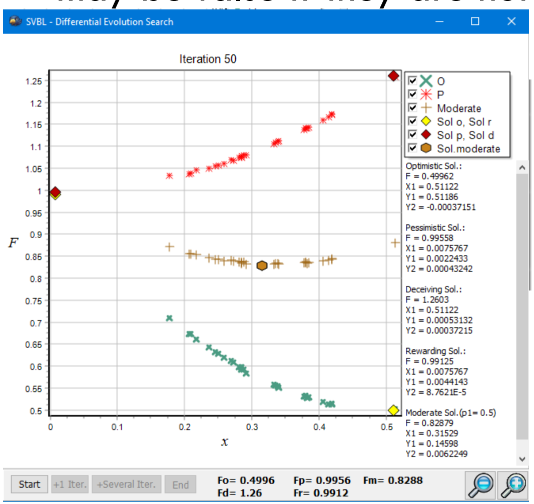
➔ the **optimistic, pessimistic, deceiving** and **rewarding** solutions delimit ranges of possible values for the leader taking into account follower's **extreme** decisions

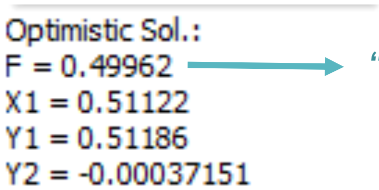
➔ **moderate** solutions may provide further decision support to the leader: optimize the expected value for the leader's objective function considering different probabilities of the follower's decision being in favor or against the interests of the leader.

But these problems are very difficult to solve…

FINAL REMARKS FOR SVBL

- Population-based meta-heuristics have been increasingly used to deal with SVBP/MOBLP
- It is very difficult to evaluate results from approximate algorithms, which may lead to **misleading results**: apparently better solutions may be false if they are not efficient to the lower level





→ *"better"* than F_o^*=0.5

Part III

Bilevel Optimization: An Application in the Energy Sector

DEMAND RESPONSE AS A BILEVEL OPTIMIZATION PROBLEM

- The electricity retailer (leader) determines dynamic time-of-use prices to maximize profit
- The consumer (follower) reacts scheduling load operation to minimize the electricity bill (to profit from periods of low energy prices), subject to comfort requirements (quality of service, associated with appliance operation)

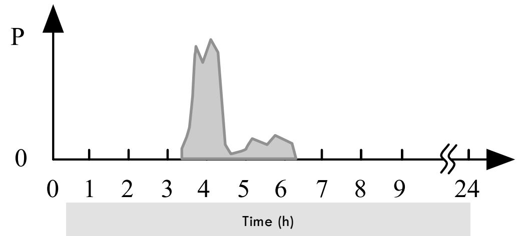
BILEVEL MODEL FOR THE INTERACTION BETWEEN RETAILER AND CONSUMER

- The retailer (*leader*) determines the prices x_i $(i=1,\dots,I)$ to be charged to the consumer in each predefined sub-period P_i of the planning period $T=\{1,\dots,T\}$ → maximize profit
- Knowing the electricity prices, the consumer (*follower*) reacts by means of re-scheduling appliance operation → minimize cost
- The consumer's energy decisions affects the retailer's profit

DEMAND RESPONSE

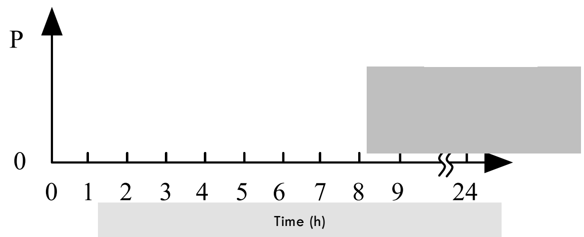
Shiftable loads (operation cycle cannot be interrupted):

- Dishwasher
- Clothes dryer
- Washing machine



Interruptible loads

- Electric water heater
- Electric vehicle



Thermostatic controlled loads

- Air conditioning system

BILEVEL MODEL

retailer's profit

$$\max_x F = \overbrace{\sum_{i=1}^{I}\sum_{t\in P_i} x_i\left(b_t + \sum_{j=1}^{J} p_{jt} + \sum_{k=1}^{K} q_{kt}\right)}^{SE=\text{sale of energy to consumers}} + \overbrace{\sum_{l=1}^{L} e_l u_l}^{CP=\text{contracted power}} \overbrace{- \sum_{t=1}^{T} \pi_t\left(b_t + \sum_{j=1}^{J} p_{jt} + \sum_{k=1}^{K} q_{kt}\right)}^{\text{cost of buying energy}}$$

s. to

$x_i \leq \overline{x_i}\,, \quad i = 1, \dots, I$

$x_i \geq \underline{x_i}\,, \quad i = 1, \dots, I$

constraints on energy prices

$\frac{1}{T}\sum_{i=1}^{I} \overline{P}_i x_i = x^{AVG}$

$\min f = \sum_{i=1}^{I} \sum_{t\in P_i} x_i (b_t + \sum_{j=1}^{J} p_{jt} + \sum_{k=1}^{K} q_{kt}) + \sum_{l=1}^{L} e_l u_l$

consumer's cost

s. to

operation of shiftable appliances

operation of interruptible appliances

power component

CASE ANALYSIS: DATA

Retailer's (UL) problem

- 6 tariff periods *Pi*
- 6 continuous variables
- 13 constraints (lower/upper bounds + average price)

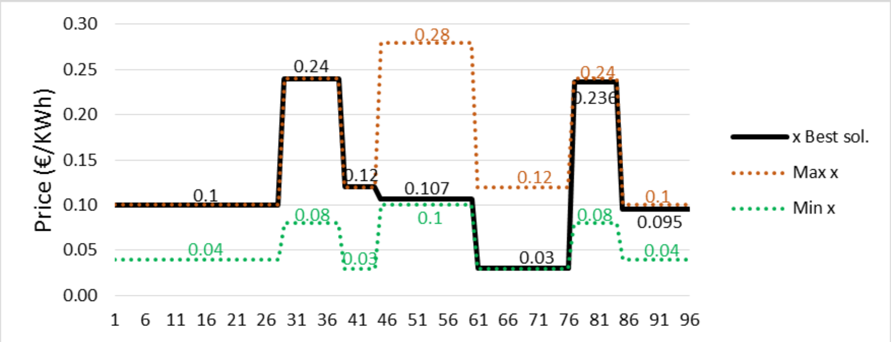
Consumer's (LL) problem

- 3 shiftable appliances (J=3) – dish + laundry +dryer machines
- 2 interruptible loads (K = 2) – EV + EWH
- non-controllable base load

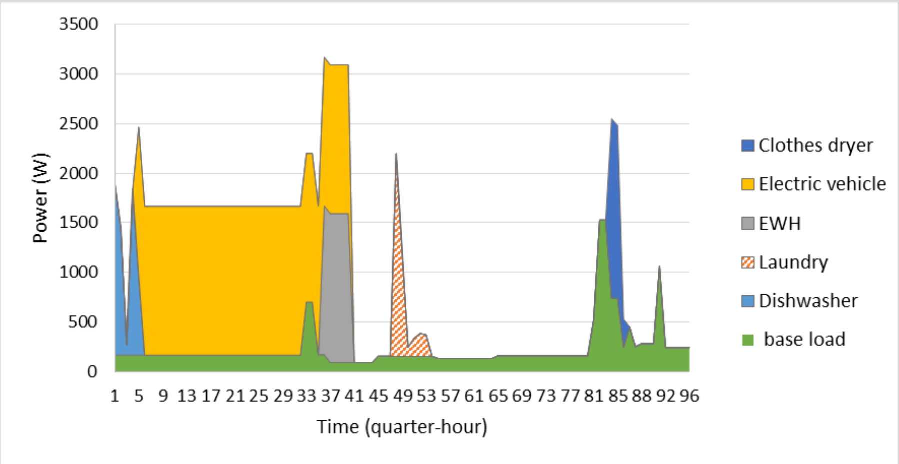
24 h planning horizon discretized in 15 min time intervals
For each instantiation of the UL variables, the LL MILP problem has 559 binary variables, 691 continuous variables and 2389 constraints.

ILLUSTRATIVE RESULTS

Electricity prices



Load diagram



CONCLUSIONS

- Main concepts on single-objective bilevel optimization
- Semi-vectorial bilevel optimization – optimistic, pessimistic, deceiving, and rewarding solutions
- An overview of an application of bilevel optimization in the design of time-of-use tariffs