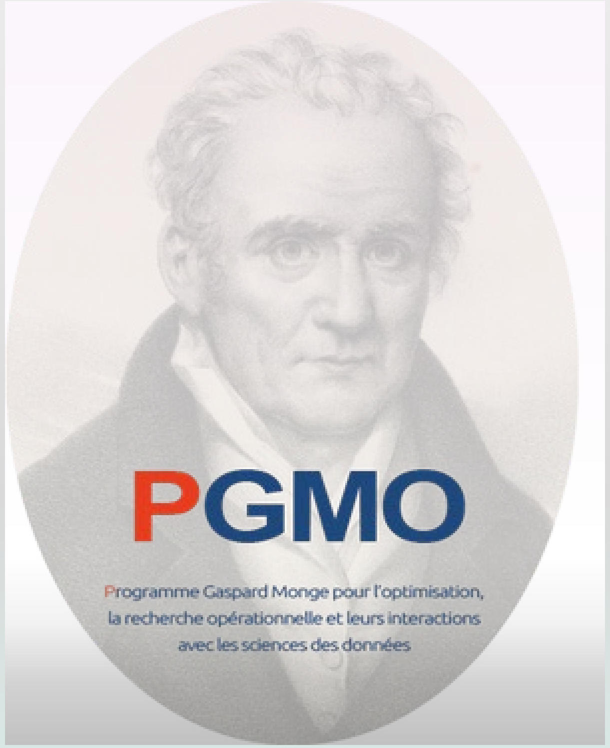
BILEVEL OPTIMIZATION UNDER UNCERTAINTY

IVANA LJUBIC

ESSEC BUSINESS SCHOOL, PARIS

PGMO DAYS, NOV 30, 2022, PARIS

ESSEC
BUSINESS SCHOOL

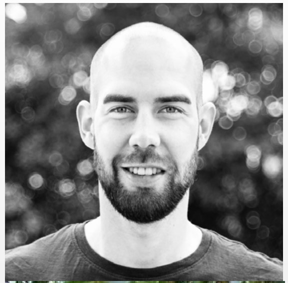


BASED ON OUR RECENT ARTICLES:

A Survey on Mixed-Integer Programming Techniques in Bilevel Optimization
EURO Journal on Computational Optimization. 2021. DOI: 10.1016/j.ejco.2021.100007
Jointly with Thomas Kleinert, Martine Labbé, and Martin Schmidt

A Survey on Bilevel Optimization Under Uncertainty
Jointly with Yasmine Beck and Martin Schmidt, Optimization Online, 2022

A Brief Introduction to Robust Bilevel Optimization
Jointly with Yasmine Beck and Martin Schmidt,
Views-and-News of the SIAM Activity Group on Optimization, to appear 2022

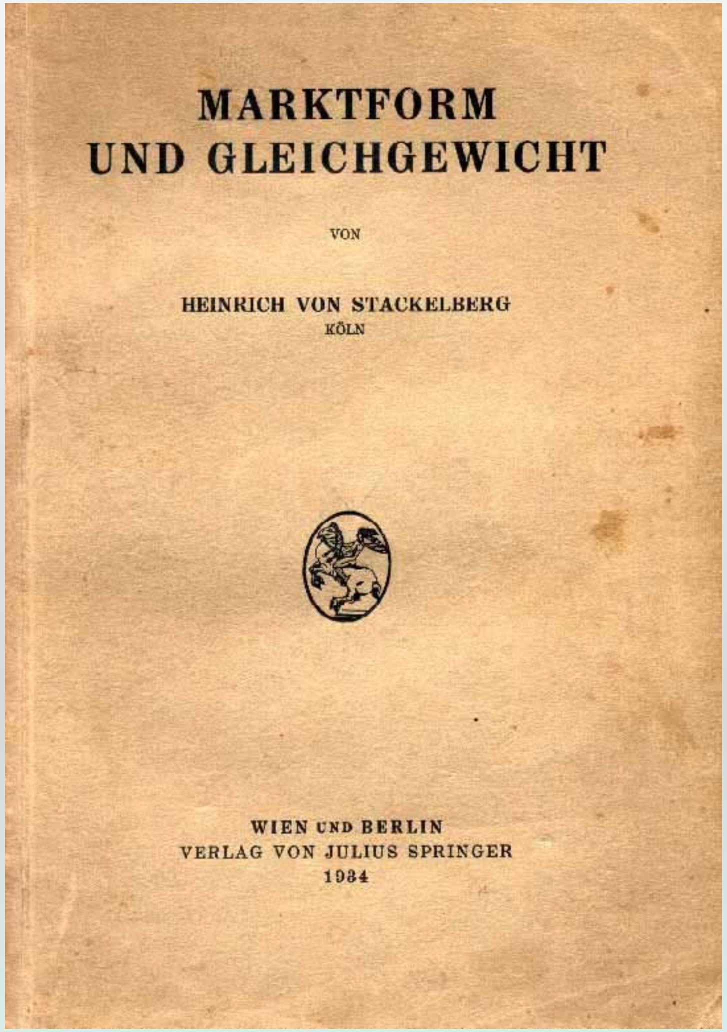


BILEVEL OPTIMIZATION

WITH DETERMINISTIC DATA

STACKELBERG GAMES

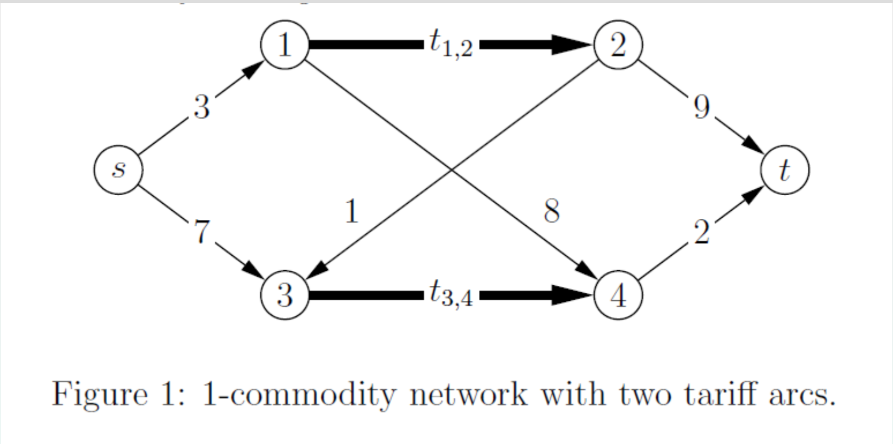
- Introduced in economy by H. v. Stackelberg in 1934
- Two-player sequential game: LEADER and FOLLOWER
- The LEADER moves before the FOLLOWER
- **Perfect information:** the leader has a perfect knowledge of the followers strategy
- The follower observes leader's action and acts rationally
- **Rationality:** agents act optimally, maximizing their payoffs
- BILEVEL OPTIMIZATION: Bracken & McGill (1973), Candler & Norton (1977)



APPLICATIONS: PRICING

Pricing: operator sets tariffs, and then customers choose the cheapest alternative

- **Tariff-setting, toll optimization (Labbé et al., 1998; Brotcorne et al., 2001; Labbé & Violin, 2016)**
- **Network Design and Pricing (Brotcorne et al., 2008)**
- **Survey (van Hoesel, 2008)**



A DETERMINISTIC BILEVEL PROBLEM

$$\text{“}\min_{x \in X}\text{”} \quad F(x,y) \tag{1a}$$
$$\text{s.t.} \quad G(x,y) \geq 0, \tag{1b}$$
$$y \in S(x), \tag{1c}$$

upper level

$S(x)$: the set of optimal solutions of the x-parameterized problem

$$\min_{y \in Y} \quad f(x,y) \tag{2a}$$
$$\text{s.t.} \quad g(x,y) \geq 0. \tag{2b}$$

lower level

- **Both levels may involve integer decision variables. Functions can be non-linear, non-convex…**
- **(1) could be ill-posed (if LL solution is not unique). “min” to be replaced by**

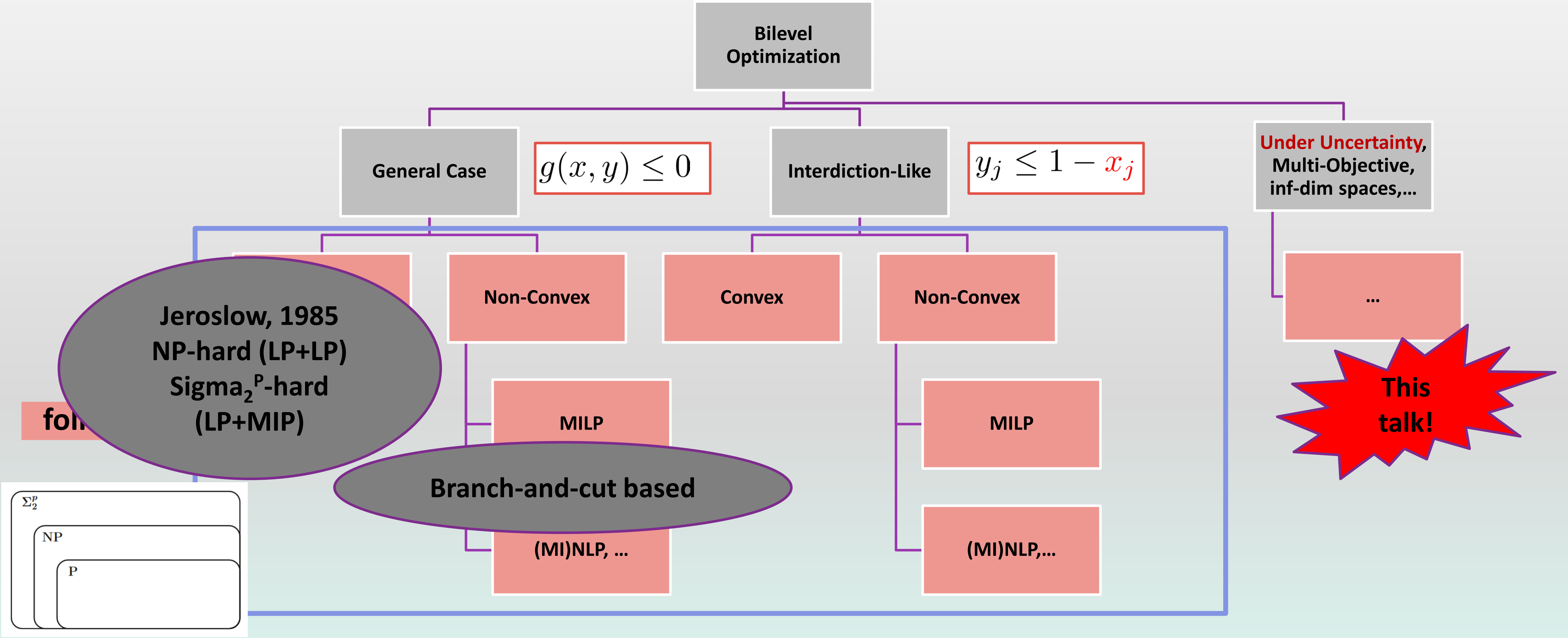
$$\min_{x \in \bar{X}} \min_{y \in S(x)} \quad F(x,y)$$

Optimistic!

$$\min_{x \in \bar{X}} \max_{y \in S(x)} \quad F(x,y)$$

Pessimistic!

OVERVIEW OF BILEVEL OPTIMIZATION PROBLEMS



THIS TALK

- **From deterministic bilevel optimization to bilevel optimization under uncertainty**

- **Sources of uncertainty**
  - **Data uncertainty**
  - **Decision uncertainty**

- **Timing for the data uncertainty**
  - **Here-and-now follower**
  - **Wait-and-see follower**

- **Challenges & opportunities**

SOURCES OF UNCERTAINTY

UNCERTAINTY: SINGLE-LEVEL VS BILEVEL

Single-level optimization:

$$\min_{x}\{c^\top x\colon Ax \geq b\}$$

- **"Only" subject to data uncertainty in A,b,c**
- **Stochastic optimization**
- **Robust optimization**
- **Distributionally robust, etc**

Bilevel optimization:

$$\text{“}\min_{x\in X}\text{”} \quad F(x,y) \tag{1a}$$

$$\text{s.t.} \quad G(x,y) \geq 0, \tag{1b}$$

$$y \in S(x), \tag{1c}$$

$S(x)$: optimal solutions of the x-parameterized problem

$$\min_{y\in Y} \quad f(x,y) \tag{2a}$$

$$\text{s.t.} \quad g(x,y) \geq 0. \tag{2b}$$

- **Subject to: data uncertainty**
- **But also: decision uncertainty. The leader is not sure about the reaction of the follower, or the follower is not certain about the observed leader's decision.**

TIMING OF UNCERTAINTY

WAIT-AND-SEE FOLLOWER

$$\text{leader } x \quad \curvearrowright \quad \text{uncertainty } u \quad \curvearrowright \quad \text{follower } y = y(x, u).$$

The leader is uncertain about the optimization parameters of the follower
Example: the leader solves a robust optimization problem

$$\text{“}\min_{x \in X} \max_{u \in \mathcal{U}}\text{”} \quad F(x, y) \quad \text{s.t.} \quad y \in S(x, u),$$

$$S(x, u) := \arg\min_{y \in Y} \quad f(x, u, y) \quad \text{s.t.} \quad g(x, u, y) \geq 0.$$

Example: the leader is risk-neutral wrt data uncertainty (discrete scenario set)

Optimistic or pessimistic leader

$$\text{“}\min_{x \in X}\text{”} \sum_{u \in \mathcal{U}} p_u F(x, y(x, u)) \quad \text{s.t.} \quad y(x, u) \in S(x, u), u \in \mathcal{U}$$

$$S(x, u) := \arg\min_{y \in Y} \quad f(x, u, y) \quad \text{s.t.} \quad g(x, u, y) \geq 0.$$

HERE-AND-NOW FOLLOWER

$$\text{leader } x \quad \curvearrowright \quad \text{follower } y = y(x) \quad \curvearrowright \quad \text{uncertainty } u.$$

The follower solves the problem under data uncertainty (stochastic, robust,…).

Optimistic vs pessimistic leader

For example: optimistic leader, the robust follower hedges against uncertainty in the objective function

$$\min_{x \in X} \min_{y \in S(x)} F(x, y)$$

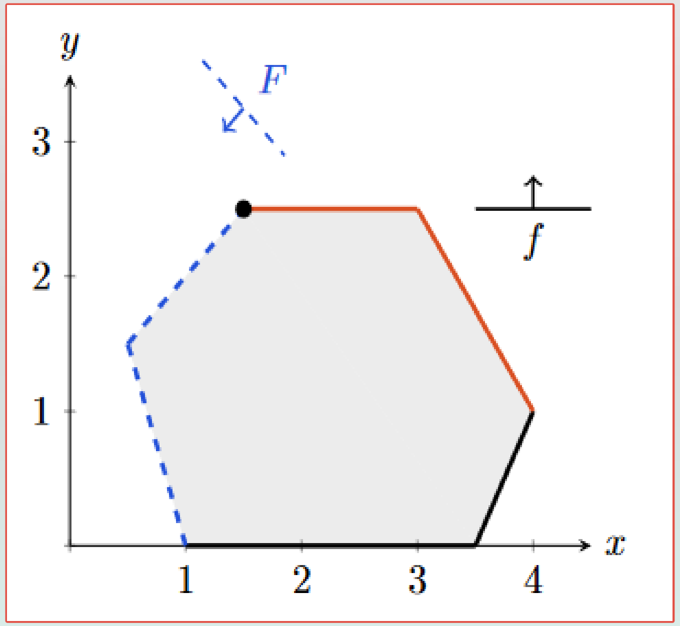
$$S(x) := \arg\min_{y' \in Y} \left\{ \max_{u \in \mathcal{U}} f(x, u, y') : g(x, y') \geq 0 \right\}.$$

A SMALL EXAMPLE

Deterministic bilevel

$$
\begin{aligned}
\text{“}\min_{x\in\mathbb{R}}\text{”} \quad & F(x,y) = x + y \\
\text{s.t.} \quad & x - y \geq -1, \\
& 3x + y \geq 3, \\
& y \in S(x),
\end{aligned}
$$

$$
\begin{aligned}
S(x) := \arg\min_{y\in\mathbb{R}} \quad & f(x,y) = -0.1y \\
\text{s.t.} \quad & -2x + y \geq -7, \\
& -3x - 2y \geq -14, \\
& 0 \leq y \leq 2.5.
\end{aligned}
$$

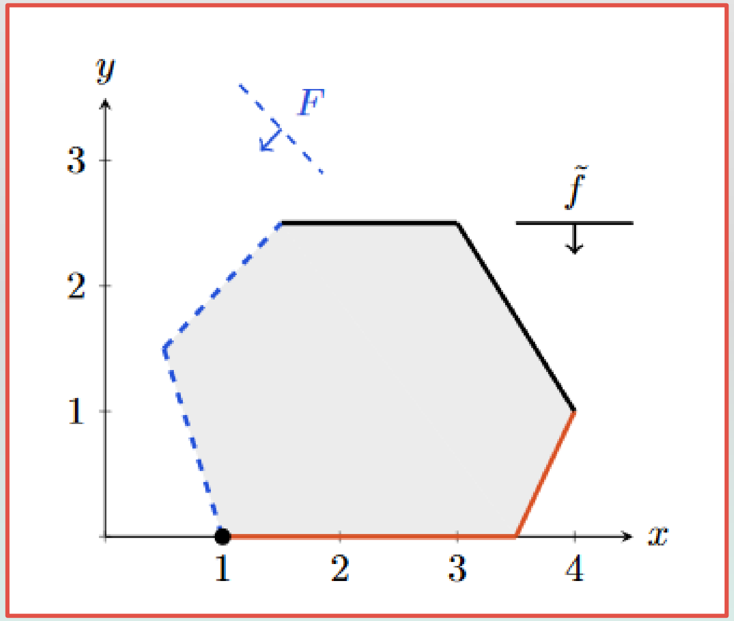


(a)

Here-and-now follower

$$
\mathcal{U} := \{u \in \mathbb{R} : |u| \leq 0.5\}
$$

$$
\begin{aligned}
S(x) := \arg\min_{y\in\mathbb{R}} \max_{u\in\mathcal{U}} \quad & \tilde{f}(x,u,y) = (-0.1 + u)y \\
\text{s.t.} \quad & -2x + y \geq -7, \\
& -3x - 2y \geq -14, \\
& 0 \leq y \leq 2.5.
\end{aligned}
$$



(b)

Wait-and-see follower

$$
\begin{aligned}
\text{“}\min_{x\in\mathbb{R}} \max_{u\in\mathcal{U}}\text{”} \quad & F(x,y) = x + y \\
\text{s.t.} \quad & x - y \geq -1, \\
& 3x + y \geq 3, \\
& y \in S(x,u),
\end{aligned}
$$

$$
\begin{aligned}
S(x,u) := \arg\min_{y\in\mathbb{R}} \quad & (-0.1 + u)y \\
\text{s.t.} \quad & -2x + y \geq -7, \\
& -3x - 2y \geq -14, \\
& 0 \leq y \leq 2.5.
\end{aligned}
$$

$$
\min_{x} \quad \hat{F}(x) \quad \text{s.t.} \quad 1.5 \leq x \leq 4
$$

$$
\hat{F}(x) = \begin{cases} x + 2.5, & 1.5 \leq x \leq 3, \\ -0.5x + 7, & 3 \leq x \leq 4. \end{cases}
$$

$u \in [-0.5, 0.1) \Rightarrow (a)$

$u \in (0.1, 0.5] \Rightarrow (b)$

$u = 0.1 \Rightarrow$ all LL sol feasible

CHALLENGES

PROBLEM COMPLEXITY

Robust single-level LPs:

- **Interval, ball, ellipsoidal, polyhedral or Gamma-uncertainty preserve "tractability" of their deterministic counterpart (Ben-Tal & Nemirovski, Bertsimas & Sim)**

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & (a+u)^T x \leq b \text{ for all } u \in \mathcal{U} \end{aligned}$$

Robust bilevel optimization:

- **Here-and-now follower: tractability of the lower-level remains preserved for these uncertainty types**
- **Continuous convex lower level: KKT-based, strong duality-based reformulations still possible**
- **Discrete lower level: branch-and-cut still possible**
- **Major challenge: much larger in size, parallelization**

Robust bilevel optimization:

- **Wait-and-see follower: the problems may climb up in the complexity hierarchy!**

ROBUST BILEVEL OPTIMIZATION

Deterministic bilevel

$$
\begin{aligned}
&\text{“}\max_{x \in X}\text{”} \quad d^T y \\
&\text{s.t.} \;\; y \in S(x) \\
&S(x) := \arg\max\{u^T y : Ay \le Bx + b\} \\
&X \subseteq \{0,1\}^{n_x}
\end{aligned}
$$

NP-hard

Robust bilevel: Wait-and-see follower

$$
\begin{aligned}
&\text{“}\max_{x \in X} \quad \min_{u \in U}\text{”} \quad d^T y \\
&\text{s.t.} \;\; y \in S(x, u) \\
&S(x, u) := \arg\max\{u^T y : Ay \le Bx + b\} \\
&X \subseteq \{0,1\}^{n_x}
\end{aligned}
$$

$$
\mathcal{U} := [u_1^-, u_1^+] \times \cdots \times [u_{n_y}^-, u_{n_y}^+]
$$

Under interval uncertainty, the robust counterpart is Sigma$_2^P$-hard
The “adversarial problem” (inner min) is NP-hard

Buchheim, Henke, Hommelsheim:
On the complexity of robust bilevel optimization with uncertain follower's objective. OR Letters 49(5): 703-707 (2021)

OPPORTUNITIES

BILEVEL STOCHASTIC MIP

Discrete scenario set

$$\min_{x\in X,y} \quad c^T x + \sum_{u\in\mathcal{U}} p_u \quad d_L^T\, y(x,u)$$

$$\text{s.t. } y(x,u) \in \arg\min_{y\in Y}\{d_F^T\, y : Ay \le B_u x + b_u\}$$

$$X \subseteq \{0,1\}^{n_x}, Y \subseteq \{0,1\}^{n_y}$$

Value function:
$\Phi(x,u) = \min_{y\in Y}\{d_F^T\, y : Ay \le B_u x + b_u\}$

Value-function reformulation (optimistic)

$$\min_{x\in X,y} \quad c^T x + \sum_{u\in\mathcal{U}} p_u \quad d_L^T\, y_u$$

$$\begin{aligned} \text{s.t.} \quad & d_F^T\, y_u \le \Phi(x,u), \quad u \in \mathcal{U} \\ & Ay_u \le B_u x + b_u, \quad u \in \mathcal{U} \\ & y_u \in Y, \quad u \in \mathcal{U} \\ & X \subseteq \{0,1\}^{n_x}, Y \subseteq \{0,1\}^{n_y} \end{aligned}$$

Single-leader, multiple independent followers

Leverage on the existing branch-and-cut methods (Fischetti et al, 2017; Tahernejad et al, 2020)

S. Bolusani, S. Coniglio, T. K. Ralphs, and S. Tahernejad, "A Unified Framework for Multistage Mixed Integer Linear Optimization," in *Bilevel optimization: advances and next challenges*, S. Dempe and A. Zemkoho, Eds., 2020, p. 513–560.

CRITICISM… AND OUTLOOK

PERFECT INFORMATION AND RATIONALITY OF DECISION MAKERS

DECISION UNCERTAINTY: EXAMPLES

- Leader hedges against sub-optimal follower reactions → **near-optimal robust bilevel models** (Besancon et al, 2019).
- If the level of cooperation/confrontation of the follower is unknown → **intermediate cases, between the optimistic and the pessimistic one** (Aboussoror & Loridan, 1995; Mallozzi & Morgan, 1996).
- The follower cannot perfectly observe the decision of the leader → hedges against all possible leader decision given the noisy observation (Bagwell, 1995; vanDamme & Hurkens,1997; Beck & Schmidt:2021).
- Limited intellectual or computational resources render it impossible for the follower to take a globally optimal decision → the **follower resorts to heuristic approaches** and the leader may be uncertain w.r.t. which heuristic is used (Zare et al, 2020).

CONCLUSIONS

- **Connections between bilevel and robust/stochastic optimization still to be better understood**
- **When can we retain the tractability of the deterministic bilevel counterpart?**
- **When can we solve uncertain bilevel problems through a serious of deterministic ones?**
- **When do the bilevel problems under uncertainty become significantly harder?**
- **How can we better exploit the existing computational frameworks for deterministic bilevel optimization? (decomposition, SAA, scenario aggregation…)**

- **Data uncertainty vs Decision uncertainty, which paradigm to follow?**

LITERATURE

- L. Brotcorne, M. Labbé, P. Marcotte, and G. Savard. A Bilevel Model for Toll Optimization on a Multicommodity Transportation Network, Transportation Science, 35(4): 345-358, 2001
- L. Brotcorne, M. Labbé, P. Marcotte, and G. Savard. Joint design and pricing on a network. Operations Research, 56 (5):1104–1115, 2008
- S. Dempe. Bilevel optimization: theory, algorithms and applications, TU Freiberg, ISSN 2512-3750. Fakultät für Mathematik und Informatik. PREPRINT 2018-11

LITERATURE, CONT.

- R.G. Jeroslow. The polynomial hierarchy and a simple model for competitive analysis. Mathematical Programming, 32(2):146–164, 1985
- M. Labbé, P. Marcotte, and G. Savard. A bilevel model of taxation and its application to optimal highway pricing. Management Science, 44(12):1608–1622, 1998
- S. van Hoesel. An overview of Stackelberg pricing in networks. European Journal of Operational Reseach, 189:1393–1492, 2008