



WORKING PAPER NO. 541

Regularization and Approximation Methods in Stackelberg Games and Bilevel Optimization

Francesco Caruso, Maria Beatrice Lignola and Jacqueline Morgan

September 2019

This version September 2020



University of Naples Federico II



University of Salerno



Bocconi

Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS – UNIVERSITY OF NAPLES
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it
ISSN: 2240-9696

WORKING PAPER NO. 541

Regularization and Approximation Methods in Stackelberg Games and Bilevel Optimization

Francesco Caruso^{*}, Maria Beatrice Lignola^{} and Jacqueline Morgan^{***}**

Abstract

In a two-stage Stackelberg game, depending on the leader's information about the choice of the follower among his optimal responses, one can associate different types of mathematical problems. We present formulations and solution concepts for such problems, together with their possible connections in bilevel optimization, and we illustrate the crucial issues concerning these solution concepts. Then, we discuss which of these issues can be positively or negatively answered and how managing the latter ones by means of two widely used approaches: regularizing the set of optimal responses of the follower, via different types of approximate solutions, or regularizing the follower's payoff function, via the Tikhonov or the proximal regularizations. The first approach allows to obviate the lack of existence and/or stability through approximating problems, whose solutions exist under not restrictive conditions and enable to construct a surrogate solution to the original problem. The second approach permits to overcome the non-uniqueness of the follower's optimal response, by constructing sequences of Stackelberg games with a unique second-stage solution which approximate in some sense the original game, and to select among the solutions by using a constructive method with behavioural motivations.

Keywords: Two-stage Stackelberg game; bilevel optimization; pessimistic and optimistic problem; intermediate Stackelberg problem; subgame perfect Nash equilibrium; existence and stability of solutions; constructive selection method; follower's optimal reaction set approximation; approximate solution; viscosity solution; follower's payoff function approximation; Tikhonov method; proximal regularization; cost-to-move.

^{*} Università di Napoli Federico II. Postal address: Department of Economics, Università di Napoli Federico II, Via Cintia (Monte S. Angelo), 80126 Napoli, Italy. E-mail: francesco.caruso@unina.it

^{**} Università di Napoli Federico II and CSEF. Postal address: Department of Economics, Università di Napoli Federico II, Via Cintia (Monte S. Angelo), 80126 Napoli, Italy. E-mail: morgan@unina.it

^{***} Università di Napoli Federico II. Postal address: Department of Mathematics and Applications R. Caccioppoli Università di Napoli Federico II, Via Claudio 21, 80125, Naples, Italy. E-mail: lignola@unina.it

Table of contents

1. Introduction

2. Crucial issues in Stackelberg games

3. Regularizing the follower's optimal reaction set in Stackelberg games

3.1 Regularization of pessimistic bilevel optimization problems

3.2 Approximation of optimistic bilevel optimization problems

3.3 Regularization of bilevel problems with equilibrium constraints

4. Regularizing the follower's payoff function in Stackelberg games

4.1 Regularizing the follower's payoff function via Tikhonov method

4.2 Regularizing the follower's payoff function via proximal method

5. Conclusions

References

1 Introduction

Aim of this introductory section is to describe the interactions occurring between two agents (or players) which act sequentially and to present the mathematical problems associated to the different types of information the agents could be equipped. Such interactions are captured by the notion of *Stackelberg game*, which we intend in this section a two-person two-stage non cooperative game, denoted by $\Gamma = (X, Y, L, F)$, where

1. **in the first stage:** one player, henceforth called the *leader*, chooses an action x in his action set X ;
2. **in the second stage:** one player, henceforth called the *follower*, observes the action x chosen by the leader in the first stage and then chooses an action y in his action set Y ;
3. **after the two-stage interaction:** the leader receives $L(x, y)$ where $L: X \times Y \rightarrow \mathbb{R}$ is the leader's payoff function, and the follower receives $F(x, y)$ where $F: X \times Y \rightarrow \mathbb{R}$ is the follower's payoff function.

At the moment, we set no assumptions on the structure of the actions sets X and Y : they could be of finite or infinite cardinalities, subsets of finite or infinite dimensional spaces, strict subsets or whole spaces... Their nature will be specified each time in the sequel of the chapter. Assume that players pursue to minimize their payoff functions, as traditionally used in optimization literature where the payoff functions embed players' costs. However, since $\max f(\cdot) = -\min\{-f(\cdot)\}$, we could assume equivalently that players maximize their payoff functions (which usually appears in economics contexts, where payoff functions represent players' profits).

The follower, once seen the action x chosen by the leader, faces the optimization problem

$$(P_x): \min_{y \in Y} F(x, y),$$

i.e. finding $y \in Y$ such that $F(x, y) = \inf_{z \in Y} F(x, z)$. Then, he reacts to the leader's choice by picking an optimal response to x , that is an action belonging to the set, called the *follower's optimal reaction set*

$$M(x) = \operatorname{Arg} \min_{y \in Y} F(x, y) = \left\{ y \in Y \text{ such that } F(x, y) = \inf_{z \in Y} F(x, z) \right\}.$$

The set-valued map $M: X \rightrightarrows Y$ that assigns to any action x of the leader the set $M(x)$ defined above is the *follower's best reply correspondence*. For sake of brevity, it will be also called *argmin map*.

On the other hand, when the leader can foresee the optimal response $y(x) \in \operatorname{Arg} \min_{y \in Y} F(x, y)$ chosen for any action x by the follower, then he will pick an action

in the set $\text{Arg min}_{x \in X} L(x, y(x))$ according to such a forecast.

However, given the informative structure of the game, predicting the follower's optimal responses (or, equivalently, the response function $y(\cdot)$ of the follower) is typically a hard task for the leader, unless some additional information is available.

Depending on the additional information of the leader about the single-valuedness of M or, when M is not single-valued, about how the follower chooses an optimal response in the set $M(x)$ for any $x \in X$, we can associate to the Stackelberg game Γ different kinds of mathematical problems.

We will illustrate five problems and solution concepts connected to the Stackelberg game defined previously. The first one appears if the follower's best reply correspondence M is assumed to be single-valued (and the leader knows it), whereas the others concern situations where M is not single-valued. Their connections will be emphasized and illustrated by examples. In this introduction only the seminal papers and the first ones on the approximation of these problems will be quoted.

A. Stackelberg problems

When the follower's optimal response to any choice of the leader is unique, the leader can fully anticipate the reaction of the follower taking this into account before choosing his action and the follower behaves answering to the leader in the optimal (expected) way. Therefore, assumed that M is single-valued and $M(x) = \{m(x)\}$ for any $x \in X$, the leader faces the so-called *Stackelberg problem*

$$(SP) \begin{cases} \min_{x \in X} L(x, m(x)) \\ \text{where } m(x) \text{ is the solution to } (P_x). \end{cases}$$

The players acting in this way are often referred to be engaged in a *classical Stackelberg game*, terminology due to von Stackelberg who first investigated such an interaction in [109] where an economic model involving two firms that compete sequentially on quantities to produce is presented. Then, properties of Stackelberg solutions together with extensions of the classical Stackelberg games to a dynamic framework have been investigated in [24, 105].

Let us recall the solution concepts associated to (SP) .

Definition 1.1

- The infimum value $\inf_{x \in X} L(x, m(x))$ is the *Stackelberg value* of (SP) .
- A leader's action $\bar{x} \in X$ is a *Stackelberg solution* to (SP) if

$$\bar{x} \in \text{Arg min}_{x \in X} L(x, m(x)).$$

- An action profile $(\bar{x}, \bar{y}) \in X \times Y$ is a *Stackelberg equilibrium* of (SP) if \bar{x} is a Stackelberg solution and $\bar{y} = m(\bar{x})$.

Example 1.1 The Stackelberg leadership model described in [109] regards two firms in a market which sell homogeneous products and have to choose the quantities to produce. Firm 1, acting as the leader, chooses first a quantity $q_1 \geq 0$; then firm 2, acting as the follower, observes the quantity q_1 and chooses a quantity $q_2 \geq 0$. The firms aim to maximize their profit functions defined on $[0, +\infty[^2$ by

$$L(q_1, q_2) = q_1 P(Q) - C_1(q_1) \quad \text{and} \quad F(q_1, q_2) = q_2 P(Q) - C_2(q_2),$$

respectively, where $Q = q_1 + q_2$ is the total quantity in the market, $P(\cdot)$ is the inverse demand function and $C_i(\cdot)$ is the cost function of firm $i \in \{1, 2\}$. As often happens in economics contexts, we assume linear inverse demand function $P(Q) = \max\{0, a - bQ\}$ with $a, b \in]0, +\infty[$ and linear cost functions $C_i(q_i) = cq_i$ for any $i \in \{1, 2\}$ with $c \in]0, a[$. Hence, the optimal response of firm 2 to any quantity chosen by firm 1 is unique and given by

$$m(q_1) = \max \left\{ 0, \frac{a - bq_1 - c}{2b} \right\}, \quad \text{for any } q_1 \in [0, +\infty[$$

and, since $\text{Arg max}_{q_1 \in [0, +\infty[} L(q_1, m(q_1)) = \{(a - c)/2b\}$, the quantities produced at equilibrium are

$$\bar{q}_1 = \frac{a - c}{2b} \quad \text{and} \quad \bar{q}_2 = \frac{a - c}{4b}.$$

In Stackelberg leadership models the equilibrium quantity of the firm acting as the leader is always greater than equilibrium quantity of the firm acting as the follower (analogously for the equilibrium profits).

We point out that in [103, 71, 67] the following more general Stackelberg problem has been treated:

$$(GSP) \begin{cases} \min_{x \in X, g(x, m(x)) \leq 0} L(x, m(x)) \\ \text{where } m(x) \text{ is the solution to } (P_x), \end{cases}$$

where $g: X \times Y \rightarrow \mathbb{R}$. In such a problem the leader's constraints depend also on the follower's best reply function m , which makes it harder to manage than (SP) from the mathematical point of view. In the sequel of the chapter we will consider only constraints for the leader which do not depend on the actions of the follower.

From now on, in the next paragraphs of this section we assume that M is not single-valued.

B. Pessimistic leader: weak Stackelberg problem

Firstly, we consider an "extreme" situation where the pessimistic leader believes that the follower could choose the worst action for him in the set of optimal responses. The

leader, who would prevent himself against the worst he can obtain when the follower plays optimally, will face the *weak Stackelberg problem*, also called *pessimistic bilevel optimization problem*

$$(PB) \begin{cases} \min_{x \in X} \sup_{y \in M(x)} L(x, y) \\ \text{where } M(x) = \text{Arg min}_{y \in Y} F(x, y). \end{cases}$$

The introduction of such a problem is due to Leitmann in [46], where it is denoted as *generalized Stackelberg problem*. The solution concepts associated to (PB) are reminded below.

Definition 1.2

- The infimum value $w = \inf_{x \in X} \sup_{y \in M(x)} L(x, y)$ is the *security (or pessimistic) value* of (PB).
- A leader's action $\bar{x} \in X$ is a *weak Stackelberg (or pessimistic) solution* to (PB) if

$$\bar{x} \in \text{Arg min}_{x \in X} \sup_{y \in M(x)} L(x, y),$$

that is, $\sup_{y \in M(\bar{x})} L(\bar{x}, y) = \inf_{x \in X} \sup_{y \in M(x)} L(x, y)$.

- An action profile $(\bar{x}, \bar{y}) \in X \times Y$ is a *weak Stackelberg (or pessimistic) equilibrium* of (PB) if \bar{x} is a weak Stackelberg solution and $\bar{y} \in M(\bar{x})$.

For first investigations about regularization and approximation of problem (PB) see [89, 69, 75, 61].

C. Optimistic leader: strong Stackelberg problem

In a second “extreme” situation, the leader is optimistic and believes that the follower will choose the best action for the leader in the set of his optimal responses (or the leader is able to force the follower to choose, among all the optimal responses, the best one for the leader). The leader will deal with the *strong Stackelberg problem*, also called *optimistic bilevel optimization problem*

$$(OB) \begin{cases} \min_{x \in X} \inf_{y \in M(x)} L(x, y) \\ \text{where } M(x) = \text{Arg min}_{y \in Y} F(x, y). \end{cases}$$

The solution concepts associated to (OB) are reminded below.

Definition 1.3

- The infimum value $s = \inf_{x \in X} \inf_{y \in M(x)} L(x, y)$ is the *optimistic value* of (OB).

- A leader's action $\bar{x} \in X$ is a *strong Stackelberg (or optimistic) solution* to (OB) if

$$\bar{x} \in \text{Arg min}_{x \in X} \inf_{y \in M(x)} L(x, y),$$

that is, $\inf_{y \in M(\bar{x})} L(\bar{x}, y) = \inf_{x \in X} \inf_{y \in M(x)} L(x, y)$.

- An action profile $(\bar{x}, \bar{y}) \in X \times Y$ is a *strong Stackelberg (or optimistic) equilibrium* of (OB) if \bar{x} is a strong Stackelberg solution and $\bar{y} \in \text{Arg min}_{y \in M(\bar{x})} L(\bar{x}, y)$.

The use of terms “weak Stackelberg” and “strong Stackelberg” goes back to [20] where a concept of sequential Stackelberg equilibrium in a general dynamic games framework is introduced. For a first investigation on regularization and approximation of problem (OB) see [60].

We point out that an action profile is a strong Stackelberg equilibrium of (OB) if and only if it is a solution of the following problem:

$$\min_{x \in X, y \in M(x)} L(x, y)$$

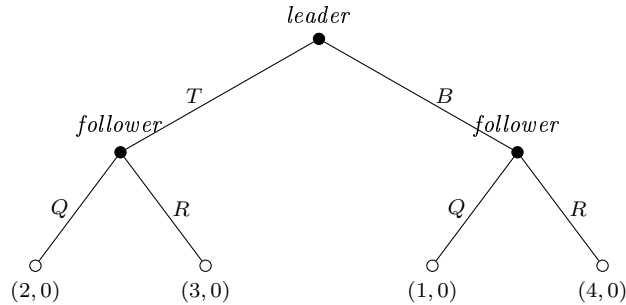
(and, moreover, its value coincides with the optimistic value of (OB)), also frequently called “bilevel optimization problem” (for first results, see for example [15, 27, 96, 85, 108, 112]).

The following examples show that the pessimistic and the optimistic behaviours of the leader can design different solutions, values and equilibria.

Example 1.2 Consider the Stackelberg game in [93, Example 2.1] where the leader has two available actions T and B , that is $X = \{T, B\}$, the follower has two available actions Q and R , that is $Y = \{Q, R\}$, and the payoff functions are defined by

$$\begin{aligned} L(T, Q) &= 2, & L(T, R) &= 3, & L(B, Q) &= 1, & L(B, R) &= 4, \\ F(T, Q) &= 0, & F(T, R) &= 0, & F(B, Q) &= 0, & F(B, R) &= 0. \end{aligned}$$

Such a game can be also described by using the following “extensive form representation”:



The follower's best reply correspondence is defined by $M(T) = \{Q, R\}$ and $M(B) =$

$\{Q, R\}$. Since

$$\begin{aligned} x = T &\implies \sup_{y \in M(x)} L(x, y) = 3 \quad \text{and} \quad \inf_{y \in M(x)} L(x, y) = 2, \\ x = B &\implies \sup_{y \in M(x)} L(x, y) = 4 \quad \text{and} \quad \inf_{y \in M(x)} L(x, y) = 1, \end{aligned}$$

then $\inf_{x \in \{T, B\}} \sup_{y \in M(x)} L(x, y) = 3$ and $\inf_{x \in \{T, B\}} \inf_{y \in M(x)} L(x, y) = 1$. Therefore, the weak Stackelberg solution is T , the pessimistic value is 3 and the weak Stackelberg equilibria are (T, Q) and (T, R) . Whereas, the strong Stackelberg solution is B , the optimistic value is 1 and the strong Stackelberg equilibrium is (B, Q) .

Example 1.3 Consider the Stackelberg game in [93, Example 3.4] where $X = [-2, 2]$, $Y = [-1, 1]$,

$$L(x, y) = -x - y \quad \text{and} \quad F(x, y) = \begin{cases} (x + 7/4)y, & \text{if } x \in [-2, -7/4[\\ 0, & \text{if } x \in [-7/4, 7/4[\\ (x - 7/4)y, & \text{if } x \in [7/4, 2]. \end{cases}$$

The follower's best reply correspondence M is defined on $[-2, 2]$ by

$$M(x) = \begin{cases} \{1\}, & \text{if } x \in [-2, -7/4[\\ [-1, 1], & \text{if } x \in [-7/4, 7/4[\\ \{-1\}, & \text{if } x \in]7/4, 2]. \end{cases}$$

On the one hand, since

$$\sup_{y \in M(x)} L(x, y) = \begin{cases} -x - 1, & \text{if } x \in [-2, -7/4[\\ -x + 1, & \text{if } x \in [-7/4, 2], \end{cases}$$

then $\inf_{x \in X} \sup_{y \in M(x)} L(x, y) = -1$. Hence, the weak Stackelberg solution is 2, the pessimistic value is -1 and the weak Stackelberg equilibrium is $(2, -1)$.

On the other hand, as

$$\inf_{y \in M(x)} L(x, y) = \begin{cases} -x - 1, & \text{if } x \in [-2, 7/4[\\ -x + 1, & \text{if } x \in]7/4, 2], \end{cases}$$

then $\inf_{x \in X} \inf_{y \in M(x)} L(x, y) = -11/4$. Therefore, the strong Stackelberg solution is $7/4$, the optimistic value is $-11/4$ and the strong Stackelberg equilibrium is $(7/4, 1)$.

D. Intermediate Stackelberg problem

The problems (PB) and (OB) reflect the two possible extreme behaviours of the leader regarding how the follower chooses his action in the second stage. However, an intermediate situation could occur: the leader has some information on the follower's choice in his set of optimal responses which allows to attribute a probability distribution reflecting

his beliefs on $M(x)$ for any $x \in X$. Assume here, in order to simplify, that $M(x)$ has a finite number of elements for any $x \in X$, namely

$$M(x) = \{y_1(x), \dots, y_{k(x)}(x)\}.$$

Therefore, denoted with $p_i(x)$ the probability (enforced by the leader) that the follower chooses $y_i(x)$ in $M(x)$ for any $x \in X$ and defined $D = \{D(x) | x \in X\}$ where $D(x) = \{p_1(x), \dots, p_{k(x)}(x)\}$, the leader will face the so-called *intermediate Stackelberg problem* with respect to D

$$(IS_D) \quad \min_{x \in X} \sum_{i=1}^{k(x)} p_i(x) L(x, y_i(x)).$$

Problem (IS_D) and the related solution concepts were introduced by Mallozzi and Morgan in [84, 82] also in the general case where $M(x)$ is not a discrete set for any $x \in X$ (see also [83] for an application to oligopolistic markets). Note that the *strong-weak Stackelberg problem* proposed in [3] includes particular intermediate problems where the leader attributes probability 0 to any point in $M(x)$ except for the best one and the worst one for him. Afterwards, an approach similar to [82] has been examined in [21] where a “partial cooperative” attitude of the follower is modelled in Stackelberg games with linear payoff functions.

Let us recall the solution concepts related to (IS_D) .

Definition 1.4

- The infimum value $v_D = \inf_{x \in X} \sum_{i=1}^{k(x)} p_i(x) L(x, y_i(x))$ is the *intermediate value* of (IS_D) .
- A leader’s action $\bar{x} \in X$ is an *intermediate Stackelberg solution* with respect to D if

$$\bar{x} \in \text{Arg min}_{x \in X} \sum_{i=1}^{k(x)} p_i(x) L(x, y_i(x)).$$

- An action profile $(\bar{x}, \bar{y}) \in X \times Y$ is an *intermediate Stackelberg equilibrium* with respect to D if \bar{x} is an intermediate Stackelberg solution with respect to D and $\bar{y} \in M(\bar{x})$.

Depending on the probability distribution $D(x)$ on $M(x)$ for any $x \in X$, problem (IS_D) could become (PB) or (OB) . Examples on the connections among weak, strong and intermediate Stackelberg solutions are illustrated in [82], one of which is presented below for the sake of completeness.

Example 1.4 Consider $X = Y = [-1, 1]$, $L(x, y) = x + y$ and $F(x, y) = |y^2 - x^4|$. The follower’s best reply correspondence M is defined on $[-1, 1]$ by $M(x) = \{x^2, -x^2\}$ and let $D(x) = \{\alpha, 1 - \alpha\}$ with $\alpha \in [0, 1]$ be the probability distribution attributed by the

leader on $M(x)$, for any $x \in X$. Hence, the intermediate Stackelberg solution is

$$\operatorname{Arg\,min}_{x \in X} (\alpha L(x, x^2) + (1 - \alpha)L(x, -x^2)) = \begin{cases} \{-1\}, & \text{if } \alpha \in [0, 3/4] \\ \{-1/2(2\alpha - 1)\}, & \text{if } \alpha \in]3/4, 1]. \end{cases}$$

We point out that the solution to (IS_D) coincides with the solution to (PB) if $\alpha = 1$ and with the solution to (OB) if $\alpha \leq 3/4$. Furthermore, regarding the values, we have

$$v_D = \inf_{x \in X} (\alpha L(x, x^2) + (1 - \alpha)L(x, -x^2)) = \begin{cases} 2\alpha - 2, & \text{if } \alpha \in [0, 3/4] \\ -1/4(2\alpha - 1), & \text{if } \alpha \in]3/4, 1], \end{cases}$$

and, reminded that we denoted with w and s , respectively, the pessimistic value and the optimistic value, it follows

$$-2 = s \leq v_D \leq w = -1/4$$

for any $\alpha \in [0, 1]$. Inequality above relies on a general property: the intermediate value is always between the optimistic value and the pessimistic value.

E. Subgame perfect Nash equilibrium problem

The equilibrium concepts presented until now consist of action profiles, that is pairs composed by one action of the leader and one action of the follower and, moreover, they are focused primarily on the leader's perspectives. Such concepts are broadly used in an engineering setting or by optimization practitioners, see for example [10, 81, 28, 41]. However, as traditional in a game-theoretical framework, we are now interested in an equilibrium concept which takes more into account the strategic aspects of the game. The equilibrium concept naturally fitting such goal is the *subgame perfect Nash equilibrium* (henceforth SPNE), which represents the most widely employed solution concept in an economic setting. The SPNE solution concept was introduced by the Nobel laureate Selten in [101] who suggested that players should act according to the so-called *principle of sequential rationality*: "the equilibria to select are those whereby the players behave optimally from any point of the game onwards". For the notions of strategy, subgame and SPNE in a general game-theoretical framework and for further discussion on the behavioural implications and procedures to find SPNEs see, for example, [39, Chapter 3] and [87, Chapter 7].

We first recall the notion of *player's strategy* only for a two-player Stackelberg game $\Gamma = (X, Y, L, F)$. In such a game, the set of leader's strategies coincides with the set of leader's actions X , whereas the set of follower's strategies is the set of all the functions from X to Y , i.e. $Y^X = \{\varphi: X \rightarrow Y\}$.

The set of strategy profiles is $X \times Y^X$ and the definition of subgame perfect Nash equilibrium is characterized in the following way.

Definition 1.5 A strategy profile $(\bar{x}, \bar{\varphi}) \in X \times Y^X$ is an SPNE of Γ if the following conditions are satisfied:

(SG1) for each choice x of the leader, the follower minimizes his payoff function, i.e.

$$\bar{\varphi}(x) \in M(x) \quad \text{for any } x \in X;$$

(SG2) the leader minimizes his payoff function taking into account his hierarchical advantage, i.e.

$$\bar{x} \in \underset{x \in X}{\text{Arg min}} L(x, \bar{\varphi}(x)).$$

Note that the denomination “subgame perfect Nash equilibrium” is due to the following key features: first the SPNE notion is a refinement of the Nash equilibrium solution concept (introduced by the Nobel laureate Nash in [95]), secondly the restriction of an SPNE to any subgame constitutes a Nash equilibrium.

We emphasize that an SPNE consists of a strategy profile, that is a pair composed by one strategy of the leader (or equivalently one action of the leader) and one strategy of the follower (which is a function from the actions set of the leader to the actions set of the follower), as illustrated in Definition 1.5 and differently from all the equilibrium concepts defined before where only action profiles are involved. However, we can connect SPNEs and the equilibrium notions described above in this section, as pointed out in the following three examples.

Example 1.5 The set of SPNEs of a Stackelberg game $\Gamma = (X, Y, L, F)$ where the follower’s optimal response to any choice of the leader is unique can be fully characterized in terms of the solutions to the associate problem (*SP*) defined in paragraph A.

In fact, just by exploiting the definitions 1.1 and 1.5, the following equivalence holds:

$$(\bar{x}, \bar{\varphi}) \in X \times Y^X \text{ is an SPNE of } \Gamma \iff \begin{array}{l} \bar{x} \text{ is a solution to (SP) and} \\ M(x) = \{\bar{\varphi}(x)\} \text{ for any } x \in X, \end{array}$$

and, moreover

$$(\bar{x}, \bar{\varphi}) \in X \times Y^X \text{ is an SPNE of } \Gamma \implies (\bar{x}, \bar{\varphi}(\bar{x})) \in X \times Y \text{ is a Stackelberg equilibrium.}$$

For instance, in the Stackelberg leadership model presented in Example 1.1, the unique SPNE is the strategy profile (\bar{q}_1, \bar{m}) where $\bar{q}_1 \in [0, +\infty[$ and $\bar{m}: [0, +\infty[\rightarrow [0, +\infty[$ are given by

$$\bar{q}_1 = \frac{a-c}{2b} \quad \text{and} \quad \bar{m}(q_1) = \max \left\{ 0, \frac{a-bq_1-c}{2b} \right\}.$$

We remind that \bar{q}_1 is the unique Stackelberg solution and, moreover, $(\bar{q}_1, \bar{m}(\bar{q}_1))$ is the unique Stackelberg equilibrium.

Example 1.6 In the case where the follower’s best reply correspondence is not single-valued, the solutions to the associate problem (*PB*) as well as to the associate problem (*OB*) defined in paragraphes B and C, respectively, can be part of an SPNE.

For instance, in the game of Example 1.2 (similarly to what is pointed out in [93,

Example 2.1]) there are two SPNEs whose first component is a weak Stackelberg solution, namely $(T, \bar{\varphi}_1)$ and $(T, \bar{\varphi}_2)$ where $\bar{\varphi}_1$ and $\bar{\varphi}_2$ are the functions defined by

$$\bar{\varphi}_1(x) = \begin{cases} R, & \text{if } x = T \\ R, & \text{if } x = B \end{cases} \quad \text{and} \quad \bar{\varphi}_2(x) = \begin{cases} Q, & \text{if } x = T \\ R, & \text{if } x = B, \end{cases}$$

respectively; there are two SPNEs whose first component is a strong Stackelberg solution, namely $(B, \bar{\varphi}_3)$ and $(B, \bar{\varphi}_4)$ where $\bar{\varphi}_3$ and $\bar{\varphi}_4$ are the functions defined by

$$\bar{\varphi}_3(x) = \begin{cases} Q, & \text{if } x = T \\ Q, & \text{if } x = B \end{cases} \quad \text{and} \quad \bar{\varphi}_4(x) = \begin{cases} R, & \text{if } x = T \\ Q, & \text{if } x = B, \end{cases}$$

respectively. Focusing on the SPNEs $(T, \bar{\varphi}_1)$ and $(B, \bar{\varphi}_3)$, since

$$\begin{aligned} \{\bar{\varphi}_1(T)\} &= \text{Arg max}_{y \in \{Q, R\}} L(T, y) & \text{and} & \quad \{\bar{\varphi}_1(B)\} = \text{Arg max}_{y \in \{Q, R\}} L(B, y), \\ \{\bar{\varphi}_3(T)\} &= \text{Arg min}_{y \in \{Q, R\}} L(T, y) & \text{and} & \quad \{\bar{\varphi}_3(B)\} = \text{Arg min}_{y \in \{Q, R\}} L(B, y), \end{aligned}$$

such SPNEs naturally reflect the pessimistic and the optimistic behaviours of the leader. Therefore, having also in mind to deal with the issue of how reducing the number of SPNEs, we will take into account these latter types of SPNEs “induced” by weak or strong Stackelberg solutions.

Definition 1.6

- A strategy profile $(\bar{x}, \bar{\varphi}) \in X \times Y^X$ is an SPNE of Γ *induced by a weak Stackelberg solution* if $(\bar{x}, \bar{\varphi})$ is an SPNE of Γ which satisfies

$$\bar{x} \text{ is a solution to problem } (PB) \quad \text{and} \quad \bar{\varphi}(x) \in \text{Arg max}_{y \in M(x)} L(x, y) \text{ for any } x \in X.$$

- A strategy profile $(\bar{x}, \bar{\varphi}) \in X \times Y^X$ is an SPNE of Γ *induced by a strong Stackelberg solution* if $(\bar{x}, \bar{\varphi})$ is an SPNE of Γ which satisfies

$$\bar{x} \text{ is a solution to problem } (OB) \quad \text{and} \quad \bar{\varphi}(x) \in \text{Arg min}_{y \in M(x)} L(x, y) \text{ for any } x \in X.$$

Coming back to Example 1.6, $(T, \bar{\varphi}_1)$ is an SPNE induced by a weak Stackelberg solution, $(B, \bar{\varphi}_3)$ is an SPNE induced by a strong Stackelberg solution, $(T, \bar{\varphi}_2)$ and $(B, \bar{\varphi}_4)$ are SPNEs that are not induced either by a weak or by a strong Stackelberg solution and the game has no further SPNEs.

Let us provide below an example where the SPNEs induced by weak or strong Stackelberg solutions are derived in a continuous setting, differently from Example 1.6.

Example 1.7 In the game of Example 1.3, there is one weak Stackelberg solution to (PB) and one strong Stackelberg solution to (OB) and each one induces one SPNE: the

strategy profiles $(2, \bar{\varphi})$ and $(7/4, \bar{\psi})$ where $\bar{\varphi}: [-2, 2] \rightarrow [-1, 1]$ and $\bar{\psi}: [-2, 2] \rightarrow [-1, 1]$ are defined respectively by

$$\bar{\varphi}(x) = \begin{cases} 1, & \text{if } x \in [-2, -7/4[\\ -1, & \text{if } x \in [-7/4, 2] \end{cases} \quad \text{and} \quad \bar{\psi}(x) = \begin{cases} 1, & \text{if } x \in [-2, 7/4[\\ -1, & \text{if } x \in]7/4, 2], \end{cases}$$

are the SPNEs induced by the weak Stackelberg solution and the strong Stackelberg solution, respectively. As in Example 1.6, also in this case there are SPNEs not induced either by weak or by strong Stackelberg solutions, like the strategy profile $(2, \bar{\vartheta})$ where $\bar{\vartheta}: [-2, 2] \rightarrow [-1, 1]$ is defined by

$$\bar{\vartheta}(x) = \begin{cases} 1, & \text{if } x \in [-2, -7/4[\\ -4x/7, & \text{if } x \in [-7/4, 7/4[\\ -1, & \text{if } x \in [7/4, 2]. \end{cases}$$

After having presented the mathematical problems associated to a Stackelberg game interaction, we conclude this section by describing the main issues arising when one deals with the different kinds of problems illustrated above. In particular, we will focus on the following topics:

- **Existence:** do the solutions of the problem exist under not too restrictive assumptions on the actions sets and on the payoff functions (possibly discontinuous, as not unusual in economic frameworks)?
- **“Variational” stability:** the data of the problem could be affected by uncertainty, hence given a perturbation of the original problem acting on the actions sets and/or on the payoff functions, does a sequence of solutions of the perturbed problems converge to a solution of the original problem? And what about the values? Such analysis, which helps to predict the “behaviour” of the game when one is not able to deal with exact equilibria, and similar analyses that compare outcomes before and after a change in exogenous parameters are known in economics as *comparative statics*.
- **Well-posedness:** does the problem have a unique solution? And does any method which constructs an “approximating sequence” automatically allows to approach a solution?
- **Approximation:** how constructing appropriate concepts of approximate solutions in order to both obviate the lack of solutions of a problem and to manage problems where infinite dimensional spaces are involved?
- **Numerical approximation:** how transforming the original problem into a better-behaved equivalent problem? And how overcoming the numerical difficulties which derive from the possible non-single-valuedness of the follower’s best reply correspondence?

- **Selection:** if the problem has more than one solution, is it possible to construct a procedure which allows to reduce the number of solutions or, better yet, to pick just one solution? Which are the motivations that would induce the players to act according to such a procedure in order to reach the designed solution?

Obviously the same issues appear also for the equilibria.

In the next section, for any kind of problem previously defined, we will discuss which of the issues just displayed can be positively or negatively answered and we will specify in detail the outline of the rest of the chapter.

2 Crucial issues in Stackelberg games

Let us remind below the list of crucial issues:

- Existence;
- Variational stability;
- Well-posedness;
- Approximation;
- Numerical approximation;
- Selection.

A. Stackelberg problems

Problem (SP) is well-behaved regarding all the topics just summarized. In particular, by applying the maximum theorem in [16, 6], the existence of Stackelberg solutions and Stackelberg equilibria is guaranteed provided that the action sets X and Y are compact subsets of two Euclidean spaces and that the payoff functions L and F are continuous over $X \times Y$.

Remark 2.1 However, we point out that the functions involved in many real world situations are not always continuous and in infinite dimensional frameworks requiring the continuity is very restrictive. So, as proved in propositions 4.1 and 5.1 in [92], one can only assume that

- L is lower semicontinuous over $X \times Y$,
- F is lower semicontinuous over $X \times Y$,
- for any $(x, y) \in X \times Y$ and any sequence $(x_n)_n$ converging to x in X , there exists a sequence $(\tilde{y}_n)_n$ in Y such that

$$\limsup_{n \rightarrow +\infty} F(x_n, \tilde{y}_n) \leq F(x, y).$$

Furthermore, it is worth to highlight that the second and the third condition stated above does not imply the continuity of F (see, for example, [58, Example 3.1.1]).

As regards variational stability of the Stackelberg solution, equilibrium and value of (SP) under perturbations of the payoff functions, consider the perturbed problems

$$(SP)_n \begin{cases} \min_{x \in X} L_n(x, m_n(x)) \\ \text{where } m_n(x) \text{ is the solution to } \min_{y \in Y} F_n(x, y), \end{cases}$$

with L_n and F_n real-valued functions defined on $X \times Y$, for any $n \in N$. Then, the following results hold.

Proposition 2.1. [68, Prop. 3.1, 3.2 and 3.3] *Assume that X and Y are compact subsets of two Euclidean spaces and that the sequences $(L_n)_n$ and $(F_n)_n$ continuously converge to L and F , respectively, i.e. for any $(x, y) \in X \times Y$ and any sequence $(x_n, y_n)_n$ converging to (x, y) in $X \times Y$, we have*

$$\lim_{n \rightarrow +\infty} L_n(x_n, y_n) = L(x, y) \quad \text{and} \quad \lim_{n \rightarrow +\infty} F_n(x_n, y_n) = F(x, y).$$

Then, about the second-stage problem, for any sequence $(x_n)_n$ converging to x in X ,

$$\lim_{n \rightarrow +\infty} m_n(x_n) = m(x) \quad \text{and} \quad \lim_{n \rightarrow +\infty} F_n(x_n, m_n(x_n)) = F(x, m(x)).$$

Moreover, regarding the first-stage problem,

- for any sequence $(\bar{x}_n)_n$ such that $\bar{x}_{n_k} \in \text{Arg min}_{x \in X} L_{n_k}(x, m_{n_k}(x))$ and $(\bar{x}_{n_k})_k$ converges to \bar{x} in X for a selection of integers $(n_k)_k$, we have

$$\bar{x} \in \text{Arg min}_{x \in X} L(x, m(x)),$$

- $\limsup_{n \rightarrow +\infty} \left(\inf_{x \in X} L_n(x, m_n(x)) \right) \leq \inf_{x \in X} L(x, m(x))$.

Finally, if $(\bar{x}_n, \bar{y}_n) \in X \times Y$ is a Stackelberg equilibrium of $(SP)_n$ for any $n \in \mathbb{N}$, then any convergent subsequence of the sequence $(\bar{x}_n, \bar{y}_n)_n$ in $X \times Y$ has a limit (\bar{x}, \bar{y}) which is a Stackelberg equilibrium of (SP) and satisfies

$$\lim_{n \rightarrow +\infty} L_n(\bar{x}_n, \bar{y}_n) = L(\bar{x}, \bar{y}) \quad \text{and} \quad \lim_{n \rightarrow +\infty} F_n(\bar{x}_n, \bar{y}_n) = F(\bar{x}, \bar{y}).$$

Remark 2.2 Stability results illustrated in Proposition 2.1 hold even by replacing the convergence assumptions with weaker convergence requirements, as proved in [68]. In particular

- the continuous convergence of $(L_n)_n$ to L can be substituted by the following two conditions:
 - (a) for any $(x, y) \in X \times Y$ and any sequence $(x_n, y_n)_n$ converging to (x, y) in $X \times Y$, we have

$$\liminf_{n \rightarrow +\infty} L_n(x_n, y_n) \geq L(x, y);$$

- (b) for any $x \in X$ there exists a sequence $(\tilde{x}_n)_n$ converging to x in X such that, for any $y \in Y$ and any sequence $(y_n)_n$ converging to y in Y , we have

$$\limsup_{n \rightarrow +\infty} L_n(\tilde{x}_n, y_n) \leq L(x, y);$$

- whereas, the continuous convergence of $(F_n)_n$ to F can be substituted by the following two conditions:

- (c) for any $(x, y) \in X \times Y$ and any sequence $(x_n, y_n)_n$ converging to (x, y) in $X \times Y$, we have

$$\liminf_{n \rightarrow +\infty} F_n(x_n, y_n) \geq F(x, y);$$

- (d) for any $(x, y) \in X \times Y$ and any sequence $(x_n)_n$ converging to x in X , there exists a sequence $(\tilde{y}_n)_n$ in Y such that

$$\limsup_{n \rightarrow +\infty} F_n(x_n, \tilde{y}_n) \leq F(x, y).$$

Examples and counterexamples on how conditions (a)–(d) are connected with known convergence notions can be found in remarks 2.2 and 2.3 in [68] where perturbations on the constraints of the leader and the follower are also considered.

Results on well-posedness of (SP) can be found in [70, 92], whereas in [28, Chapter 6] numerical approximation and algorithmic issues are widely investigated.

Concerning problem (GSP) defined in paragraph A of Section 1, we mention that in [103] a first approximation technique based on a barrier method has been proposed, then in [71] a general approximation scheme involving conditions of minimal character has been introduced together with applications to barrier methods, whereas in [67] also external penalty methods have been considered.

B. Weak Stackelberg (or pessimistic bilevel optimization) problem

Problem (PB) is the worst-behaved among the problems illustrated in Section 1. In fact, the compactness of the action sets and the continuity of the payoff functions do not guarantee, in general, the existence of weak Stackelberg solutions and weak Stackelberg equilibria even in a finite dimensional setting, as proved in many folk examples in literature (see [14, 80] and [10, Remark 4.3]) and in the following one.

Example 2.1 Consider $X = Y = [-1, 1]$, $L(x, y) = -x + y$ and $F(x, y) = -xy$. The follower's best reply correspondence M is defined on $[-1, 1]$ by

$$M(x) = \begin{cases} \{-1\}, & \text{if } x \in [-1, 0[\\ [-1, 1], & \text{if } x = 0 \\ \{1\}, & \text{if } x \in]0, 1]. \end{cases} \quad (1)$$

Since

$$\sup_{y \in M(x)} L(x, y) = \begin{cases} -x - 1, & \text{if } x \in [-1, 0[\\ -x + 1, & \text{if } x \in [0, 1], \end{cases}$$

then $\text{Arg min}_{x \in X} \sup_{y \in M(x)} L(x, y) = \emptyset$. Hence such a (PB) has no solution and no weak Stackelberg equilibrium.

Nevertheless, existence results for very special classes of problem (PB) can be found in [80, 2, 4].

In a more general framework, to overcome the lack of solutions, for any $\epsilon > 0$ the concept of ϵ -approximate solution of (PB) has been investigated in a sequential setting by Loridan and Morgan in [75], where the existence of such solutions is analyzed under mild convexity assumptions on the follower's payoff function, and in [77] where quasi-convexity assumptions are considered. Then, in [76] the notion of strict ϵ -approximate solution has been presented and related existence results have been provided without requiring convexity assumptions. A crucial property of these two concepts is the convergence of the approximate pessimistic values towards the pessimistic value of (PB) as ϵ tends to zero. We point out that all the results in the papers just mentioned have been obtained in the case of non-parametric constraints.

Afterwards, Lignola and Morgan in [61] extensively investigated all the possible kind of constraints (including those defined by parametric inequalities) in a topological setting. More recently, new notions of approximate solutions have been introduced in [57, 56] in an easier-to-manage sequential setting which allow to construct a surrogate for the solution of (PB) called *viscosity solution*. Such concepts will be presented in Section 3 together with their properties.

Regarding the variational stability issue of (PB) , consider the following perturbed problems

$$(PB)_n \begin{cases} \min_{x \in X} \sup_{y \in M_n(x)} L_n(x, y) \\ \text{where } M_n(x) = \text{Arg min}_{y \in Y} F_n(x, y), \end{cases}$$

with L_n and F_n real-valued functions defined on $X \times Y$, for any $n \in \mathbb{N}$. Even requiring compactness and heavy convergence assumptions, variational stability of the weak Stackelberg solutions and the pessimistic values under perturbations of the payoff functions does not hold in general, as described in the example below.

Example 2.2 Consider $X = Y = [-1, 1]$, $L(x, y) = x + y$, $F(x, y) = 0$ and let $L_n(x, y) = x + y + 1/n$ and $F_n(x, y) = y/n$ for any $n \in \mathbb{N}$.

Although the sequence of functions $(L_n)_n$ continuously converges to L and the sequence $(F_n)_n$ continuously converges to F , the sequence of pessimistic values of $(PB)_n$ does not converge to the pessimistic value of (PB) .

In fact, since the follower's best reply correspondences in $(PB)_n$ and in (PB) are defined

respectively by $M_n(x) = \{-1\}$ and by $M(x) = [-1, 1]$ for any $x \in X$, then

$$\inf_{x \in X} \sup_{y \in M_n(x)} L_n(x, y) = -2 + 1/n \quad \text{and} \quad \inf_{x \in X} \sup_{y \in M(x)} L(x, y) = 0.$$

Hence

$$-2 = \lim_{n \rightarrow +\infty} \left(\inf_{x \in X} \sup_{y \in M_n(x)} L_n(x, y) \right) \neq \inf_{x \in X} \sup_{y \in M(x)} L(x, y) = 0.$$

However, convergence results of the exact solutions to $(PB)_n$ towards a solution to (PB) have been obtained in [69] for a special class of problems.

In a more general framework, in order to obviate the lack of stability, a first attempt to manage the stability issue (intended in the Hausdorff's sense) of (PB) can be found in [89] by means of approximate solutions. Afterwards, "variational" stability (in the sense of Example 2.2) for such approximate solutions has been investigated under conditions of minimal character and assuming non-parametric constraints in [75] (under mild convexity assumptions on the follower's payoff function) and in [77] (under quasi-convexity assumptions). For applications to interior penalty methods see [73, Section 5] and to exterior penalty methods see [77, Section 5].

Then, a comprehensive analysis on all the types of follower's constraints has been provided in [61] for a general topological framework. Such results will be presented in Section 3.

Furthermore, (PB) is hard to manage even regarding well-posedness properties, investigated in [92].

Concerning the numerical approximation issue of (PB) , various regularization methods involving both the follower's optimal reaction set and the follower's payoff function have been proposed in order to approach problem (PB) via a sequence of Stackelberg problems; for first results, see [88, 78, 74, 32, 79]. These methods, their properties and the results achieved will be presented in Section 4.

C. Strong Stackelberg (or optimistic bilevel optimization) problem

Problem (OB) is ensured to have strong Stackelberg solutions under the same assumptions stated for the existence of solutions of (SP) , i.e. by assuming that the action sets X and Y are compact subsets of two Euclidean spaces and that the payoff functions L and F are continuous over $X \times Y$.

Remark 2.3 The continuity assumptions can be weakened as in Remark 2.1, by applying Proposition 4.1.1 in [58] about the lower semicontinuity of the marginal function.

We remind that results concerning the existence of strong Stackelberg solutions have been first obtained in [55].

As regards to the variational stability issue of (OB) , consider the following perturbed

problems

$$(OB)_n \begin{cases} \min_{x \in X} \inf_{y \in M_n(x)} L_n(x, y) \\ \text{where } M_n(x) = \operatorname{Arg} \min_{y \in Y} F_n(x, y), \end{cases}$$

with L_n and F_n real-valued functions defined on $X \times Y$, for any $n \in \mathbb{N}$.

We point out that problem (OB) , as well as (PB) , can exhibit under perturbation, a lack of convergence of the strong Stackelberg solutions and of the optimistic values, as illustrated in Example 4.1 in [60], rewritten below for the sake of completeness.

Example 2.3 Consider $X = Y = [0, 1]$, $L(x, y) = x - y + 1$, $F(x, y) = x$ and let $L_n(x, y) = L(x, y)$ and $F_n(x, y) = y/n + x$ for any $n \in \mathbb{N}$.

Though the sequences of functions $(L_n)_n$ and $(F_n)_n$ continuously converge to L and F , respectively, the sequence of optimistic values of $(OB)_n$ does not converge to the optimistic value of (OB) .

Indeed, as the follower's best reply correspondences in $(OB)_n$ and in (OB) are defined respectively by $M_n(x) = \{0\}$ and by $M(x) = [0, 1]$ for any $x \in X$, then

$$\inf_{x \in X} \inf_{y \in M_n(x)} L_n(x, y) = 1 \quad \text{and} \quad \inf_{x \in X} \inf_{y \in M(x)} L(x, y) = 0.$$

Therefore,

$$\lim_{n \rightarrow +\infty} \left(\inf_{x \in X} \inf_{y \in M_n(x)} L_n(x, y) \right) \neq \inf_{x \in X} \inf_{y \in M(x)} L(x, y).$$

However, notions of approximate solutions have been introduced and investigated in [60] in order to face the lack of stability in (OB) , whereas achievements on well-posedness can be derived by applying results obtained in [64, 50].

Obviously, as well as in (PB) , the non-single-valuedness of the follower's best reply correspondence M gives rise to difficulties from the numerical point of view in (OB) .

D. Intermediate Stackelberg problem

Sufficient conditions for the existence of intermediate Stackelberg solutions of (IS_D) are investigated in [82, Section 3] in the general case where $M(x)$ is not a discrete set for at least one $x \in X$. However, we point out that problem (IS_D) inherits all the difficulties illustrated for (PB) about existence, stability, well-posedness and approximation issues.

E. Subgame perfect Nash equilibrium problem

When the follower's best reply correspondence M is not a single-valued map, infinitely many SPNEs could come up in Stackelberg games, as shown in the following example.

Example 2.4 Consider $X = Y = [-1, 1]$, $L(x, y) = x + y$ and $F(x, y) = -xy$. The follower's best reply correspondence M is given in (1). Denoted with φ^σ the function

defined on $[-1, 1]$ by

$$\varphi^\sigma(x) = \begin{cases} -1, & \text{if } x \in [-1, 0[\\ \sigma, & \text{if } x = 0 \\ 1, & \text{if } x \in]0, 1], \end{cases}$$

where $\sigma \in [-1, 1]$, it follows that $\varphi^\sigma(x) \in M(x)$ for any $x \in [-1, 1]$ and $\text{Argmin}_{x \in X} L(x, \varphi^\sigma(x)) = \{-1\}$. Therefore, the strategy profile $(-1, \varphi^\sigma)$ is an SPNE of Γ for any $\sigma \in [-1, 1]$, so Γ has infinitely many SPNEs.

Hence, restricting the number of SPNEs becomes essential and the selection issue arises. For further discussion on the theory of equilibrium selection in games, see [40].

Remark 2.4 We showed via Examples 1.6 and 1.7 that starting from a solution to the associated (*PB*) or to the associated (*OB*) one can induce an SPNE motivated according to each of the two different behaviours of the leader. Finding such SPNEs induced by weak or strong Stackelberg solutions, as defined in Definition 1.6, provides two methods to select an SPNE in Stackelberg games. Nevertheless such methods can be exploited only in the two extreme situations illustrated before. Anyway they require to the leader of knowing the follower's best reply correspondence and, at the best of our knowledge, they do not exhibit a manageable constructive approach to achieve an SPNE.

Furthermore, also the numerical approximation issue matters for the SPNEs in Stackelberg games since M can be not single-valued. Therefore, it would be desirable to design constructive methods in order to select an SPNE with the following features: relieving the leader of knowing M , allowing to overcome the difficulties deriving from the possible non-single-valuedness of M , providing some behavioural motivations of the players (different from the extreme situations described above). Results in this direction have been first obtained in [93], where an SPNE selection method based on Tikhonov regularization is proposed, and more recently in [22], where proximal regularization is employed.

In order to overcome the drawbacks illustrated in this section, we identify two basic and widely used approaches that lead to two classes of regularization methods.

- Regularizing the follower's optimal reaction sets;
- Regularizing the follower's payoff function.

The first of the above approaches will be presented in Section 3 for problems (*PB*) and (*OB*), in the case where the follower's constraints involve a finite number of inequalities, and for other kinds of bilevel problems, such as two-stage games where the second-stage problem is defined by a variational or a quasi-variational inequality, by a Nash equilibrium problem and, more generally, by a quasi-equilibrium problem.

In Section 4 we will investigate the second approach which allows to approximate a

Stackelberg game via a sequence of Stackelberg games whose follower's best reply correspondence is single-valued. A conclusive section concerns various extensions of the static Stackelberg game above considered.

In the rest of the chapter we will again present results under easy-to-manage continuity or convergence assumptions, but we point out that the papers quoted in the statements involve conditions of minimal character, as in remarks 2.1 and 2.2.

3 Regularizing the follower's optimal reaction set in Stackelberg games

The first attempts for approximating a Stackelberg game $\Gamma = (X, Y, L, F)$ by a general sequence of perturbed or regularized games have been carried out by Loridan and Morgan. They started in [71, 70, 68] with the case where the second-stage problem (P_x) has a unique solution for any x and there is no difference between the optimistic and the pessimistic bilevel optimization problem; then, they extensively investigated in [69, 75, 73, 77, 76, 78, 74] pessimistic bilevel problems in which the follower's constraints do not depend on the leader's actions.

In this section, we illustrate regularization methods for the bilevel problems (PB) and (OB) and also for bilevel problems in which the follower's problem is modeled by an *equilibrium* or a *quasi-equilibrium* problem. Due to the relevance of inequality constraints investigation in optimization theory, we assume that the follower's constraints are described by a finite number of inequalities whereas, for the sake of brevity, the leader's constraints are assumed to be constant. Then, we set

$$K(x) = \bigcap_{i=1}^k K_i(x) = \bigcap_{i=1}^k \{y \in Y \text{ such that } g_i(x, y) \leq 0\}, \quad (2)$$

where X and Y are nonempty subsets of two Euclidean spaces, g_i is a function from $X \times Y$ to \mathbb{R} for $i = 1, \dots, k$, and we denote by:

- \mathcal{S} the set of solutions to the optimistic bilevel problem (OB) , i.e. the set of strong Stackelberg solutions of (OB) .
- \mathcal{W} the set of solutions to the pessimistic bilevel problem (PB) , i.e. the set of weak Stackelberg solutions of (PB) .

3.1 Regularization of pessimistic bilevel optimization problems

First investigations on pessimistic bilevel optimization problems mainly consisted in trying to obviate the lack of solutions, which may arise, as shown in Example 2.1, even in very simple models. In fact, in that example the marginal function

$$e(x) = \sup_{y \in M(x)} L(x, y)$$

is not lower semicontinuous and this is a possible consequence of the fact that the map M does not enjoy an important property, namely *lower semicontinuity*.

Definition 3.1 [6] Let X and Y be nonempty subsets of two Euclidean spaces.

A set-valued map $T : x \in X \rightrightarrows T(x) \subseteq Y$ is lower semicontinuous over X if for any $x \in X$, for any sequence $(x_n)_n$ converging to x in X and any $y \in T(x)$ there exists a sequence $(y_n)_n$ converging to y in Y such that $y_n \in T(x_n)$ for n sufficiently large, that is

$$T(x) \subseteq \operatorname{Lim\,inf}_n T(x_n).$$

A set-valued map T from X to Y is closed over X if the graph of T is closed in $X \times Y$, that is for any sequence $(x_n, y_n)_n$ converging to $(x, y) \in X \times Y$, such that $y_n \in T(x_n)$ for $n \in \mathbb{N}$, one has $y \in T(x)$, equivalent to

$$\operatorname{Lim\,sup}_n T(x_n) \subseteq T(x).$$

Here, $\operatorname{Lim\,inf}$ and $\operatorname{Lim\,sup}$ denote, respectively, the Painlevé-Kuratowski lower and upper limit of a sequence of sets.

Then, let us recall the well-known *maximum theorem* [16, 6].

Lemma 3.1. *If the following assumptions are satisfied:*

- *the function L is lower semicontinuous over $X \times Y$;*
 - *the set-valued map $T : X \rightrightarrows Y$ is lower semicontinuous over X ;*
- then the marginal function defined by*

$$\sup_{y \in T(x)} L(x, y)$$

is lower semicontinuous over X .

3.1.1 Approximate and strict approximate solutions

An attempt to present a complete discussion, which aimed to unify all previous results in a general scheme, was given in [61] in the setting of topological vector spaces and by analyzing the second-stage problem in the presence and in the absence of general or inequality constraints. Unfortunately, using Γ -limits and working in a topological setting have restricted the readability and the applicability of that paper, so, here we recall the main idea and the principal results in a simplified form but under more restrictive assumptions.

Let $\varepsilon > 0$. The ε -argmin map

$$M^\varepsilon : x \in X \rightrightarrows M^\varepsilon(x) = \left\{ y \in K(x) \text{ such that } F(x, y) \leq \inf_{z \in K(x)} F(x, z) + \varepsilon \right\},$$

can be a potential candidate to substitute the argmin map M but, unfortunately, also M^ε may fail to be lower semicontinuous even when the constraints do not depend on x .

Example 3.1 [57, Ex. 2.3] Let $X = Y = [0, 1]$, $K(x) = [0, 1]$ for any $x \in X$. Consider $F(x, y) = -y^2 + (1 + x)y - x$ for every $(x, y) \in [0, 1]^2$ and any $\varepsilon \in [0, 1/4[$. Then, one gets:

- $M(0) = \{0, 1\}$ and $M(x) = \{0\}$ for every $x \in]0, 1]$;
- $M^\varepsilon(x) = [0, \alpha^\varepsilon(x)] \cup]\beta^\varepsilon(x), 1]$, if $x \in [0, \varepsilon]$;
- $M^\varepsilon(x) = [0, \alpha^\varepsilon(x)]$, if $x \in]\varepsilon, 1]$;

where $\alpha^\varepsilon(x) = \frac{1}{2} \left(x + 1 - \sqrt{(x+1)^2 - 4\varepsilon} \right)$ and $\beta^\varepsilon(x) = \frac{1}{2} \left(x + 1 + \sqrt{(x+1)^2 - 4\varepsilon} \right)$ solve the equation

$$y^2 - (x+1)y + \varepsilon = 0.$$

So, both the maps M and M^ε are not lower semicontinuous over X : M lacks the lower semicontinuity property at the point $x = 0$ whereas M^ε at the point $x = \varepsilon$.

Employing strict inequalities and then passing to large inequalities, when possible, is the right way to proceed in order to get the lower semicontinuity property of the ε -minima, so we consider

$$\widetilde{M}^\varepsilon : x \in X \Rightarrow \widetilde{M}^\varepsilon(x) = \left\{ y \in K(x) \text{ such that } F(x, y) < \inf_{z \in K(x)} F(x, z) + \varepsilon \right\},$$

and the example above can be completed computing the strict ε -minima

- $\widetilde{M}^\varepsilon(x) = [0, \alpha^\varepsilon(x)[\cup]\beta^\varepsilon(x), 1]$, if $x \in [0, \varepsilon[$
- $\widetilde{M}^\varepsilon(x) = [0, \alpha^\varepsilon(x)[$, if $x \in]\varepsilon, 1]$

and realizing that $\widetilde{M}^\varepsilon$ is lower semicontinuous over X .

In order to prove the lower semicontinuity of $\widetilde{M}^\varepsilon$, a crucial role is played both by the lower semicontinuity and closedness of the constraints map K and by the continuity of the follower's payoff function F , as shown by the next lemmas, which concern, respectively, the continuity and convexity properties of inequality constraints maps and the lower semicontinuity of $\widetilde{M}^\varepsilon$ and M^ε . The first lemma can be found in [13, theorems 3.1.1 and 3.1.6].

Lemma 3.2. *If the set X is closed and the following assumption is satisfied:*

$C_1)$ *for every $i = 1, \dots, k$, the function g_i is lower semicontinuous over $X \times Y$;*

then, the set-valued map K is closed on X .

If the following assumptions are satisfied:

$C_2)$ *for every $i = 1, \dots, k$ and $y \in Y$, the function $g_i(\cdot, y)$ is upper semicontinuous over X ;*

$C_3)$ *for every $i = 1, \dots, k$ and $x \in X$, the function $g_i(x, \cdot)$ is strictly quasiconvex (see [13]) and upper semicontinuous on Y , which is assumed to be convex;*

$C_4)$ *for every $x \in X$ there exists $y \in Y$ such that $\max_{i=1, \dots, k} g_i(x, y) < 0$;*

then, the set-valued map K is lower semicontinuous and convex-valued over X .

Remark 3.1 We emphasize that the strict quasiconvexity of the function $g_i(x, \cdot)$ guarantees that

$$K_i(x) \subseteq \text{cl}(\text{int } K_i(x))$$

where the map K_i is defined in (2). This condition is crucial in proving the lower semicontinuity of the map K_i , but K_i can be lower semicontinuous even if the function $g_i(x, \cdot)$ is not strictly quasiconvex as it occurs, for example, with $g_i(x, y) = x^2 - y^2$.

Lemma 3.3. [61, Th. 4.3] *Under the assumptions $C_1) - C_4)$, if the set X is closed, the set Y is convex and compact and the function F is continuous over $X \times Y$, then the map $\widetilde{M}^\varepsilon$ is lower semicontinuous over X .*

In order to get the lower semicontinuity of M^ε , a crucial role is also played by suitable *convexity* properties of the data.

Lemma 3.4. [61, Th. 5.4] *Under the assumptions of Lemma 3.3, if also the following condition holds:*

$H_1)$ *the function $F(x, \cdot)$ is strictly quasiconvex on $K(x)$, for every $x \in X$;*

then, the map M^ε is lower semicontinuous over X .

The next step consists in achieving the lower semicontinuity of the marginal functions

$$e^\varepsilon : x \in X \rightarrow \sup_{y \in M^\varepsilon(x)} L(x, y) \quad \widetilde{e}^\varepsilon : x \in X \rightarrow \sup_{y \in \widetilde{M}^\varepsilon(x)} L(x, y),$$

using Lemma 3.1, and then the existence of two types of approximate solutions for the problem (PB), \mathcal{W}^ε and $\widetilde{\mathcal{W}}^\varepsilon$, defined by

$$\begin{aligned} x^\varepsilon \in \mathcal{W}^\varepsilon &\iff e(x^\varepsilon) = \inf_{x \in X} e^\varepsilon(x) \iff \sup_{y \in M^\varepsilon(x^\varepsilon)} L(x^\varepsilon, y) = \inf_{x \in X} \sup_{y \in M^\varepsilon(x)} L(x, y), \\ \widetilde{x}^\varepsilon \in \widetilde{\mathcal{W}}^\varepsilon &\iff e(\widetilde{x}^\varepsilon) = \inf_{x \in X} \widetilde{e}^\varepsilon(x) \iff \sup_{y \in \widetilde{M}^\varepsilon(\widetilde{x}^\varepsilon)} L(\widetilde{x}^\varepsilon, y) = \inf_{x \in X} \sup_{y \in \widetilde{M}^\varepsilon(x)} L(x, y). \end{aligned}$$

Proposition 3.1. *Assume that the set X is compact and the function L is lower semicontinuous over $X \times Y$.*

If assumptions of Lemma 3.3 hold, then the approximate solutions set $\widetilde{\mathcal{W}}^\varepsilon$ is nonempty for any $\varepsilon > 0$.

If assumptions of Lemma 3.4 hold, then the approximate solutions set \mathcal{W}^ε is nonempty for any $\varepsilon > 0$.

Now, two questions naturally arise:

- do the approximate security values w^ε and $\widetilde{w}^\varepsilon$, defined respectively by

$$w^\varepsilon = \inf_{x \in X} \sup_{y \in M^\varepsilon(x)} L(x, y) \quad \widetilde{w}^\varepsilon = \inf_{x \in X} \sup_{y \in \widetilde{M}^\varepsilon(x)} L(x, y),$$

converge towards the security value w of the original problem when ε tends to zero?

- are such approximate solutions *stable* with respect to perturbations of the data, for

any fixed $\varepsilon > 0$?

The term *stable* should be intended in the following sense. Assume that the data of the problem (PB) are asymptotically approached by the data of the perturbed problems

$$(PB)_n \quad \text{find } x_n \in X \text{ such that } \sup_{y \in M_n(x_n)} L_n(x_n, y) = \inf_{x \in X} \sup_{y \in M_n(x)} L_n(x, y),$$

where $(F_n)_n$ and $(L_n)_n$ are sequences of functions from $X \times Y$ to \mathbb{R} . Let, for any n , M_n^ε be the ε -argmin map of F_n with constraints K_n . More specifically, given k sequences of real valued functions $(g_{i,n})_n$ defined on $X \times Y$, we set

$$K_n(x) = \bigcap_{i=1}^k \{y \in Y \text{ such that } g_{i,n}(x, y) \leq 0\} \quad (3)$$

and we denote by M_n^ε the map

$$M_n^\varepsilon : x \in X \Rightarrow M_n^\varepsilon(x) = \left\{ y \in K_n(x) : F_n(x, y) \leq \inf_{z \in K_n(x)} F_n(x, z) + \varepsilon \right\}$$

and by $\mathcal{W}_n^\varepsilon$ the set

$$\mathcal{W}_n^\varepsilon = \left\{ x \in X : \sup_{y \in M_n^\varepsilon(x)} L_n(x, y) = \inf_{u \in X} \sup_{y \in M_n^\varepsilon(u)} L_n(u, y) \right\}.$$

Then, one asks under which assumptions any convergent sequence $(x_n^\varepsilon)_n$ of approximate solutions for $(PB)_n$ converges to $x^\varepsilon \in \mathcal{W}^\varepsilon$ and one has

$$\text{Limsup}_n \mathcal{W}_n^\varepsilon \subseteq \mathcal{W}^\varepsilon, \quad \forall \varepsilon > 0. \quad (4)$$

The following propositions answer to both questions.

Proposition 3.2. [57, Cor. 2.2] *Under the assumptions of Lemma 3.3, if in addition the function L is upper semicontinuous over $X \times Y$, then*

$$\lim_{\varepsilon \rightarrow 0} \tilde{w}^\varepsilon = w = \lim_{\varepsilon \rightarrow 0} w^\varepsilon.$$

We remark that:

- under the assumptions of Proposition 3.2, the map M^ε may be not lower semicontinuous over X , so this property is not necessary for the approximate security values convergence;
- no convexity assumption is made on g_i , F and L .

Proposition 3.3. [61, Th. 7.6] *Assume that X is compact and Y is convex and compact. If the sequences of functions $(F_n)_n$ and $(g_{i,n})_n$ continuously converge over $X \times Y$ to the functions F and g_i respectively, L is lower semicontinuous over $X \times Y$, assumptions $C_1) - C_4)$ and the following hold for any $n \in \mathbb{N}$:*

$C_{3,n})$ for every $i = 1, \dots, k$ and $x \in X$, the function $g_{i,n}(x, \cdot)$ is strictly quasi-convex on Y ;

$C_{4,n}$) for every $x \in X$ there exists $y \in Y$ such that $\max_{i=1,\dots,k} g_{i,n}(x, y) < 0$;

$H_{2,n}$) for any $(x, y) \in X \times Y$ and any sequence $(x_n, y_n)_n$ converging to (x, y) in $X \times Y$ one has

$$\limsup_{n \rightarrow +\infty} L_n(x_n, y_n) \leq L(x, y),$$

then the ε -solutions are stable, i.e inclusion in (4) holds, for any $\varepsilon > 0$.

We stress that, in general, one cannot prove a result analogous to (4) also for the strict approximate solutions $\widetilde{W}^\varepsilon$, due to the *open* nature of them. A natural extension of the previous investigations was the research of possible candidates to substitute the lacking weak Stackelberg solutions to (PB), which will be presented in the next subsection.

We mention that the approximate solutions presented above have been afterwards employed in [3, 1] (about the *strong-weak Stackelberg problem*, which generalizes problems (PB) and (OB)) and in [45].

3.1.2 Inner regularizations and viscosity solutions

The following procedure seems to be quite natural to obviate the lack of solutions for pessimistic bilevel optimization problems:

- defining regularizing problems which have solutions under not too restrictive conditions;
- utilizing sequences of solutions to the regularizing problems admitting convergent subsequences;
- considering a limit point of one of such subsequences as a *surrogate* solution to (PB) provided that the supremum values of the objective functions calculated in such approximating solutions converge to the security value w .

To accomplish the first step, we identify the set-valued map properties that are helpful in this regularization process.

Definition 3.2 [57, Def. 2.1] A family of set-valued maps $\mathfrak{T} = \{\mathcal{T}^\varepsilon, \varepsilon > 0\}$, where

$$\mathcal{T}^\varepsilon : x \in X \rightrightarrows \mathcal{T}^\varepsilon(x) \subseteq Y,$$

is an inner regularization for the family of minimum problems $\{(P_x), x \in X\}$ if the conditions below are satisfied:

R_1) $M(x) \subseteq \mathcal{T}^\varepsilon(x) \subseteq \mathcal{T}^\eta(x)$ for every $x \in X$ and $0 < \varepsilon < \eta$;

R_2) for any $x \in X$, any sequence $(x_n)_n$ converging to x in X and any sequence of positive numbers $(\varepsilon_n)_n$ decreasing to zero, one has

$$\text{Lim sup}_n \mathcal{T}^{\varepsilon_n}(x_n) \subseteq M(x);$$

R_3) \mathcal{T}^ε is a lower semicontinuous set-valued map on X , for every $\varepsilon > 0$.

The above definition is related to the pessimistic nature of the problem (PB) and we are aware that different problems at the upper level would require different definitions of inner regularizations for the family $\{(P_x), x \in X\}$.

The strict and large ε -argmin maps, $\widetilde{\mathfrak{M}} = \{\widetilde{M}^\varepsilon, \varepsilon > 0\}$ and $\mathfrak{M} = \{M^\varepsilon, \varepsilon > 0\}$, constitute inner regularization classes under appropriate conditions.

Proposition 3.4. [57, Cor. 2.1] *The family $\widetilde{\mathfrak{M}}$ is an inner regularization whenever the assumptions of Lemma 3.3 are satisfied.*

The family \mathfrak{M} is an inner regularization whenever the assumptions of Lemma 3.4 are satisfied.

Other inner regularization classes can be found in [57] and also in [56], where the constraints of the second-stage problem are allowed to be violated. Here, we only mention the following ones:

$$M_d^\varepsilon: x \in X \Rightarrow M_d^\varepsilon(x) = \left\{ y \in Y : d(y, K(x)) \leq \varepsilon \text{ and } F(x, y) \leq \inf_{z \in K(x)} F(x, z) + \varepsilon \right\},$$

$$\widetilde{M}_d^\varepsilon: x \in X \Rightarrow \widetilde{M}_d^\varepsilon(x) = \left\{ y \in Y : d(y, K(x)) < \varepsilon \text{ and } F(x, y) < \inf_{z \in K(x)} F(x, z) + \varepsilon \right\}.$$

Proposition 3.5. *The family $\widetilde{\mathfrak{M}}_d = \{\widetilde{M}_d^\varepsilon, \varepsilon > 0\}$ is an inner regularization whenever the assumptions of Lemma 3.3 are satisfied.*

The family $\mathfrak{M}_d = \{M_d^\varepsilon, \varepsilon > 0\}$ is an inner regularization whenever the assumptions of Lemma 3.4 are satisfied.

Proof. The result follows by arguing as in [56] and using Lemma 3.2. □

However, assumption C_4), which is crucial for the lower semicontinuity of the constraints map K , could be not satisfied, so a good alternative to M_d^ε and $\widetilde{M}_d^\varepsilon$ could be the maps:

$$G^\varepsilon: x \in X \Rightarrow G^\varepsilon(x) = \bigcap_{i=1}^k \left\{ y \in Y : g_i(x, y) \leq \varepsilon \text{ and } F(x, y) \leq \inf_{z \in K(x)} F(x, z) + \varepsilon \right\},$$

$$\widetilde{G}^\varepsilon: x \in X \Rightarrow \widetilde{G}^\varepsilon(x) = \bigcap_{i=1}^k \left\{ y \in Y : g_i(x, y) < \varepsilon \text{ and } F(x, y) < \inf_{z \in K(x)} F(x, z) + \varepsilon \right\}.$$

Proposition 3.6. *Assume that Y is convex and compact and F is continuous over $X \times Y$. Then:*

- *the family $\widetilde{\mathfrak{G}} = \{\widetilde{G}^\varepsilon, \varepsilon > 0\}$ is an inner regularization whenever assumptions $C_1) - C_3)$ and the following hold:*

$C_5)$ *for every $\varepsilon > 0$ and $x \in X$ there exists $y \in Y$ such that $\max_{i=1, \dots, k} g_i(x, y) < \varepsilon$;*

- *the family $\mathfrak{G} = \{G^\varepsilon, \varepsilon > 0\}$ is an inner regularization whenever the assumptions $C_1) - C_3)$, $C_5)$ and $H_1)$ are satisfied.*

Proof. Assumptions $C_1)$, $C_2)$ and $C_5)$ guarantee that the map

$$x \in X \Rightarrow \bigcap_{i=1}^k \{y \in Y : g_i(x, y) < \varepsilon\}$$

is lower semicontinuous over $X \times Y$, as well assumptions $C_1) - C_3)$ and $C_5)$ guarantee that the map

$$x \in X \Rightarrow \bigcap_{i=1}^k \{y \in Y : g_i(x, y) \leq \varepsilon\}$$

is lower semicontinuous over $X \times Y$. Then, the result can be proved arguing as in propositions 2.1, 2.2 and 2.4 in [56]. \square

The next steps consist in introducing the concept of *viscosity solution* for pessimistic bilevel optimization problems related to an inner regularization class and in stating a related existence result.

Definition 3.3 [57, Def. 2.2] Let \mathfrak{T} be an inner regularization for the family $\{P_x, x \in X\}$. A point $\bar{x} \in X$ is a \mathfrak{T} -viscosity solution for (PB) if for every sequence $(\varepsilon_n)_n$ decreasing to zero there exists $(x^{\varepsilon_n})_n$, $x^{\varepsilon_n} \in X$ for any $n \in \mathbb{N}$, such that:

- $V_1)$ a subsequence of $(x^{\varepsilon_n})_n$ converges to \bar{x} ;
- $V_2)$ $\sup_{y \in \mathcal{T}^{\varepsilon_n}(x^{\varepsilon_n})} L(x^{\varepsilon_n}, y) = \inf_{x \in X} \sup_{y \in \mathcal{T}^{\varepsilon_n}(x)} L(x, y) \quad \forall n \in \mathbb{N}$;
- $V_3)$ $\lim_{n \rightarrow +\infty} \sup_{y \in \mathcal{T}^{\varepsilon_n}(x^{\varepsilon_n})} L(x^{\varepsilon_n}, y) = \inf_{x \in X} \sup_{y \in M(x)} L(x, y) = w$.

Theorem 3.1. [57, Prop. 4.1] *If $\mathfrak{T} = \{\mathcal{T}^\varepsilon, \varepsilon > 0\}$ is an inner regularization for the family $\{(P_x), x \in X\}$, the sets X and Y are compact and the function L is continuous over $X \times Y$, then there exists a \mathfrak{T} -viscosity solution for the pessimistic problem (PB) .*

Suitable existence results of viscosity solutions with respect to each of the families considered above can be derived from Theorem 3.1 and propositions 3.4, 3.5 and 3.6.

The next example illustrates the above procedure in a simple case.

Example 3.2 Let $X = [0, 1]$, $Y = [-1, 1]$, $L(x, y) = x + y$ and $F(x, y) = xy$ for every $(x, y) \in X \times Y$. Consider $i = 1$ and $g_1(x, y) = x^2 - y^2$, so that $K(x) = [-1, -x] \cup [x, 1]$. If $\varepsilon \in]0, 2]$, it is easy to check that:

- the argmin map M of the minima to (P_x) is not lower semicontinuous at $x = 0$, since $M(0) = Y$ and $M(x) = \{-1\}$ for every $x \in]0, 1]$;
- the marginal function $e(x)$ is not lower semicontinuous at $x = 0$, since $e(0) = 1$ and $e(x) = x - 1$ for every $x \in]0, 1]$;
- the problem (PB) does not have a solution since $w = \inf_{x \in X} \sup_{y \in M(x)} L(x, y) = -1$ but $e(x) > -1$ for every $x \in [0, 1]$;

- the map $\widetilde{M}_d^\varepsilon$ is lower semicontinuous over X since $\widetilde{M}_d^\varepsilon(x) = [-1, 1]$ if $x \in [0, \varepsilon/2[$ and $\widetilde{M}_d^\varepsilon(x) = [-1, -1 + \varepsilon/x[$ if $x \in [\varepsilon/2, 1]$;
- the minimum point of the marginal function $\sup_{y \in \widetilde{M}_d^\varepsilon(\cdot)} L(\cdot, y)$ is $\widetilde{x}_d^\varepsilon = \sqrt{\varepsilon}$ and

$$\widetilde{w}_d^\varepsilon = \inf_{x \in X} \sup_{y \in \widetilde{M}_d^\varepsilon(x)} L(x, y) = -1 + 2\sqrt{\varepsilon};$$

- the family $\widetilde{\mathfrak{M}}_d$ is an inner regularization even if the function $g_1(x, \cdot)$ is not strictly quasiconvex (see Remark 3.1);
- the data satisfy the conditions in Theorem 3.1;
- the $\widetilde{\mathfrak{M}}_d$ -viscosity solution for (PB) is $x = 0$.

3.2 Approximation of optimistic bilevel optimization problems

As we observed in paragraph C of Section 2, a strong Stackelberg equilibrium exists under not too restrictive conditions. However, when such a problem has to be approached by a sequence of perturbed problems (e.g. in discretization process of (OB) in an infinite dimensional setting), the perturbed solutions may not converge towards a solution of the original problem, that is such problems are lacking in stability in the sense specified in Subsection 3.1.1.

More precisely, assume that the data of the problem (OB) are asymptotically approached by the data of the perturbed problems

$$(OB)_n \text{ find } x_n \in X \text{ such that } \inf_{y \in M_n(x_n)} L_n(x_n, y) = \inf_{x \in X} \inf_{y \in M_n(x)} L_n(x, y),$$

where $(F_n)_n$ and $(L_n)_n$ are sequences of functions from $X \times Y$ to \mathbb{R} and, for any n , K_n is defined as in (3) and M_n is the argmin map of F_n with constraints K_n . Then, denoted by \mathcal{S}_n the set of solutions to $(OB)_n$, for any $n \in \mathbb{N}$, one asks if one has

$$\text{Lim sup}_n \mathcal{S}_n \subseteq \mathcal{S}$$

and the answer is negative, in general.

Example 3.3 Let $X = Y = [0, 1]$, $K(x) = K_n(x) = Y$, $F_n(x, y) = y/n$ and $L_n(x, y) = -xy + 1/n$. It is easy to see that:

- $(F_n)_n$ and $(L_n)_n$ uniformly converge, and therefore also continuously converge, to the functions defined by $F(x, y) = 0$ and $L(x, y) = -xy$ respectively;
- $M(x) = [0, 1]$ and, for any $n \in \mathbb{N}$, $M_n(x) = \{0\}$ for every $x \in [0, 1]$;
- $\mathcal{S} = \{1\}$ and, for any $n \in \mathbb{N}$, $\mathcal{S}_n = [0, 1]$.

Then, $\text{Lim sup}_n \mathcal{S}_n \not\subseteq \mathcal{S}$.

Nevertheless, also in the optimistic case, regularizing the follower's reaction set turns out to be useful to obtain satisfactory approximation results for the strong Stackelberg values and solutions.

In line with the pessimistic case, we define:

$$\begin{aligned}
s^\varepsilon &= \inf_{x \in X} \inf_{y \in M^\varepsilon(x)} L(x, y) \quad \text{and} \quad \tilde{s}^\varepsilon = \inf_{x \in X} \inf_{y \in \tilde{M}^\varepsilon(x)} L(x, y), \\
\mathcal{S}^\varepsilon &= \left\{ x \in X : \inf_{y \in M^\varepsilon(x)} L(x, y) = \inf_{u \in X} \inf_{y \in M^\varepsilon(u)} L(u, y) \right\}, \\
\tilde{\mathcal{S}}^\varepsilon &= \left\{ x \in X : \inf_{y \in \tilde{M}^\varepsilon(x)} L(x, y) = \inf_{u \in X} \inf_{y \in \tilde{M}^\varepsilon(u)} L(u, y) \right\}, \\
\mathcal{S}_n^\varepsilon &= \left\{ x \in X : \inf_{y \in M_n^\varepsilon(x)} L_n(x, y) = \inf_{u \in X} \inf_{y \in M_n^\varepsilon(u)} L_n(u, y) \right\}.
\end{aligned}$$

Proposition 3.7. [60, Th. 3.2] *Under the assumptions of Lemma 3.3, if in addition: the function L is lower semicontinuous over $X \times Y$, then*

$$\lim_{\varepsilon \rightarrow 0} s^\varepsilon = s = \lim_{\varepsilon \rightarrow 0} \tilde{s}^\varepsilon \quad \text{and} \quad \text{Lim sup}_{\varepsilon \rightarrow 0} \mathcal{S}^\varepsilon \subseteq \mathcal{S}.$$

Proposition 3.8. [60, Th. 5.6] *Assume that X is compact and Y is convex and compact. If the sequences of functions $(F_n)_n$ and $(g_{i,n})_n$ continuously converge over $X \times Y$ to the functions F and g_i respectively, assumptions $C_1) - C_4)$, $C_{3,n})$, $C_{4,n})$, $H_1)$, $H_{2,n})$ and the following hold for any $n \in \mathbb{N}$:*

$H_{3,n})$ for any $(x, y) \in X \times Y$ and any sequence $(x_n, y_n)_n$ converging to (x, y) in $X \times Y$ one has

$$L(x, y) \leq \liminf_{n \rightarrow +\infty} L_n(x_n, y_n);$$

then, for any $\varepsilon > 0$

$$\text{Lim sup}_n \mathcal{S}_n^\varepsilon \subseteq \mathcal{S}^\varepsilon.$$

Moreover, one has

$$\text{Lim sup}_{\varepsilon \rightarrow 0} \left(\text{Lim sup}_n \mathcal{S}_n^\varepsilon \right) \subseteq \mathcal{S}.$$

However, in this approximation process, the role of ε and n cannot be inverted and, in general, one cannot get a result of the type

$$\text{Lim sup}_n \mathcal{S}^{\varepsilon_n} \subseteq \mathcal{S}, \tag{5}$$

with $(\varepsilon_n)_n$ decreasing to zero, as shown by the next example.

Example 3.4 Let $X = Y = [0, 1]$, $K(x) = K_n(x) = Y$, $F_n(x, y) = y/n$ and $L_n(x, y) = x(x - y) + 1/n$. It can be computed that:

- $(F_n)_n$ and $(L_n)_n$ continuously converge on $X \times Y$ to the functions defined by $F(x, y) = 0$ and $L(x, y) = x(x - y)$ respectively;

- $M(x) = M^\varepsilon(x) = [0, 1]$ for every $x \in X$;
- for any $n \in \mathbb{N}$, $M_n^\varepsilon(x) = [0, n\varepsilon]$ if $n\varepsilon \leq 1$ and $M_n^\varepsilon(x) = [0, 1]$ if $n\varepsilon > 1$;
- $\mathcal{S} = \mathcal{S}^\varepsilon = \{1/2\}$ and, for any $n \in \mathbb{N}$, $\mathcal{S}_n^\varepsilon = \{n\varepsilon/2\}$ if $n\varepsilon \leq 1$ and $\mathcal{S}_n^\varepsilon = \{1/2\}$ if $n\varepsilon > 1$.

Therefore,

$$\limsup_{\varepsilon \rightarrow 0} \left(\limsup_n \mathcal{S}_n^\varepsilon \right) = \mathcal{S} \quad \text{and} \quad \limsup_n \left(\limsup_{\varepsilon \rightarrow 0} \mathcal{S}_n^\varepsilon \right) = \{0\} \not\subseteq \mathcal{S},$$

and we remark that $\limsup_n \mathcal{S}^{\varepsilon_n} \subseteq \mathcal{S}$ if the sequence $(\varepsilon_n)_n$ is infinitesimal of the first order.

It would be interesting defining suitable viscosity solutions also for optimistic bilevel optimization problems aimed to reach an inclusion analogous to (5).

This is a still unexplored topic, but given an optimistic bilevel problem approached by a sequence of approximating optimistic bilevel problems $(OB)_n$, it seems to be reasonable investigating the limit points of sequences of solutions to $(OB)_n$ that are not so far from the solution set to (OB) .

We mention that algorithms to approach approximate solutions to (OB) have been developed in [66], where the reformulation of (OB) via the optimal value function is exploited, and in [44], where a reformulation of (OB) as a generalized Nash equilibrium problem is employed.

3.3 Regularization of bilevel problems with equilibrium constraints

Stackelberg games are models that can be naturally extended to the case where there are one leader and one, or more than one, follower who solve a second-stage problem described by a parametric variational inequality, Nash equilibrium, quasi-variational inequality or more generally quasi-equilibrium. The critical issues of pessimistic and optimistic approach are still present and regularizing the problem solved by the follower(s) is useful also in this case. First we present the case where one solves a variational inequality in the second stage. The set of solutions of such a parametric variational inequality can be considered as the follower's constraints in the first stage, so traditionally the associate bilevel problem is called *bilevel problems with variational inequality constraints*, and analogously for the other problems considered in the second stage.

3.3.1 Variational inequality constraints

Assume that A is a function from $X \times Y$ to Y and consider the family of variational inequalities $\{(V_x), x \in X\}$, where any problem (V_x) consists in finding, see [12],

$$y \in K(x) \quad \text{such that} \quad \langle A(x, y), y - w \rangle \leq 0 \quad \forall w \in K(x).$$

Then, the pessimistic and the optimistic bilevel problems with variational inequality in the second stage are respectively defined by:

$$(PBVI) \quad \text{find } x \in X \text{ such that } \sup_{y \in \mathcal{V}(x)} L(x, y) = \inf_{u \in X} \sup_{y \in \mathcal{V}(u)} L(u, y),$$

$$(OBVI) \quad \text{find } x \in X \text{ such that } \inf_{y \in \mathcal{V}(x)} L(x, y) = \inf_{u \in X} \inf_{y \in \mathcal{V}(u)} L(u, y),$$

where, for any x , $\mathcal{V}(x)$ is the set of solutions to the variational inequality (V_x) .

In this case, given a positive number ε , fruitful regularizations of the follower's reaction set consist in finding

$$y \in K(x) \quad \text{such that} \quad \langle A(x, y), y - w \rangle \leq \varepsilon \quad \forall w \in K(x)$$

or

$$y \in K(x) \quad \text{such that} \quad \langle A(x, w), y - w \rangle \leq \varepsilon \quad \forall w \in K(x),$$

together with their *strict* versions. This double possibility of regularizing the map \mathcal{V} follows from using the Minty type variational inequality, which consists in finding

$$y \in K(x) \quad \text{such that} \quad \langle A(x, w), y - w \rangle \leq 0 \quad \forall w \in K(x),$$

use which is essential in infinite dimensional spaces, as displayed in [12].

The above regularizations have been employed for investigating, in a theoretical setting, well-posedness in [64] and stability properties in [51] of $(OBVI)$ in Banach spaces; moreover, we mention that an exact penalization scheme has been proposed in [110] for deriving necessary optimality conditions of $(OBVI)$. Convergence properties of the approximate weak Stackelberg values of $(PBVI)$ can be found in [52] in finite dimensional spaces. They have been also considered, in applied framework as truss topology optimization in [36] or climate regulation policy in [34].

3.3.2 Nash equilibrium constraints

Consider $Y = \prod_{j=1}^l Y_j$ and l real-valued functions F_1, \dots, F_l defined on $X \times Y$. Then, for any $x \in X$, the Nash equilibrium problem (NE_x) consists in finding

$$\mathbf{y} \in Y \quad \text{such that} \quad F_j(x, \mathbf{y}) \leq \inf_{u_j \in Y_j} F_j(x, u_j, \mathbf{y}_{-j}) \quad \forall j = 1, \dots, l,$$

where $\mathbf{y}_{-j} = (y_1, \dots, y_{j-1}, y_{j+1}, \dots, y_l)$, and the pessimistic and the optimistic bilevel problems with Nash equilibrium problem in the second stage are respectively defined by:

$$(PBNE) \quad \text{find } x \in X \text{ such that } \sup_{\mathbf{y} \in \mathcal{N}(x)} L(x, \mathbf{y}) = \inf_{u \in X} \sup_{\mathbf{y} \in \mathcal{N}(u)} L(u, \mathbf{y}),$$

$$(OBNE) \quad \text{find } x \in X \text{ such that } \inf_{\mathbf{y} \in \mathcal{N}(x)} L(x, \mathbf{y}) = \inf_{u \in X} \inf_{\mathbf{y} \in \mathcal{N}(u)} L(u, \mathbf{y}),$$

where, for any x , $\mathcal{N}(x)$ is the set of solutions to the Nash equilibrium problem (NE_x). Regularizing the follower's reaction set consists, in this case, in considering:

$$\mathcal{N}^\varepsilon(x) = \left\{ \mathbf{y} \in Y : F_j(x, \mathbf{y}) \leq \inf_{u_j \in Y_j} F_j(x, u_j, \mathbf{y}_{-j}) + \varepsilon \quad \forall j = 1, \dots, l \right\},$$

that is the classical set of approximate Nash equilibria (see, for example, [10]), or

$$\widehat{\mathcal{N}}^\varepsilon(x) = \left\{ \mathbf{y} \in Y : \sum_{j=1}^l F_j(x, \mathbf{y}) \leq \inf_{u_j \in Y_j} \sum_{j=1}^l F_j(x, u_j, \mathbf{y}_{-j}) + \varepsilon \right\},$$

introduced and investigated in [94]. These sets do not coincide in general, as shown in [94, Ex. 3.6], but both can be used to study parametric well-posedness properties of Nash equilibrium problems since, as proven in [50],

$$\lim_{\varepsilon \rightarrow 0} \text{diam} \mathcal{N}^\varepsilon(x) = 0 \iff \lim_{\varepsilon \rightarrow 0} \text{diam} \widehat{\mathcal{N}}^\varepsilon(x) = 0.$$

Well posedness of optimistic bilevel problems with Nash equilibrium problem in the second stage has been investigated in that paper, whereas viscosity solutions for pessimistic ones have been introduced in [63] as a prototype for the concepts of viscosity solutions for (PB) later developed.

3.3.3 Quasi-variational inequality constraints

Assume that T is a set-valued map from $X \times Y$ to Y , A is a function from $X \times Y$ to Y , and consider the family of quasi-variational inequalities $\{(Q_x), x \in X\}$, where any (Q_x) consists in finding, see [12],

$$y \in Y \text{ such that } y \in T(x, y) \text{ and } \langle A(x, y), y - w \rangle \leq 0 \quad \forall w \in T(x, y).$$

Then, the pessimistic and the optimistic bilevel problems with quasi-variational inequality in the second stage are defined respectively by:

$$(PBQI) \quad \text{find } x \in X \text{ such that } \sup_{y \in \mathcal{Q}(x)} L(x, y) = \inf_{u \in X} \sup_{y \in \mathcal{Q}(u)} L(u, y),$$

$$(OBQI) \quad \text{find } x \in X \text{ such that } \inf_{y \in \mathcal{Q}(x)} L(x, y) = \inf_{u \in X} \inf_{y \in \mathcal{Q}(u)} L(u, y),$$

where, for any x , $\mathcal{Q}(x)$ is the set of solutions to the quasi-variational inequality (Q_x) .

In this case, given a positive number ε , fruitful regularizations of the follower's reaction set consist in finding

$$y \in Y \text{ such that } d(y, T(x, y)) \leq \varepsilon \text{ and } \langle A(x, y), y - w \rangle \leq \varepsilon \quad \forall w \in T(x, y)$$

or

$$y \in Y \text{ such that } d(y, T(x, y)) \leq \varepsilon \text{ and } \langle A(x, w), y - w \rangle \leq \varepsilon \quad \forall w \in T(x, y)$$

together with their *strict* versions.

Requiring that an approximate solution to (Q_x) may violate the constraints, provided

that it remains in a neighborhood of them, is quite necessary in quasi-variational inequality setting where a fixed-point problem is involved, as shown in [48, Ex. 2.1]. The above regularizations have been used in [53] for establishing convergence results of the weak Stackelberg values of $(PBQI)$ in finite dimensional spaces, and in [54] for a complete investigation in infinite dimensional spaces of convergence properties of approximate solutions and values of $(OBQI)$ problems, there called *semi-quasivariational optimistic bilevel problems*.

3.3.4 Quasi-equilibrium constraints

Let h be a real-valued function defined in $X \times Y \times Y$ and let T be a set-valued map from $X \times Y$ to Y with nonempty values. For any $x \in X$, the quasi-equilibrium problem (QE_x) (also called *quasi-variational problem* in [59]) consists in finding

$$y \in Y \text{ such that } y \in T(x, y) \text{ and } h(x, y, w) \leq 0 \quad \forall w \in T(x, y).$$

The class of such problems encompasses several problems arising in optimization, in game theory and in variational analysis, as illustrated in [59].

In [49], the concept of viscosity solutions relative to an inner regularization class for the family $\{(QE_x), x \in X\}$ was introduced and it was proved that the map \tilde{Q}^ε defined by

$$\tilde{Q}^\varepsilon(x) = \{y \in Y : d(y, T(x, y)) < \varepsilon \text{ and } h(x, y, w) < \varepsilon \quad \forall w \in T(x, y)\}$$

generate an inner regularization class under suitable assumptions and that the *large* version of \tilde{Q}^ε may fail to be an inner regularization under the same hypotheses. The results were established in infinite dimensional Banach spaces, so it had been necessary to balance the use of weak and strong convergences in the assumptions, but, in finite dimensional spaces, the statements can be simplified and made more readable.

4 Regularizing the follower's payoff function in Stackelberg games

The general non-single-valuedness of the follower's best reply correspondence M brings out the difficulties in the numerical approximation of problems (PB) and (OB) , as mentioned in paragraph C of Section 2, as well as the need to define constructive methods for selecting an SPNE in Stackelberg games, as emphasized in paragraph E of Section 2. For these reasons and possibly also for behavioural motivations (see, for example, [22]), regularization methods involving the follower's payoff function have been introduced. Such methods allow to construct sequences of perturbed problems where the solution of the second-stage problem is unique by exploiting well-known regularization techniques in convex optimization. For the sake of brevity, we will present only two approaches for regularizing the follower's payoff function: the first one is based on the Tikhonov regularization, the second one on the proximal regularization.

In this section we assume that the leader's and follower's action sets X and Y are

subsets of the Euclidean spaces \mathbb{X} and \mathbb{Y} , respectively, and we denote by $\|\cdot\|_{\mathbb{X}}$ and $\|\cdot\|_{\mathbb{Y}}$ the norm of \mathbb{X} and \mathbb{Y} , respectively. Furthermore, for the sake of brevity, both leader's and follower's constraints are assumed to be constant.

4.1 Regularizing the follower's payoff function via Tikhonov method

Let us recall preliminarily the Tikhonov regularization method for the approximation of solutions to convex optimization problems. Then, we illustrate how it has been employed to regularize bilevel optimization problems (PB) or (OB) and to select SPNEs.

4.1.1 Tikhonov regularization in convex optimization

Assume we deal with the minimization problem

$$(P): \min_{a \in A} J(a),$$

where J is a real-valued function defined on a subset A of a Euclidean space \mathbb{A} with norm $\|\cdot\|_{\mathbb{A}}$. In [107] Tikhonov introduced, in an optimal control framework, the following perturbed minimization problems

$$(P)_{\lambda_n}^{\mathcal{T}}: \min_{a \in A} \left(J(a) + \frac{1}{2\lambda_n} \|a\|_{\mathbb{A}}^2 \right),$$

where $n \in \mathbb{N}$ and $\lambda_n > 0$, and he proved the connections between the solutions to problem $(P)_{\lambda_n}^{\mathcal{T}}$ and the solutions to problem (P) , which are recalled in the next well-known result.

Theorem 4.1. *Assume that the set A is compact and convex, the function J is lower semicontinuous and convex over A and $\lim_{n \rightarrow +\infty} \lambda_n = +\infty$.*

Then, problem $(P)_{\lambda_n}^{\mathcal{T}}$ has a unique solution $\bar{a}_n^{\mathcal{T}} \in A$, for any $n \in \mathbb{N}$, and the sequence $(\bar{a}_n^{\mathcal{T}})_n$ is convergent to the minimum norm element of the set of the minimizers of J over A , so

$$\left\| \lim_{n \rightarrow +\infty} \bar{a}_n^{\mathcal{T}} \right\|_{\mathbb{A}} = \inf_{a \in V} \|a\|_{\mathbb{A}},$$

where $V = \text{Arg min}_{z \in A} J(z)$.

Therefore, the Tikhonov regularization method allows to approach a solution of a convex minimization problem by constructing a sequence of perturbed problems, in general better-behaved than the original problem, that have a unique solution. Furthermore, the limit of the sequence generated accordingly is uniquely characterized in the set of solutions to the original problem.

We mention that, afterwards, Tikhonov regularization has been broadly exploited for finding Nash equilibria in one-stage games where players move simultaneously: see [7, 47] for zero-sum games (with applications also to differential games) and [93, Section 4], [43, 25] for general N -players games.

4.1.2 Tikhonov regularization of the follower's payoff function in Stackelberg games

Let $\Gamma = (X, Y, L, F)$ be a Stackelberg game. The first attempt to regularize the follower's payoff function by employing Tikhonov regularization is due to Loridan and Morgan in [78]. Having in mind to approximate problem (PB) , for any $x \in X$ they considered the following Tikhonov-regularized second-stage problems

$$(P_x)_{\lambda_n}^T : \min_{y \in Y} \left(F(x, y) + \frac{1}{2\lambda_n} \|y\|_Y^2 \right),$$

where $n \in \mathbb{N}$ and $\lambda_n > 0$. The next result states preliminarily properties about the connections between the perturbed problems $(P_x)_{\lambda_n}^T$ and problem (P_x) as n goes to $+\infty$.

Proposition 4.1. [78, Prop. 3.1] *Assume that the set Y is compact and convex, the function F is continuous over $X \times Y$ and $\lim_{n \rightarrow +\infty} \lambda_n = +\infty$.*

Then, for any $x \in X$ and any sequence $(x_n)_n$ converging to x in X , we have

- $\lim_{n \rightarrow +\infty} \inf_{y \in Y} \left(F(x_n, y) + \frac{1}{2\lambda_n} \|y\|_Y^2 \right) = \inf_{y \in Y} F(x, y);$
- $\text{Lim sup}_n \text{Arg min}_{y \in Y} \left(F(x_n, y) + \frac{1}{2\lambda_n} \|y\|_Y^2 \right) \subseteq M(x).$

Remark 4.1 We highlight that the sequence of functions generated by Tikhonov-regularizing the follower's payoff function satisfies assumptions $(F1)$, $(F2)$ and $(\tilde{F}2)$ in [75, Section 4], therefore additional results involving ε -solutions and strict ε -solutions hold by applying propositions 4.2, 4.4 and 6.1 in [75]. For instance, the inclusion stated in Proposition 4.1 above is still valid even when considering ε -argmin maps.

More specific results concerning problems $(P_x)_{\lambda_n}^T$ and the connections with (P_x) are achieved by adding a convexity assumption, usual in a Tikhonov regularization framework.

Proposition 4.2. [78, Prop. 3.2 and 3.4] *Assume that the function $F(x, \cdot)$ is convex over Y for any $x \in X$ and hypotheses of Proposition 4.1 are satisfied.*

Then, for any $x \in X$,

- *problem $(P_x)_{\lambda_n}^T$ has a unique solution $\bar{\varphi}_n^T(x) \in Y$ for any $n \in \mathbb{N}$, that is*

$$\{\bar{\varphi}_n^T(x)\} = \text{Arg min}_{y \in Y} \left(F(x, y) + \frac{1}{2\lambda_n} \|y\|_Y^2 \right); \quad (6)$$

- $\lim_{n \rightarrow +\infty} \bar{\varphi}_n^T(x) = \hat{\varphi}(x)$, where $\hat{\varphi}(x)$ is the minimum norm element of the set $M(x)$, that is

$$\{\hat{\varphi}(x)\} = \text{Arg min}_{y \in M(x)} \|y\|_Y. \quad (7)$$

Furthermore, for any $n \in \mathbb{N}$ and any sequence $(x_k)_k$ converging to x in X , we have

$$\lim_{k \rightarrow +\infty} \bar{\varphi}_n^T(x_k) = \bar{\varphi}_n^T(x) \quad \text{and} \quad \lim_{k \rightarrow +\infty} F(x_k, \bar{\varphi}_n^T(x_k)) = F(x, \bar{\varphi}_n^T(x)).$$

Remark 4.2 Under the assumptions of Proposition 4.2 additional results regarding ε -solutions have been proved in [78, Proposition 3.3] for the Tikhonov-regularized problem $(P_x)_{\lambda_n}^T$.

We point out that, in general, the sequence $(\bar{\varphi}_n^T)_n$ is not continuously convergent to $\hat{\varphi}$ (for the definition of continuous convergence see Proposition 2.1), as shown in [78, Remark 3.1] and in the example below.

Example 4.1 Consider $X = Y = [-1, 1]$ and $F(x, y) = -xy$. The follower's best reply correspondence M is given in (1). Choosing $\lambda_n = n$, for any $x \in X$ we obtain

$$\bar{\varphi}_n^T(x) = \begin{cases} -1, & \text{if } x \in [-1, -1/n[\\ nx, & \text{if } x \in [-1/n, 1/n] \\ 1, & \text{if } x \in]1/n, 1] \end{cases} \quad \text{and} \quad \hat{\varphi}(x) = \begin{cases} -1, & \text{if } x \in [-1, 0[\\ 0, & \text{if } x = 0 \\ 1, & \text{if } x \in]0, 1]. \end{cases}$$

Therefore, the sequence of functions $(\bar{\varphi}_n^T)_n$ is not continuously convergent to $\hat{\varphi}$, as the function $\hat{\varphi}$ is not continuous at $x = 0$.

Now, let us consider the following Tikhonov-regularized bilevel problems:

$$(SP)_{\lambda_n}^T \begin{cases} \min_{x \in X} L(x, \bar{\varphi}_n^T(x)) \\ \text{where } \bar{\varphi}_n^T(x) \text{ is the solution to } (P_x)_{\lambda_n}^T. \end{cases}$$

Provided that assumptions of Proposition 4.2 are satisfied, $(SP)_{\lambda_n}^T$ is a Stackelberg problem for any $n \in \mathbb{N}$ since the solution to the second-stage problem is unique. Two questions arise:

- 1) Are there solutions to $(SP)_{\lambda_n}^T$?
- 2) What happens when $n \rightarrow +\infty$?

The next proposition provides the answer to the first question.

Proposition 4.3. [78, Prop. 3.5] *Assume that the set X is compact, Y is compact and convex, the functions L and F are continuous over $X \times Y$ and the function $F(x, \cdot)$ is convex over Y for any $x \in X$.*

Then, the Tikhonov-regularized Stackelberg problem $(SP)_{\lambda_n}^T$ has at least one solution, for any $n \in \mathbb{N}$.

As regards to investigation on the connections between $(SP)_{\lambda_n}^T$ and (PB) as n goes to $+\infty$, the second question is addressed in the following result.

Proposition 4.4. [78, Prop. 3.6] *Assume that $\lim_{n \rightarrow +\infty} \lambda_n = +\infty$ and hypotheses of Proposition 4.3 are satisfied. Denote with \bar{x}_n^T a solution to $(SP)_{\lambda_n}^T$ for any $n \in \mathbb{N}$.*

Then, any convergent subsequence of the sequence $(\bar{x}_n^T, \bar{\varphi}_n^T(\bar{x}_n^T))_n$ in $X \times Y$ has a limit (\bar{x}, \bar{y}) which satisfies

$$L(\bar{x}, \bar{y}) \leq w \quad \text{and} \quad \bar{y} \in M(\bar{x}), \tag{8}$$

where w is the security value of (PB) defined in Definition 1.2.

Definition 4.1 [69] An action profile $(\bar{x}, \bar{y}) \in X \times Y$ satisfying (8) is a *lower Stackelberg equilibrium pair* for problem (PB) .

The lower Stackelberg equilibrium pair solution concept was introduced in [69, Remark 5.3] in the context of approximation of (PB) via ε -solutions. Note that a result analogous to Proposition 4.4 has been obtained in [74] for the *least-norm regularization* of the second-stage problem of (PB) . In fact such a regularization, which is an adaptation of the method introduced in [104] and which involves the regularization both of the follower's optimal reaction set and of the follower's payoff function, generates sequences of action profiles whose limit points are lower Stackelberg equilibrium pairs, as proved in propositions 2.6 and 3.5 in [74].

We observe that if (\bar{x}, \bar{y}) is a weak Stackelberg equilibrium of (PB) , then (\bar{x}, \bar{y}) is a lower Stackelberg equilibrium pair. The converse is not true, in general, as illustrated in the following example.

Example 4.2 Consider $X = [1/2, 2]$, $Y = [-1, 1]$,

$$L(x, y) = -x - y \quad \text{and} \quad F(x, y) = \begin{cases} 0, & \text{if } x \in [1/2, 1[\\ (x - 1)y, & \text{if } x \in [1, 2]. \end{cases}$$

The pair $(1, 1)$ is a lower Stackelberg equilibrium, whereas the unique weak Stackelberg equilibrium is $(2, -1)$.

Remark 4.3 It is worth to emphasize that, in general, the sequence $(\bar{x}_n^T)_n$, where \bar{x}_n^T is a solution to $(SP)_{\lambda_n}^T$, converges neither to a weak Stackelberg (or pessimistic) solution of (PB) nor to a strong Stackelberg (or optimistic) solution of (OB) ; analogously, the sequence $(\bar{x}_n^T, \bar{\varphi}_n^T(\bar{x}_n^T))_n$ converges, in general, neither to a weak Stackelberg equilibrium nor to a strong Stackelberg equilibrium, as illustrated in the next example also used in [93].

Example 4.3 Consider $X = [-1/2, 1/2]$, $Y = [-1, 1]$,

$$L(x, y) = -x - y \quad \text{and} \quad F(x, y) = \begin{cases} (x + 1/4)y, & \text{if } x \in [-1/2, -1/4[\\ 0, & \text{if } x \in [-1/4, 1/4] \\ (x - 1/4)y, & \text{if } x \in]1/4, 1/2]. \end{cases}$$

The sequences $(\bar{x}_n^T)_n$ and $(\bar{x}_n^T, \bar{\varphi}_n^T(\bar{x}_n^T))_n$ converge to $-1/4$ and $(-1/4, 1)$, respectively. Instead, the strong Stackelberg solution is $1/4$ and the strong Stackelberg equilibrium is $(1/4, 1)$, whereas the weak Stackelberg solution and the weak Stackelberg equilibrium do not exist.

Afterwards, the Tikhonov regularization approach illustrated up to now has been employed in [26] where, by requiring stronger assumptions on the payoff functions, fur-

ther regularity properties for the Tikhonov-regularized second-stage problem $(P_x)_{\lambda_n}^{\mathcal{T}}$ have been shown. Moreover, an algorithm have been designed for the solutions of the Tikhonov-regularized Stackelberg problem $(SP)_{\lambda_n}^{\mathcal{T}}$ as n goes to $+\infty$.

We mention that a regularization approach in the same spirit of $(P_x)_{\lambda_n}^{\mathcal{T}}$ consists in considering the following perturbed second-stage problems

$$\min_{y \in Y} \left(F(x, y) + \frac{1}{2\lambda_n} L(x, y) \right),$$

which have a unique solution for any $n \in \mathbb{N}$, provided that the function $L(x, \cdot)$ is strongly convex on Y for any $x \in X$. In [32] such a Tikhonov-like approach has been investigated for problem (OB) , in order to circumvent the non-uniqueness of the solutions to the second-stage problem, and an algorithm has been proposed. Further discussions on this kind of regularization can be found in [28, Subsection 7.3.2].

We highlight that the idea of using the leader's payoff function in the regularization of the second-stage problem goes back to Molodtsov in [88]: he introduced a "mixed" regularization method, by combining both the regularization of the follower's reaction set and the regularization of the follower's payoff function, which allows to approximate problem (PB) via a sequence of strong Stackelberg problems. Then, more general properties of the *Molodtsov regularization* have been obtained in [79], where perturbations of the data of problem (PB) have been also considered.

4.1.3 Selection of SPNEs via Tikhonov regularization

A Stackelberg game where the follower's best reply correspondence M is not single-valued could have infinitely many SPNEs, as illustrated in Example 2.4. Therefore, let us focus on how restricting the number of SPNEs or, better yet, picking just one. The first attempt to select an SPNE via a constructive approach that allows to overcome the non-single-valuedness of M is due to Morgan and Patrone in [93], where the Tikhonov regularization approach presented in Subsection 4.1.2 was exploited.

Let $\Gamma = (X, Y, L, F)$ be a Stackelberg game and consider the following Tikhonov-regularized Stackelberg games

$$\Gamma_{\lambda_n}^{\mathcal{T}} = (X, Y, L, F_{\lambda_n}^{\mathcal{T}}),$$

where $F_{\lambda_n}^{\mathcal{T}} : X \times Y \rightarrow \mathbb{R}$ is defined on $X \times Y$ by

$$F_{\lambda_n}^{\mathcal{T}}(x, y) = F(x, y) + \frac{1}{2\lambda_n} \|y\|_Y^2,$$

for any $n \in \mathbb{N}$, i.e. $\Gamma_{\lambda_n}^{\mathcal{T}}$ is the game obtained from Γ by replacing the follower's payoff function F with the objective function of the Tikhonov-regularized problem $(P_x)_{\lambda_n}^{\mathcal{T}}$. Propositions 4.2 and 4.3 imply the following properties for $\Gamma_{\lambda_n}^{\mathcal{T}}$.

Corollary 4.1. *Under the assumptions of Proposition 4.3 we have, for any $n \in \mathbb{N}$,*

- the follower's best reply correspondence in $\Gamma_{\lambda_n}^T$ is single-valued;
- the game $\Gamma_{\lambda_n}^T$ has at least one SPNE: the strategy profile $(\bar{x}_n^T, \bar{\varphi}_n^T(\cdot))$, where $\bar{\varphi}_n^T: X \rightarrow Y$ is defined in (6) and $\bar{x}_n^T \in X$ is a solution to $(SP)_{\lambda_n}^T$, is an SPNE of $\Gamma_{\lambda_n}^T$.

Moreover, the sequence of strategies $(\bar{\varphi}_n^T)_n$ is pointwise convergent to the function $\hat{\varphi}$ defined in (7).

It is natural to ask if the limit of the sequence of SPNEs $(\bar{x}_n^T, \bar{\varphi}_n^T(\cdot))_n$, associated to the sequence of perturbed Stackelberg games $(\Gamma_{\lambda_n}^T)_n$, is an SPNE of the original game Γ . The answer is negative, in general, as displayed in the example below.

Example 4.4 Consider X , Y , L and F defined as in Example 2.1. The follower's best reply correspondence M is given in (1). Choosing $\lambda_n = n$, the SPNE of $\Gamma_{\lambda_n}^T$ is $(\bar{x}_n^T, \bar{\varphi}_n^T(\cdot))$ where $\bar{x}_n^T = -1/n$ and the function $\bar{\varphi}_n^T$ is derived in Example 4.1; moreover

$$\lim_{n \rightarrow +\infty} \bar{x}_n^T = 0 \quad \text{and} \quad \lim_{n \rightarrow +\infty} \bar{\varphi}_n^T(x) = \hat{\varphi}(x) \text{ for any } x \in X,$$

where the function $\hat{\varphi}$ is derived in Example 4.1. Regarding the strategy profile $(0, \hat{\varphi}(\cdot))$, we have

- $\hat{\varphi}(x) \in M(x)$ for any $x \in X$, so condition (SG1) in Definition 1.5 is satisfied,
- $0 \notin \text{Arg min}_{x \in X} L(x, \hat{\varphi}(x)) = \emptyset$, so condition (SG2) in Definition 1.5 does not hold,

hence, $(0, \hat{\varphi}(\cdot))$ is not an SPNE of Γ .

In order to achieve an SPNE of Γ we need to take into account the limit of the sequence of action profiles $(\bar{x}_n^T, \bar{\varphi}_n^T(\bar{x}_n^T))_n$.

Theorem 4.2. [93, Th. 3.1] *Assume that $\lim_{n \rightarrow +\infty} \lambda_n = +\infty$ and hypotheses of Proposition 4.3 are satisfied. Denote with $(\bar{x}_n^T, \bar{\varphi}_n^T(\cdot))$ an SPNE of $\Gamma_{\lambda_n}^T$, as defined in Corollary 4.1, for any $n \in \mathbb{N}$.*

If the sequence $(\bar{x}_n^T, \bar{\varphi}_n^T(\bar{x}_n^T))_n$ converges to (\bar{x}^T, \bar{y}^T) in $X \times Y$, then the strategy profile $(\bar{x}^T, \tilde{\varphi}^T(\cdot))$ where

$$\tilde{\varphi}^T(x) = \begin{cases} \bar{y}^T, & \text{if } x = \bar{x}^T \\ \hat{\varphi}(x), & \text{if } x \neq \bar{x}^T \end{cases}$$

with $\hat{\varphi}(x)$ is defined in (7), is an SPNE of Γ .

Remark 4.4 Coming back to Example 4.4, we have that $(\bar{x}_n^T, \bar{\varphi}_n^T(\bar{x}_n^T)) = (-1/n, -1)$ for any $n \in \mathbb{N}$ and $(\bar{x}^T, \bar{y}^T) = (0, -1)$. Therefore, the SPNE selected according to Theorem 4.2 is $(0, \tilde{\varphi}^T(\cdot))$ where the function $\tilde{\varphi}^T: X \rightarrow Y$ is defined on X by

$$\tilde{\varphi}^T(x) = \begin{cases} -1, & \text{if } x \in [-1, 0] \\ 1, & \text{if } x \in]0, 1]. \end{cases}$$

In fact, $\tilde{\varphi}^T(x) \in M(x)$ for any $x \in X$ and $0 \in \text{Arg min}_{x \in X} L(x, \tilde{\varphi}^T(x))$.

The following example shows that the SPNE selected according to Theorem 4.2 does not coincide, in general, with the SPNEs induced by a weak or a strong Stackelberg solution, as defined in Definition 1.6.

Example 4.5 Consider X , Y , L and F defined as in Example 1.7. The sequence $(\bar{x}_n^T, \bar{\varphi}_n^T(\bar{x}_n^T))_n$ converges to $(7/4, 0)$, so the SPNE selected according to Theorem 4.2 is the strategy profile $(7/4, \tilde{\varphi}^T(\cdot))$ where

$$\tilde{\varphi}^T(x) = \begin{cases} 1, & \text{if } x \in [-2, -7/4[\\ 0, & \text{if } x \in [-7/4, 7/4] \\ -1, & \text{if } x \in]7/4, 2]. \end{cases}$$

Such an SPNE is different from the SPNEs induced by weak and strong Stackelberg solutions, which are given in Example 1.7. Moreover, since Γ has multiple SPNEs and only one of which is obtained via the method displayed above, the selection method for SPNEs designed via Tikhonov regularization is effective.

The Tikhonov approach to select SPNEs presented in this subsection has been extended also to the class of *one-leader two-follower Stackelberg games* in [93, Section 4]. In this case the two followers, after having observed the action x chosen by the leader, face a parametric two-player Nash equilibrium problem (NE_x) (see Subsection 3.3.2). By Tikhonov-regularizing the followers' payoff functions in problem (NE_x) , a sequence of perturbed one-leader two-follower Stackelberg games has been defined where the solution to the Nash equilibrium problem in the second-stage is unique. This allows to select an SPNE similarly to Theorem 4.2, as proved in theorems 4.1 and 4.2 in [93].

4.2 Regularizing the follower's payoff function via proximal method

Let us recall firstly the proximal regularization method for approximating the solutions of a convex optimization problem. Then, we show how it can be employed to regularize bilevel optimization problems (PB) or (OB) and to select SPNEs involving behavioural motivations.

4.2.1 Proximal-point algorithm in convex optimization

Assume we deal with the minimization problem

$$(P): \min_{a \in A} J(a),$$

where J is a real-valued function defined on a subset A of a Euclidean space \mathbb{A} with norm $\|\cdot\|_{\mathbb{A}}$. The algorithm defined below has been introduced by Martinet in [86] for regularizing variational inequalities and by Rockafellar in [99] for finding zeros of maximal monotone operators.

Proximal Point Algorithm (PA)

Fix an initial point $\bar{a}_0 \in A$ and define for any $n \in \mathbb{N}$

$$\{\bar{a}_n^{\mathcal{P}}\} = \underset{a \in A}{\text{Arg min}} \left(J(a) + \frac{1}{2\gamma_n} \|a - \bar{a}_{n-1}^{\mathcal{P}}\|_{\mathbb{A}}^2 \right),$$

where $\gamma_n > 0$ for any $n \in \mathbb{N}$.

The well-definedness of algorithm (PA) and its convergence properties are recalled in the following well-known result stated in [99].

Theorem 4.3. *Assume that the set A is compact and convex, the function J is lower semicontinuous and convex over A and $\sum_{n=1}^{+\infty} \gamma_n = +\infty$.*

Then, algorithm (PA) is well-defined and the sequence $(\bar{a}_n^{\mathcal{P}})_n$ is convergent to a solution of problem (P), that is

$$\lim_{n \rightarrow +\infty} \bar{a}_n^{\mathcal{P}} \in \underset{a \in A}{\text{Arg min}} J(a).$$

Remark 4.5 We point out that algorithm (PA) and its denomination rely on the *Moreau-Yosida regularization* of J with parameter $\gamma > 0$, which is the function $J_\gamma^{\mathcal{P}} : A \rightarrow [-\infty, +\infty]$ defined on A by

$$J_\gamma^{\mathcal{P}}(a) = \inf_{x \in A} \left(J(x) + \frac{1}{2\gamma} \|x - a\|_{\mathbb{A}}^2 \right),$$

introduced previously by Moreau in [90], and also called *Moreau envelope* or *proximal approximation*.

Therefore, the proximal point algorithm allows to approximate a solution of a convex minimization problem by constructing recursively a sequence of proximal-perturbed problems, namely

$$(P)_{\gamma_n}^{\mathcal{P}} : \min_{a \in A} \left(J(a) + \frac{1}{2\gamma_n} \|a - \bar{a}_{n-1}^{\mathcal{P}}\|_{\mathbb{A}}^2 \right),$$

in general better-behaved than the original problem and that have a unique solution. Looking at the differences with respect to the Tikhonov approach, we have

- the proximal-regularized problems $(P)_{\gamma_n}^{\mathcal{P}}$ are recursively defined, differently from the Tikhonov-regularized problems $(P)_{\lambda_n}^{\mathcal{T}}$ in Subsection 4.1.1;
- the limit of the sequence generated by (PA) is “just” a minimizer of J over A , while in the Tikhonov approach the limit point is characterized as the minimum norm element in the set of minimizers of J over A ;
- as regards to the assumptions ensuring the convergence of algorithm (PA) and of Tikhonov method, the hypothesis on the sequence of parameters $(\gamma_n)_n$ in Theorem 4.3 is weaker than the corresponding hypothesis on $(\lambda_n)_n$ in Theorem 4.1.

In a one-stage games framework, we mention that algorithm (PA) has been used in [86, 99] also for approximating saddle-points in zero-sum games and in [38] in general N -player games. More recently, in [5] an alternating proximal minimization algorithm has been proposed for two-player weighted potential games and in [17] Moreau-Yosida regularization has been exploited to define a new equilibrium concept to model strategic uncertainty among players.

4.2.2 Proximal regularization of the follower's payoff function in Stackelberg games

Let $\Gamma = (X, Y, L, F)$ be a Stackelberg game. We first display some preliminary results on employing proximal regularization only in the second-stage problem similarly as presented in Subsection 4.1.2 for the Tikhonov regularization. We do this in order to make easier to illustrate the methodology for selecting an SPNE of Γ we will use in the next subsection which involves the regularization of both players' payoff functions. Consider the following approach: fixed an initial point $\bar{y}_0 \in Y$, define for any $n \in \mathbb{N}$

$$\{\bar{\varphi}_n^{\mathcal{P}}(x)\} = \text{Arg min}_{y \in Y} \left(F(x, y) + \frac{1}{2\gamma_n} \|y - \bar{\varphi}_{n-1}^{\mathcal{P}}(x)\|_Y^2 \right) \quad \text{for any } x \in X, \quad (9)$$

where $\gamma_n > 0$ for any $n \in \mathbb{N}$ and $\bar{\varphi}_0^{\mathcal{P}}(x) = \bar{y}_0$ for any $x \in X$.

Remark 4.6 Assume that Y is compact and convex, the function F is continuous over $X \times Y$, the function $F(x, \cdot)$ is convex over Y for any $x \in X$ and $\sum_{n=1}^{+\infty} \gamma_n = +\infty$. From Theorem 4.3 it straightforwardly follows that, for any $x \in X$, the sequence $(\bar{\varphi}_n^{\mathcal{P}}(x))_n$ is well-defined and $\lim_{n \rightarrow +\infty} \bar{\varphi}_n^{\mathcal{P}}(x) \in M(x)$.

Therefore, we can approach the follower's problem (P_x) via a sequence of proximal-regularized second-stage problems recursively defined by

$$(P_x)_{\gamma_n}^{\mathcal{P}} : \text{Arg min}_{y \in Y} \left(F(x, y) + \frac{1}{2\gamma_n} \|y - \bar{\varphi}_{n-1}^{\mathcal{P}}(x)\|_Y^2 \right),$$

and consequently we can define the following proximal-regularized bilevel problems:

$$(SP)_{\gamma_n}^{\mathcal{P}} \begin{cases} \min_{x \in X} L(x, \bar{\varphi}_n^{\mathcal{P}}(x)) \\ \text{where } \bar{\varphi}_n^{\mathcal{P}}(x) \text{ is the solution to } (P_x)_{\gamma_n}^{\mathcal{P}} \end{cases}$$

which come out to be Stackelberg problems, since the solution to the second-stage problem of $(SP)_{\gamma_n}^{\mathcal{P}}$ is unique for any $n \in \mathbb{N}$. Analogously to Subsection 4.1.2, we are interested in two questions.

- 1) Are there solutions to $(SP)_{\gamma_n}^{\mathcal{P}}$?
- 2) What happens when $n \rightarrow +\infty$?

By means of the maximum theorem, the function $\bar{\varphi}_n^{\mathcal{P}} : X \rightarrow Y$ defined in (9) is continuous over X , so the next proposition provides the answer to the first question.

Proposition 4.5. *Assume that the set X is compact, Y is compact and convex, the functions L and F are continuous over $X \times Y$ and the function $F(x, \cdot)$ is convex over Y for any $x \in X$.*

Then, the proximal-regularized Stackelberg problem $(SP)_{\gamma_n}^{\mathcal{P}}$ has at least one solution, for any $n \in \mathbb{N}$.

Concerning the second question, the limit of the sequence generated by the solutions to $(SP)_{\gamma_n}^{\mathcal{P}}$ for any $n \in \mathbb{N}$ has no connections, in general, either with weak Stackelberg (or pessimistic) solutions to (PB) or with strong Stackelberg (or optimistic) solutions to (OB) , as illustrated in the following example also used in [22].

Example 4.6 Consider $X = [1/2, 2]$, $Y = [-1, 1]$,

$$L(x, y) = x + y \quad \text{and} \quad F(x, y) = \begin{cases} 0, & \text{if } x \in [1/2, 1[\\ (x-1)y, & \text{if } x \in [1, 2]. \end{cases}$$

Choosing $\bar{y}_0 = 1$ and $\gamma_n = n$, we obtain

$$\bar{\varphi}_n^{\mathcal{P}}(x) = \begin{cases} 1, & \text{if } x \in [1/2, 1] \\ c_n - c_n x + 1, & \text{if } x \in]1, 1 + 2/c_n] \\ -1, & \text{if } x \in]1 + 2/c_n, 2], \end{cases} \quad \text{where} \quad \begin{cases} c_1 = 1 \\ c_{n+1} = c_n + n + 1 \end{cases} \quad \text{for any } n \geq 1.$$

Hence, the solution to $(SP)_{\gamma_n}^{\mathcal{P}}$ is

$$\bar{u}_n^{\mathcal{P}} = \begin{cases} 1/2, & \text{if } n = 1 \\ 1 + 2/c_n, & \text{if } n \geq 2. \end{cases}$$

Since $\lim_{n \rightarrow +\infty} c_n = +\infty$, the sequence $(\bar{u}_n^{\mathcal{P}})_n$ converges to 1, which is neither a weak nor a strong Stackelberg solution. In fact, a solution to (PB) does not exist, whereas the solution to (OB) is $1/2$. Consequently, even the sequence $(\bar{u}_n^{\mathcal{P}}, \bar{\varphi}_n^{\mathcal{P}}(\bar{u}_n^{\mathcal{P}}))_n$ does not converge either to a weak or to a strong Stackelberg equilibrium.

Example above shows also that, in general, the sequence of functions $(\bar{\varphi}_n^{\mathcal{P}})_n$ is not continuously convergent. In fact, coming back to Example 4.6, it is sufficient to note that the pointwise limit of $(\bar{\varphi}_n^{\mathcal{P}})_n$ defined by

$$\lim_{n \rightarrow +\infty} \bar{\varphi}_n^{\mathcal{P}}(x) = \begin{cases} 1, & \text{if } x \in [1/2, 1] \\ -1, & \text{if } x \in]1, 2], \end{cases}$$

is not a continuous function.

We mention that methods involving proximal regularization have been recently proposed in [100, 102] for solving the *simple bilevel optimization problem*

$$(SBP) \quad \begin{cases} \min_{x \in E} l(x) \\ \text{where } E = \text{Arg min}_{x \in \Delta} f(x), \end{cases}$$

with Δ convex subset of \mathbb{R}^n and l and f real-valued convex functions defined on \mathbb{R}^n . Such a problem, firstly studied in [106] and labelled as “simple” in [29], involves a non-parametric second-stage problem differently from the general definitions of (PB) and (OB) , and its solutions clearly coincide with the weak and strong Stackelberg solutions. Consequently, problem (SBP) does not entail the same kind of inherent difficulties that bilevel optimization problems (PB) and (OB) exhibit regarding the regularization issue, as illustrated in this subsection and in Subsection 4.1.2.

4.2.3 Selection of SPNEs via proximal regularization

Let us present now how proximal regularization has been employed in Stackelberg games in order to select SPNEs. Caruso, Ceparano and Morgan introduced in [22] a behavioural motivated constructive method which involves the proximal regularization both in the second-stage and in the first-stage problem (the additional regularization of the leader’s payoff function has been done in order to strongly motivate the behavioural interpretation of the method). Anyway, if the proximal regularization is carried out only in the second-stage problem, by arguing analogously as in the proofs in [22], a selection result for SPNEs can be proved even in such a case.

Let $\Gamma = (X, Y, L, F)$ be a Stackelberg game and consider the following procedure.

Stackelberg Game Proximal Procedure (SGP)

Fix an initial point $(\bar{x}_0, \bar{y}_0) \in X \times Y$ and define for any $n \in \mathbb{N}$

$$(\mathcal{A}_n) \quad \begin{cases} \{\bar{\varphi}_n^{\mathcal{P}}(x)\} = \text{Arg min}_{y \in Y} F_{\gamma_n}^{\mathcal{P}}(x, y), \text{ for any } x \in X \\ \bar{x}_n^{\mathcal{P}} \in \text{Arg min}_{x \in X} L_{\beta_n}^{\mathcal{P}}(x, \bar{\varphi}_n^{\mathcal{P}}(x)) \end{cases}$$

where for any $(x, y) \in X \times Y$

$$\begin{aligned} F_{\gamma_n}^{\mathcal{P}}(x, y) &= F(x, y) + \frac{1}{2\gamma_n} \|y - \bar{\varphi}_{n-1}^{\mathcal{P}}(x)\|_Y^2 \\ L_{\beta_n}^{\mathcal{P}}(x, y) &= L(x, y) + \frac{1}{2\beta_n} \|x - \bar{x}_{n-1}^{\mathcal{P}}\|_X^2, \end{aligned}$$

with $\gamma_n > 0$ for any $n \in \mathbb{N}$ and $\lim_{n \rightarrow +\infty} \gamma_n = +\infty$,

$\beta_n > 0$ for any $n \in \mathbb{N}$ and $\lim_{n \rightarrow +\infty} \beta_n = +\infty$,

and $\bar{\varphi}_0^{\mathcal{P}}(x) = \bar{y}_0$ for any $x \in X$.

Procedure (SGP) allows to construct recursively a sequence of strategies $(\bar{\varphi}_n^{\mathcal{P}}(\cdot))_n$, where $\bar{\varphi}_n^{\mathcal{P}}$ is a function from X to Y , and a sequence of leader’s actions $(\bar{x}_n^{\mathcal{P}})_n$, where $\bar{x}_n^{\mathcal{P}} \in X$. We point out that looking only at the follower’s stage in procedure (SGP) , one can recognize the same approach illustrated in Subsection 4.2.2 (for this reason here we kept unchanged the notation $\bar{\varphi}_n^{\mathcal{P}}$ used previously).

In the next proposition, the well-definedness of procedure (SGP) and its regularity and convergence properties are stated.

Proposition 4.6. [22, Prop. 1 and 2] *Under the assumptions of Proposition 4.5 we have*

- *procedure (SGP) is well-defined;*
- *for any sequence $(x_k)_k$ converging to x in X , we have*

$$\lim_{k \rightarrow +\infty} \bar{\varphi}_n^{\mathcal{P}}(x_k) = \bar{\varphi}_n^{\mathcal{P}}(x) \quad \text{for any } n \in \mathbb{N};$$

- *for any $x \in X$ the sequence $(\bar{\varphi}_n^{\mathcal{P}}(x))_n$ is convergent in Y to a solution of problem (P_x) .*

Thanks to the well-definedness of procedure (SGP), we can construct recursively a sequence of proximal-regularized Stackelberg games $(\Gamma_{v_n}^{\mathcal{P}})_n$ where

$$\Gamma_{v_n}^{\mathcal{P}} = (X, Y, L_n^{\mathcal{P}}, F_n^{\mathcal{P}})$$

is the game obtained from Γ by replacing the players' payoff functions L and F with the proximal-regularized functions $L_{\beta_n}^{\mathcal{P}}$ and $F_{\gamma_n}^{\mathcal{P}}$ and $v_n = (\gamma_n, \beta_n)$. Then, Proposition 4.6 shows in particular that for any $n \in \mathbb{N}$ the follower's best reply correspondence in $\Gamma_{v_n}^{\mathcal{P}}$ is single-valued and that the strategy profile $(\bar{x}_n^{\mathcal{P}}, \bar{\varphi}_n^{\mathcal{P}}(\cdot)) \in X \times Y^X$ generated at step (\mathcal{A}_n) of procedure (SGP) is an SPNE of $\Gamma_{v_n}^{\mathcal{P}}$ obtained as update of $(\bar{x}_{n-1}^{\mathcal{P}}, \bar{\varphi}_{n-1}^{\mathcal{P}}(\cdot))$, SPNE of $\Gamma_{v_{n-1}}^{\mathcal{P}}$.

Before stating the SPNE selection result, we highlight that procedure (SGP) has a behavioural interpretation linked to the costs that players face when deviating from their current actions, which is presented below.

Interpretation of the procedure At the generic step (\mathcal{A}_n) of procedure (SGP), the follower chooses his strategy $\bar{\varphi}_n^{\mathcal{P}}$ taking into account his previous strategy $\bar{\varphi}_{n-1}^{\mathcal{P}}$. In making such a choice, he finds an action that compromises between minimizing $F(x, \cdot)$ and being near to $\bar{\varphi}_{n-1}^{\mathcal{P}}(x)$, for any $x \in X$. The latter purpose is motivated according to an *anchoring effect*, explained in [5, Subsection 1.4], which is formulated by subtracting a quadratic slight cost to move that reflects the difficulty of changing the previous action. The coefficient γ_n is linked to the *per unit of distance cost to move* of the follower and it is related to the trade-off parameter between minimizing $F(x, \cdot)$ and minimizing the distance from $\bar{\varphi}_{n-1}^{\mathcal{P}}(x)$. Since the same arguments apply for the preceding steps until going up to step (\mathcal{A}_1) , it follows that $\bar{\varphi}_n^{\mathcal{P}}(x)$ as well as the limit of $\bar{\varphi}_n^{\mathcal{P}}(x)$ embed the willingness of being near to \bar{y}_0 . Analogous observations hold also for the leader, who chooses an action having in mind to be near to his previous choices, and therefore even with the purpose of being near to \bar{x}_0 .

By taking into account the limit of the sequence of action profiles $(\bar{x}_n^{\mathcal{P}}, \bar{\varphi}_n^{\mathcal{P}}(\bar{x}_n^{\mathcal{P}}))_n$, the following existence result of an SPNE's selection holds.

Theorem 4.4. [22, Th. 1] *Assume that hypotheses of Proposition 4.5 are satisfied and let $(\bar{x}_n^{\mathcal{P}}, \bar{\varphi}_n^{\mathcal{P}}(\cdot))_n$ be the sequence of strategy profiles generated by procedure (SGP).*

If the sequence $(\bar{x}_n^{\mathcal{P}}, \bar{\varphi}_n^{\mathcal{P}}(\bar{x}_n^{\mathcal{P}}))_n$ converges to $(\bar{x}^{\mathcal{P}}, \bar{y}^{\mathcal{P}})$ in $X \times Y$, then the strategy profile $(\bar{x}^{\mathcal{P}}, \tilde{\varphi}^{\mathcal{P}}(\cdot))$ where

$$\tilde{\varphi}^{\mathcal{P}}(x) = \begin{cases} \bar{y}^{\mathcal{P}}, & \text{if } x = \bar{x}^{\mathcal{P}} \\ \lim_{n \rightarrow +\infty} \bar{\varphi}_n^{\mathcal{P}}(x), & \text{if } x \neq \bar{x}^{\mathcal{P}}, \end{cases}$$

is an SPNE of Γ .

Theorem 4.4 points out that the sequence $(\bar{x}_n^{\mathcal{P}}, \bar{\varphi}_n^{\mathcal{P}}(\cdot))_n$ of SPNEs of the proximal-regularized games $(\Gamma_{v_n}^{\mathcal{P}})_n$ could not converge to an SPNE of Γ . This happens because the follower's strategy $\tilde{\varphi}^{\mathcal{P}}$ in the SPNE defined according to Theorem 4.4 could differ from the pointwise limit, whose existence is ensured by Proposition 4.6, of sequence $(\bar{\varphi}_n^{\mathcal{P}})_n$ in one point. In fact, denoted with $\bar{\varphi}: X \rightarrow Y$ such a pointwise limit, if the two limits

$$\lim_{n \rightarrow +\infty} \bar{\varphi}_n^{\mathcal{P}}(\bar{x}_n^{\mathcal{P}}) \quad \text{and} \quad \lim_{n \rightarrow +\infty} \bar{\varphi}_n^{\mathcal{P}}(\bar{x}^{\mathcal{P}}), \quad (10)$$

where $\bar{x}^{\mathcal{P}} = \lim_{n \rightarrow +\infty} \bar{x}_n^{\mathcal{P}}$, coincide, then $\tilde{\varphi}^{\mathcal{P}}(x) = \bar{\varphi}(x)$ for any $x \in X$ and the strategy profile $(\bar{x}^{\mathcal{P}}, \bar{\varphi})$ is an SPNE of Γ in light of Theorem 4.4. Instead, if the two limits in (10) do not coincide, then $\tilde{\varphi}^{\mathcal{P}}(\bar{x}^{\mathcal{P}}) \neq \bar{\varphi}(\bar{x}^{\mathcal{P}})$ and the strategy profile $(\bar{x}^{\mathcal{P}}, \bar{\varphi})$ could be not an SPNE of Γ , hence we need the follower's strategy $\tilde{\varphi}^{\mathcal{P}}$ as in statement of Theorem 4.4 in order to get an SPNE. Examples 2 and 3 in [22] illustrate the two cases described above: in the first one the two limits in (10) are equal, whereas, in the second one the two limits in (10) are different and the pair $(\bar{x}^{\mathcal{P}}, \bar{\varphi})$ comes out to be not an SPNE of the game. Such examples show also that, in general, the sequence $(\bar{\varphi}_n^{\mathcal{P}})_n$ is not continuously convergent to $\bar{\varphi}$.

As regards to the connections with other methods for selecting SPNEs, the following example shows that the selection method based on proximal regularization presented in this subsection, the selection method relying on Tikhonov regularization described in Subsection 4.1.3 and the way of selecting via weak and strong Stackelberg solutions illustrated in Remark 2.4 do not generate, in general, the same SPNE.

Example 4.7 [22, Ex. 3, 5 and 6] Consider X, Y, L and F defined as in Example 4.6. We have that

- the SPNE selected via proximal regularization is $(1, \tilde{\varphi}^{\mathcal{P}})$, where

$$\tilde{\varphi}^{\mathcal{P}}(x) = \begin{cases} 1, & \text{if } x \in [1/2, 1[\\ -1, & \text{if } x \in [1, 2]; \end{cases}$$

- the SPNE selected by using the Tikhonov regularization is $(1, \tilde{\varphi}^{\mathcal{T}})$, where

$$\tilde{\varphi}^{\mathcal{T}}(x) = \begin{cases} 0, & \text{if } x \in [1/2, 1[\\ -1, & \text{if } x \in [1, 2]; \end{cases}$$

- there are no SPNEs induced by a weak Stackelberg solution, as problem (PB) has no solutions;
- there exists one SPNE induced by the (unique) strong Stackelberg solution: the pair $(1/2, \bar{\psi})$ where $\bar{\psi}(x) = -1$ for any $x \in [1/2, 2]$.

Finally, we mention that convergence results involving “costs of change” and proximal methods in Stackelberg games have been recently investigated in [37, Section 5].

5 Conclusion

In this chapter we considered two-stage Stackelberg games, the natural environment for different kinds of mathematical problems that can arise depending on the leader’s information about the optimal responses of the follower. We discussed crucial issues related to such problems and we provided two approaches for managing them. The first one consists in regularizing the follower’s optimal reaction set via the introduction of appropriate solution concepts which allowed to obviate both the lack of existence and stability of the weak Stackelberg (or pessimistic bilevel) problem and the lack of stability of the strong Stackelberg (or optimistic bilevel) problem. The second one consists in regularizing the follower’s payoff function by employing the Tikhonov and the proximal regularizations which enabled both to overcome the non-single-valuedness of the follower’s best reply correspondence and to select subgame perfect Nash equilibria.

We note that the issues and the approaches (for facing such issues) we focused in this chapter are not the unique ones. For instance, the question of *sensitivity analysis* for problems (PB) and (OB) , has been recently investigated in [30, 31]. Moreover, beyond the two regularizing approaches examined in this chapter, one can construct classes of *mixed* regularization methods which exploit simultaneously both the idea of regularizing the follower’s optimal reaction set and of regularizing the follower’s payoff function; first approximation schemes of this type have been defined in [88, 74, 79].

For the sake of brevity, we dealt with Stackelberg games only in a static framework (that is any player has actions which do not evolve over time) where players have a single objective and just one leader acts in the first stage. However, such a Stackelberg game models has been extended in many directions:

- **Dynamic:** players’ actions are functions depending on time. First results on *dynamic Stackelberg games* can be found in [24, 105, 70, 20, 111, 10], while applications to economic models are described in [11, 33].
- **Multiobjective:** the follower has a vector-valued payoff function or solves vector (quasi-)variational inequalities. Methods for solving (optimistic) *multiobjective bilevel optimization problems* have been investigated in [35, 113]. Concerning

results on approximate solutions which could be useful in the analysis of Stackelberg games with second stage described by vector problems, we remind [72, 91, 62, 65]. Furthermore, a penalty scheme has been introduced to approach (optimistic) *semivectorial bilevel optimization problems* in [19] and, in a dynamic framework, existence results have been first obtained for the (optimistic and pessimistic) *semivectorial bilevel optimal control problems* in [18].

- **Multileader:** more than one leader acts in the first stage. For general *multi-leader-follower games* see, for example, [98, 42]. The situation where each follower can observe the action of one leader has been investigated in [97] for an economic model and in [23] for deriving a selection result. An increasing literature is also devoted to the class of *multi-leader-common-follower games* and to its applications to electricity markets, see for example [8, 9].

We emphasize that for the situations illustrated above it would be interesting to provide new concepts and results which derive from the analogous exploitation of the two types of regularization methods presented in the chapter.

References

- [1] A. Aboussoror, S. Adly, and F.E. Saissi. “Strong-weak nonlinear bilevel problems: existence of solutions in a sequential setting”. In: *Set-Valued Var. Anal.* 25 (2017), pp. 113–132. DOI: [10.1007/s11228-016-0369-4](https://doi.org/10.1007/s11228-016-0369-4).
- [2] A. Aboussoror and P. Loridan. “Existence of solutions to two-level optimization problems with nonunique lower-level solutions”. In: *J. Math. Anal. Appl.* 254 (2001), pp. 348–357. DOI: [10.1006/jmaa.2000.7001](https://doi.org/10.1006/jmaa.2000.7001).
- [3] A. Aboussoror and P. Loridan. “Strong-weak Stackelberg problems in finite dimensional spaces”. In: *Serdica Math. J.* 21 (1995), pp. 151–170.
- [4] A. Aboussoror and A. Mansouri. “Weak linear bilevel programming problems: existence of solutions via a penalty method”. In: *J. Math. Anal. Appl.* 304 (2005), pp. 399–408. DOI: [10.1016/j.jmaa.2004.09.033](https://doi.org/10.1016/j.jmaa.2004.09.033).
- [5] H. Attouch, P. Redont, and A. Soubeyran. “A new class of alternating proximal minimization algorithms with costs-to-move”. In: *SIAM J. Optim.* 18 (2007), pp. 1061–1081. DOI: [10.1137/060657248](https://doi.org/10.1137/060657248).
- [6] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Birkhäuser, Boston, 1990.
- [7] A. Auslender. “Recherche des points de selle d’une fonction”. In: *Symposium on Optimization*. Ed. by A.V. Balakrishnan, M. Contensou, B.F. de Veubeke, P. Krée, J.L. Lions, and N.N. Moiseev. Berlin, Heidelberg: Springer Berlin Heidelberg, 1970, pp. 37–52.
- [8] D. Aussel, P. Bendotti, and M. Pištěk. “Nash equilibrium in a pay-as-bid electricity market: Part 1 - existence and characterization”. In: *Optimization* 66 (2017), pp. 1013–1025. DOI: [10.1080/02331934.2016.1227981](https://doi.org/10.1080/02331934.2016.1227981).

- [9] D. Aussel, P. Bendotti, and M. Pištěk. “Nash equilibrium in a pay-as-bid electricity market Part 2 - best response of a producer”. In: *Optimization* 66 (2017), pp. 1027–1053. DOI: [10.1080/02331934.2016.1227982](https://doi.org/10.1080/02331934.2016.1227982).
- [10] T. Başar and G. Olsder. *Dynamic Noncooperative Game Theory*. Society for Industrial and Applied Mathematics, 1998.
- [11] A. Bagchi. *Stackelberg differential games in economic models*. Vol. 64. Lecture Notes in Control and Information Sciences. Springer-Verlag Berlin Heidelberg, 1984.
- [12] C. Baiocchi and A. Capelo. “Variational and Quasi-Variational Inequalities”. In: (1984).
- [13] B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer. *Non-Linear Parametric Optimization*. Birkhäuser Basel, 1983.
- [14] J. F. Bard and J. E. Falk. “An explicit solution to the multi-level programming problem”. In: *Comput. & Oper. Res.* 9 (1982), pp. 77–100. DOI: [10.1016/0305-0548\(82\)90007-7](https://doi.org/10.1016/0305-0548(82)90007-7).
- [15] J.F. Bard. “An algorithm for solving the general bilevel programming problem”. In: *Math. Oper. Res.* 8 (1983), pp. 260–272. DOI: [10.1287/moor.8.2.260](https://doi.org/10.1287/moor.8.2.260).
- [16] C. Berge. *Espaces Topologiques et Fonctions Multivoques*. [Translation: E.M. Patterson. *Topological Spaces*. Oliver & Boyd, 1963]. Paris: Dunod, 1959.
- [17] P. Bich. “Strategic uncertainty and equilibrium selection in discontinuous games”. In: *J. Econ. Theory* 183 (2019), pp. 786–822. DOI: [10.1016/j.jet.2019.08.001](https://doi.org/10.1016/j.jet.2019.08.001).
- [18] H. Bonnel and J. Morgan. “Semivectorial bilevel convex optimal control problems: existence results”. In: *SIAM J. Control Optim.* 50 (2012), pp. 3224–3241. DOI: [10.1137/100795450](https://doi.org/10.1137/100795450).
- [19] H. Bonnel and J. Morgan. “Semivectorial bilevel optimization problem: penalty approach”. In: *J. Optim. Theory Appl.* 131 (2006), pp. 365–382. DOI: [10.1007/s10957-006-9150-4](https://doi.org/10.1007/s10957-006-9150-4).
- [20] M. Breton, A. Alj, and A. Haurie. “Sequential Stackelberg equilibria in two-person games”. In: *J. Optim. Theory Appl.* 59 (1988), pp. 71–97. DOI: [10.1007/BF00939867](https://doi.org/10.1007/BF00939867).
- [21] D. Cao and L.C. Leung. “A partial cooperation model for non-unique linear two-level decision problems”. In: *Eur. J. Oper. Res.* 140 (2002), pp. 134–141. DOI: [10.1016/S0377-2217\(01\)00225-9](https://doi.org/10.1016/S0377-2217(01)00225-9).
- [22] F. Caruso, M.C. Ceparano, and J. Morgan. “Subgame perfect Nash equilibrium: a learning approach via costs to move”. In: *Dyn. Games Appl.* 9 (2019), pp. 416–432. DOI: [10.1007/s13235-018-0277-3](https://doi.org/10.1007/s13235-018-0277-3).
- [23] M.C. Ceparano and J. Morgan. “Equilibrium selection in multi-leader-follower games with vertical information”. In: *TOP* 25 (2017), pp. 526–543. DOI: [10.1007/s11750-017-0444-5](https://doi.org/10.1007/s11750-017-0444-5).
- [24] C. Chen and J.B. Cruz. “Stackelberg solution for two-person games with biased information patterns”. In: *IEEE Trans. Autom. Control* 17 (1972), pp. 791–798.
- [25] J.B. Clempner and A.S. Poznyak. “Computing the strong Nash equilibrium for Markov chains games”. In: *Appl. Math. Comput.* 265 (2015), pp. 911–927. DOI: [10.1016/j.amc.2015.06.005](https://doi.org/10.1016/j.amc.2015.06.005).

- [26] S. Dempe. “A bundle algorithm applied to bilevel programming problems with non-unique lower level solutions”. In: *Comput. Optim. Appl.* 15 (2000), pp. 145–166. DOI: [10.1023/A:1008735010803](https://doi.org/10.1023/A:1008735010803).
- [27] S. Dempe. “A simple algorithm for the-linear bilevel programming problem”. In: *Optimization* 18 (1987), pp. 373–385. DOI: [10.1080/02331938708843247](https://doi.org/10.1080/02331938708843247).
- [28] S. Dempe. *Foundations of Bilevel Programming*. Springer US, 2002.
- [29] S. Dempe, N. Dinh, and J. Dutta. “Optimality conditions for a simple convex bilevel programming problem”. In: *Variational Analysis and Generalized Differentiation in Optimization and Control: In Honor of Boris S. Mordukhovich*. Ed. by R.S. Burachik and J.-C. Yao. New York, NY: Springer New York, 2010, pp. 149–161. DOI: [10.1007/978-1-4419-0437-9_7](https://doi.org/10.1007/978-1-4419-0437-9_7).
- [30] S. Dempe, B. Mordukhovich, and A.B. Zemkoho. “Sensitivity analysis for two-level value functions with applications to bilevel programming”. In: *SIAM J. Optim.* 22 (2012), pp. 1309–1343. DOI: [10.1137/110845197](https://doi.org/10.1137/110845197).
- [31] S. Dempe, B.S. Mordukhovich, and A.B. Zemkoho. “Two-level value function approach to non-smooth optimistic and pessimistic bilevel programs”. In: *Optimization* 68 (2019), pp. 433–455. DOI: [10.1080/02331934.2018.1543294](https://doi.org/10.1080/02331934.2018.1543294).
- [32] S. Dempe and H. Schmidt. “On an algorithm solving two-level programming problems with nonunique lower level solutions”. In: *Comput. Optim. Appl.* 6 (1996), pp. 227–249. DOI: [10.1007/BF00247793](https://doi.org/10.1007/BF00247793).
- [33] E.J. Dockner, S. Jorgensen, N.V. Long, and G. Sorger. *Differential Games in Economics and Management Science*. Cambridge: Cambridge University Press, 2000.
- [34] L. Drouet, A. Haurie, F. Moresino, J.-P. Vial, M. Vielle, and L. Viguier. “An oracle based method to compute a coupled equilibrium in a model of international climate policy”. In: *Comput. Manag. Sci.* 5 (2008), pp. 119–140. DOI: [10.1007/s10287-007-0043-2](https://doi.org/10.1007/s10287-007-0043-2).
- [35] G. Eichfelder. “Multiobjective bilevel optimization”. In: *Math. Program.* 123 (2010), pp. 419–449. DOI: [10.1007/s10107-008-0259-0](https://doi.org/10.1007/s10107-008-0259-0).
- [36] A. Evgrafov and M. Patriksson. “On the convergence of stationary sequences in topology optimization”. In: *Int. J. Numer. Methods Eng.* 64 (2005), pp. 17–44. DOI: [10.1002/nme.1359](https://doi.org/10.1002/nme.1359).
- [37] S.D. Flâm. “Games and cost of change”. In: *Ann. Oper. Res.* (2020). DOI: <https://doi.org/10.1007/s10479-020-03585-w>.
- [38] S.D. Flâm and G.H. Greco. “Non-cooperative games; methods of subgradient projection and proximal point”. In: *Advances in Optimization*. Ed. by W. Oettli and D. Pallaschke. Berlin, Heidelberg: Springer Berlin Heidelberg, 1992, pp. 406–419.
- [39] D. Fudenberg and J. Tirole. *Game Theory*. Cambridge, MA: MIT Press, 1991.
- [40] J.C. Harsanyi and R. Selten. *A General Theory of Equilibrium Selection in Games*. The MIT Press, 1988.
- [41] A. Haurie, J.B. Krawczyk, and G. Zaccour. *Games and Dynamic Games*. World Scientific Publishing Company, 2012.

- [42] M. Hu and M. Fukushima. “Existence, uniqueness, and computation of robust Nash equilibria in a class of multi-leader-follower games”. In: *SIAM J. Optim.* 23 (2013), pp. 894–916. DOI: [10.1137/120863873](https://doi.org/10.1137/120863873).
- [43] A. Kannan and U. Shanbhag. “Distributed computation of equilibria in monotone Nash games via iterative regularization techniques”. In: *SIAM J. Optim.* 22 (2012), pp. 1177–1205. DOI: [10.1137/110825352](https://doi.org/10.1137/110825352).
- [44] L. Lampariello and S. Sagratella. “Numerically tractable optimistic bilevel problems”. Preprint. Nov. 2017. URL: http://www.optimization-online.org/DB_FILE/2017/10/6265.pdf.
- [45] L. Lampariello, S. Sagratella, and O. Stein. “The standard pessimistic bilevel problem”. In: *SIAM J. Optim.* 29 (2019), pp. 1634–1656. DOI: [10.1137/18M119759X](https://doi.org/10.1137/18M119759X).
- [46] G. Leitmann. “On generalized Stackelberg strategies”. In: *J. Optim. Theory Appl.* 26 (1978), pp. 637–643. DOI: [10.1007/BF00933155](https://doi.org/10.1007/BF00933155).
- [47] B. Lemaire. “Jeux dans les equations aux derivees partielles”. In: *Symposium on Optimization*. Ed. by A. V. Balakrishnan, M. Contensou, B. F. de Veubeke, P. Krée, J. L. Lions, and N. N. Moiseev. Berlin, Heidelberg: Springer Berlin Heidelberg, 1970, pp. 181–195.
- [48] M.B. Lignola. “Well-posedness and L-well-posedness for quasivariational inequalities”. In: *J. Optim. Theory Appl.* 128 (2006), pp. 119–138. DOI: [10.1007/s10957-005-7561-2](https://doi.org/10.1007/s10957-005-7561-2).
- [49] M.B. Lignola and J. Morgan. “A method to bypass the lack of solutions in MinSup problems under quasi-equilibrium constraints”. In: *Optim. Lett.* 10 (2016), pp. 833–846. DOI: [10.1007/s11590-015-0956-6](https://doi.org/10.1007/s11590-015-0956-6).
- [50] M.B. Lignola and J. Morgan. “ α -well-posedness for Nash equilibria and for optimization problems with Nash equilibrium constraints”. In: *J. Glob. Optim.* 36 (2006), pp. 439–459. DOI: [10.1007/s10898-006-9020-5](https://doi.org/10.1007/s10898-006-9020-5).
- [51] M.B. Lignola and J. Morgan. “Approximate solutions to variational inequalities and applications”. In: *Le Matematiche* 49 (1995), pp. 281–293. URL: <https://www.dmi.unict.it/ojs/index.php/lematematiche/article/viewFile/528/497>.
- [52] M.B. Lignola and J. Morgan. “Approximate values for mathematical programs with variational inequality constraints”. In: *Comput. Optim. Appl.* 53.2 (2012), pp. 485–503. DOI: [10.1007/s10589-012-9470-2](https://doi.org/10.1007/s10589-012-9470-2).
- [53] M.B. Lignola and J. Morgan. “Approximating values of MinSup problems with quasivariational inequality constraints”. In: *Pac. J. Optim.* 10 (2014), pp. 749–765.
- [54] M.B. Lignola and J. Morgan. “Asymptotic behavior of semi-quasivariational optimistic bilevel problems in Banach spaces”. In: *J. Math. Anal. Appl.* 424 (2015), pp. 1–20. DOI: [10.1016/j.jmaa.2014.10.059](https://doi.org/10.1016/j.jmaa.2014.10.059).
- [55] M.B. Lignola and J. Morgan. “Existence of solutions to generalized bilevel programming problem”. In: *Multilevel Optimization: Algorithms and Applications*. Ed. by A. Migdalas, P.M. Pardalos, and P. Värbrand. Boston, MA: Springer US, 1998, pp. 315–332. DOI: [10.1007/978-1-4613-0307-7_14](https://doi.org/10.1007/978-1-4613-0307-7_14).

- [56] M.B. Lignola and J. Morgan. “Further on inner regularizations in bilevel optimization”. In: *J. Optim. Theory Appl.* 180 (2019), pp. 1087–1097. DOI: [10.1007/s10957-018-1438-7](https://doi.org/10.1007/s10957-018-1438-7).
- [57] M.B. Lignola and J. Morgan. “Inner regularizations and viscosity solutions for pessimistic bilevel optimization problems”. In: *J. Optim. Theory Appl.* 173 (2017), pp. 183–202. DOI: [10.1007/s10957-017-1085-4](https://doi.org/10.1007/s10957-017-1085-4).
- [58] M.B. Lignola and J. Morgan. “Semi-continuities of marginal functions in a sequential setting”. In: *Optimization* 24 (1992), pp. 241–252. DOI: [10.1080/02331939208843793](https://doi.org/10.1080/02331939208843793).
- [59] M.B. Lignola and J. Morgan. “Stability in regularized quasi-variational wettings”. In: *J. Convex Anal.* 19 (2012), pp. 1091–1107.
- [60] M.B. Lignola and J. Morgan. “Stability of regularized bilevel programming problems”. In: *J. Optim. Theory Appl.* 93 (1997), pp. 575–596. DOI: [10.1023/A:1022695113803](https://doi.org/10.1023/A:1022695113803).
- [61] M.B. Lignola and J. Morgan. “Topological existence and stability for Stackelberg problems”. In: *J. Optim. Theory Appl.* 84 (1995), pp. 145–169. DOI: [10.1007/BF02191740](https://doi.org/10.1007/BF02191740).
- [62] M.B. Lignola and J. Morgan. “Vector quasi-variational inequalities: approximate solutions and well-posedness”. In: *J. Convex Anal.* 13 (2006), pp. 373–384.
- [63] M.B. Lignola and J. Morgan. “Viscosity solutions for bilevel problems with Nash equilibrium constraints”. In: *Far Est J. Appl. Math.* 88 (2014), pp. 15–34.
- [64] M.B. Lignola and J. Morgan. “Well-posedness for optimization problems with constraints defined by variational inequalities having a unique solution”. In: *J. Glob. Optim.* 16 (2000), pp. 57–67. DOI: [10.1023/A:1008370910807](https://doi.org/10.1023/A:1008370910807).
- [65] M.B. Lignola, J. Morgan, and V. Scalzo. “Lower convergence of approximate solutions to vector quasi-variational problems”. In: *Optimization* 59 (2010), pp. 821–832. DOI: [10.1080/02331930902863707](https://doi.org/10.1080/02331930902863707).
- [66] G.-H. Lin and J.J. Xu M.and Ye. “On solving simple bilevel programs with a nonconvex lower level program”. In: *Math. Program.* 144 (2014), pp. 277–305. DOI: [10.1007/s10107-013-0633-4](https://doi.org/10.1007/s10107-013-0633-4).
- [67] P. Loridan and J. Morgan. “A sequential stability result for constrained Stackelberg problems”. In: *Ric. Mat.* 38.1 (1989), pp. 19–32.
- [68] P. Loridan and J. Morgan. “A theoretical approximation scheme for Stackelberg problems”. In: *J. Optim. Theory Appl.* 61 (1989), pp. 95–110. DOI: [10.1007/BF00940846](https://doi.org/10.1007/BF00940846).
- [69] P. Loridan and J. Morgan. “Approximate solutions for two-level optimization problems”. In: *Trends in Mathematical Optimization: 4th French-German Conference on Optimization 1986*. Ed. by K.-H. Hoffmann, J. Zowe, J.-B. Hiriart-Urruty, and C. Lemarechal. Basel: Birkhäuser Basel, 1988, pp. 181–196. DOI: [10.1007/978-3-0348-9297-1_13](https://doi.org/10.1007/978-3-0348-9297-1_13).
- [70] P. Loridan and J. Morgan. “Approximation of the Stackelberg problem and applications in control theory”. In: *Annu. Rev. Autom. Program.* 13 (1985), pp. 121–124. DOI: [10.1016/0066-4138\(85\)90472-0](https://doi.org/10.1016/0066-4138(85)90472-0).

- [71] P. Loridan and J. Morgan. “Approximation results for two-level optimization problem and application to penalty methods”. Preprint n.52-1985. Dipartimento di Matematica e Applicazioni R. Caccioppoli, Università di Napoli. Dec. 1984. URL: https://www.researchgate.net/publication/335619596_Approximation_results_for_a_two-level_optimization_problem_and_application_to_penalty_methods.
- [72] P. Loridan and J. Morgan. “Convergence of approximate solutions and values in parametric vector optimization”. In: *Vector Variational Inequalities and Vector Equilibria: Mathematical Theories*. Ed. by F. Giannessi. Boston, MA: Springer US, 2000, pp. 335–350. DOI: [10.1007/978-1-4613-0299-5_19](https://doi.org/10.1007/978-1-4613-0299-5_19).
- [73] P. Loridan and J. Morgan. “ ϵ -Regularized two-level optimization problems: approximation and existence results”. In: *Optimization: Proceedings of the Fifth French-German Conference Held in Castel-Navet (Vareiz), France, Oct. 3–8, 1988*. Ed. by S. Dolecki. Berlin, Heidelberg: Springer Berlin Heidelberg, 1989, pp. 99–113. DOI: [10.1007/BFb0083589](https://doi.org/10.1007/BFb0083589).
- [74] P. Loridan and J. Morgan. “Least-norm regularization for weak two-level optimization problems”. In: *Optimization, Optimal Control and Partial Differential Equations*. Ed. by V. Barbu, D. Tiba, and J.F. Bonnans. Basel: Birkhäuser Basel, 1992, pp. 307–318. DOI: [10.1007/978-3-0348-8625-3_28](https://doi.org/10.1007/978-3-0348-8625-3_28).
- [75] P. Loridan and J. Morgan. “New results on approximate solution in two-level optimization”. In: *Optimization* 20 (1989), pp. 819–836. DOI: [10.1080/02331938908843503](https://doi.org/10.1080/02331938908843503).
- [76] P. Loridan and J. Morgan. “On Strict-Solution for a Two Level Optimization Problem”. In: *Proceedings of The international Conference On Operation Research*. Ed. by G. Feichtinger, W. Bühler, F. Radermacher, and P. Stähly. Vienna: Springer Verlag, 1992, pp. 165–172. DOI: [10.1007/978-3-642-77254-2_19](https://doi.org/10.1007/978-3-642-77254-2_19).
- [77] P. Loridan and J. Morgan. “Quasi convex lower level problem and applications in two level optimization”. In: *Generalized Convexity and Fractional Programming with Economic Applications: Proceedings of the International Workshop on “Generalized Convexity, Fractional Programming and Economic Applications” Held at the University of Pisa, Italy, May 30 – June 1, 1988*. Ed. by A. Cambini, E. Castagnoli, L. Martein, P. Mazzoleni, and S. Schaible. Berlin, Heidelberg: Springer Berlin Heidelberg, 1990, pp. 325–341. DOI: [10.1007/978-3-642-46709-7_23](https://doi.org/10.1007/978-3-642-46709-7_23).
- [78] P. Loridan and J. Morgan. “Regularizations for two-level optimization problems”. In: *Advances in Optimization*. Ed. by W. Oettli and D. Pallaschke. Berlin, Heidelberg: Springer Berlin Heidelberg, 1992, pp. 239–255.
- [79] P. Loridan and J. Morgan. “Weak via strong Stackelberg problem: new results”. In: *J. Glob. Optim.* (1996), pp. 263–287. DOI: [10.1007/BF00121269](https://doi.org/10.1007/BF00121269).
- [80] R. Lucchetti, F. Mignanego, and G. Pieri. “Existence theorems of equilibrium points in Stackelberg”. In: *Optimization* 18 (1987), pp. 857–866. DOI: [10.1080/02331938708843300](https://doi.org/10.1080/02331938708843300).
- [81] Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, 1996.

- [82] L. Mallozzi and J. Morgan. “Hierarchical systems with weighted reaction set”. In: *Non-linear Optimization and Applications*. Ed. by G. Di Pillo and F. Giannessi. Boston, MA: Springer US, 1996, pp. 271–282. DOI: [10.1007/978-1-4899-0289-4_19](https://doi.org/10.1007/978-1-4899-0289-4_19).
- [83] L. Mallozzi and J. Morgan. “Oligopolistic markets with leadership and demand functions possibly discontinuous”. In: *J. Optim. Theory Appl.* 125 (2005), pp. 393–407. DOI: [10.1007/s10957-004-1856-6](https://doi.org/10.1007/s10957-004-1856-6).
- [84] L. Mallozzi and J. Morgan. “Problema di Stackelberg con risposte pesate”. In: *Atti del XIX Convegno AMASES* (1995), pp. 416–425.
- [85] P. Marcotte and G. Savard. “A note on the Pareto optimality of solutions to the linear bilevel programming problem”. In: *Comput. & Oper. Res.* 18 (1991), pp. 355–359. DOI: [10.1016/0305-0548\(91\)90096-A](https://doi.org/10.1016/0305-0548(91)90096-A).
- [86] B. Martinet. “Brève communication. Régularisation d’inéquations variationnelles par approximations successives”. In: *Rev. Fr. Inform. Rech. Opér. Sér. Rouge* 4 (1970), pp. 154–158.
- [87] M. Maschler, E. Solan, and S. Zamir. *Game Theory*. Cambridge: Cambridge University Press, 2013.
- [88] D.A. Molodtsov. “The solution of a class of non-antagonistic games”. In: *USSR Comput. Math. Math. Phys.* 16 (1976), pp. 67–72. DOI: [10.1016/0041-5553\(76\)90042-2](https://doi.org/10.1016/0041-5553(76)90042-2).
- [89] D.A. Molodtsov and V.V. Fedorov. “Approximation of two-person games with information exchange”. In: *USSR Comput. Math. Math. Phys.* 13 (1973), pp. 123–142. DOI: [10.1016/0041-5553\(73\)90010-4](https://doi.org/10.1016/0041-5553(73)90010-4).
- [90] J.J. Moreau. “Proximité et dualité dans un espace hilbertien”. In: *Bull. Soc. Math. Fr.* 93 (1965), pp. 273–299.
- [91] J. Morgan. “Approximations and well-posedness in multicriteria games”. In: *Ann. Oper. Res.* 137 (2005), pp. 257–268. DOI: [10.1007/s10479-005-2260-9](https://doi.org/10.1007/s10479-005-2260-9).
- [92] J. Morgan. “Constrained well-posed two-level optimization problems”. In: *Nonsmooth Optimization and Related Topics*. Ed. by F.H. Clarke, V.F. Dem’yanov, and F. Giannessi. Boston, MA: Springer US, 1989, pp. 307–325. DOI: [10.1007/978-1-4757-6019-4_18](https://doi.org/10.1007/978-1-4757-6019-4_18).
- [93] J. Morgan and F. Patrone. “Stackelberg problems: subgame perfect equilibria via Tikhonov regularization”. In: *Advances in Dynamic Games: Applications to Economics, Management Science, Engineering, and Environmental Management*. Ed. by A. Haurie, S. Muto, L.A. Petrosjan, and T.E.S. Raghavan. Boston, MA: Birkhäuser Boston, 2006, pp. 209–221. DOI: [10.1007/0-8176-4501-2_12](https://doi.org/10.1007/0-8176-4501-2_12).
- [94] J. Morgan and R. Raucchi. “Lower semicontinuity for approximate social Nash equilibria”. In: *Int. J. Game Theory* 31 (2003), pp. 499–509. DOI: [10.1007/s001820300134](https://doi.org/10.1007/s001820300134).
- [95] J.F. Nash. “Equilibrium points in n -person games”. In: *Proc. Natl. Acad. Sci.* 36 (1950), pp. 48–49.
- [96] J.V. Outrata. “On the numerical solution of a class of Stackelberg problems”. In: *Z. Oper. Res.* 34 (1990), pp. 255–277. DOI: [10.1007/BF01416737](https://doi.org/10.1007/BF01416737).

- [97] M. Pagnozzi and S. Piccolo. “Vertical separation with private contracts”. In: *Econ. J.* 122 (2012), pp. 173–207. DOI: [10.1111/j.1468-0297.2011.02471.x](https://doi.org/10.1111/j.1468-0297.2011.02471.x).
- [98] J.-S. Pang and M. Fukushima. “Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games”. In: *Comput. Manag. Sci.* 2 (2005), pp. 21–56. DOI: [10.1007/s10287-004-0010-0](https://doi.org/10.1007/s10287-004-0010-0).
- [99] R.T. Rockafellar. “Monotone operators and the proximal point algorithm”. In: *SIAM J. Control Optim.* 14 (1976), pp. 877–898. DOI: [10.1137/0314056](https://doi.org/10.1137/0314056).
- [100] S. Sabach and S. Shtern. “A first order method for solving convex bilevel optimization problems”. In: *SIAM J. Optim.* 27 (2017), pp. 640–660. DOI: [10.1137/16M105592X](https://doi.org/10.1137/16M105592X).
- [101] R. Selten. “Spieltheoretische behandlung eines oligopolmodells mit nachfragerträgeit”. In: *Z. Gesamte Staatswiss.* (1965), pp. 301–324.
- [102] Y. Shehu, P.T. Vuong, and A. Zemkoho. “An inertial extrapolation method for convex simple bilevel optimization”. In: *Optim Methods Softw.* (2019), pp. 1–19. DOI: [10.1080/10556788.2019.1619729](https://doi.org/10.1080/10556788.2019.1619729).
- [103] K. Shimizu and E. Aiyoshi. “A new computational method for Stackelberg and min-max problems by use of a penalty method”. In: *IEEE Trans. Autom. Control* 26 (1981), pp. 460–466.
- [104] V.F. Sholohovich. “Unstable extremal problems and geometric properties of Banach spaces”. In: *Dokl. Akad. Nauk* 195 (1970), pp. 289–291.
- [105] M. Simaan and J.B. Cruz. “On the Stackelberg strategy in nonzero-sum games”. In: *J. Optim. Theory Appl.* 11 (1973), pp. 533–555. DOI: [10.1007/BF00935665](https://doi.org/10.1007/BF00935665).
- [106] M. Solodov. “An explicit descent method for bilevel convex optimization”. In: *J. Convex Anal.* 14 (2007).
- [107] A.N. Tikhonov. “Methods for the regularization of optimal control problems”. In: *Sov. Math. Dokl.* 6 (1965), pp. 761–763.
- [108] L. Vicente, G. Savard, and J. Júdice. “Descent approaches for quadratic bilevel programming”. In: *J. Optim Theory Appl.* 81 (1994), pp. 379–399. DOI: [10.1007/BF02191670](https://doi.org/10.1007/BF02191670).
- [109] H. von Stackelberg. *Marktform und Gleichgewicht*. [Translation: D. Bazin, R. Hill and L. Urch. *Market Structure and Equilibrium*. Springer-Verlag Berlin Heidelberg, 2011]. Berlin: Verlag von Julius Springer, 1934.
- [110] J. Ye, D. Zhu, and Q. Zhu. “Exact penalization and necessary optimality conditions for generalized bilevel programming problems”. In: *SIAM J. Optim.* 7 (1997), pp. 481–507. DOI: [10.1137/S1052623493257344](https://doi.org/10.1137/S1052623493257344).
- [111] J.J. Ye. “Necessary conditions for bilevel dynamic optimization problems”. In: *SIAM J. Control Optim.* 33 (1995), pp. 1208–1223. DOI: [10.1137/S0363012993249717](https://doi.org/10.1137/S0363012993249717).
- [112] J.J. Ye and D.L. Zhu. “Optimality conditions for bilevel programming problems”. In: *Optimization* 33 (1995), pp. 9–27. DOI: [10.1080/02331939508844060](https://doi.org/10.1080/02331939508844060).
- [113] A.B. Zemkoho. “Solving ill-posed bilevel programs”. In: *Set-Valued Var. Anal.* 24 (2016), pp. 423–448. DOI: [10.1007/s11228-016-0371-x](https://doi.org/10.1007/s11228-016-0371-x).