WORKING PAPER NO. 636

On construction of subgame perfect Nash equilibria in Stackelberg games

Francesco Caruso, Maria Carmela Ceparano, and Jacqueline Morgan

January 2022

This version July 2023

The previous version of the paper circulated under the title:
"A general method for selecting subgame perfect Nash equilibria in two-player Stackelberg games"

University of Naples Federico II **University of Salerno** **Bocconi University, Milan**

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

CSEF

Centre for Studies in Economics and Finance

Working Paper no. 636

On construction of subgame perfect Nash equilibria in Stackelberg games

Francesco Caruso*, Maria Carmela Ceparano†, and Jacqueline Morgan‡

Abstract

Identifying a Subgame Perfect Nash Equilibrium (henceforth SPNE) of a two-player Stackelberg game could be not a manageable task, especially when the players have a continuum of actions and the follower's best reply correspondence is not single-valued. Aim of the paper is to investigate the issue of construction of SPNEs in Stackelberg games by exploiting perturbations of both the action sets and the payoff functions of the leader and the follower. To achieve the goal, since the limit of SPNEs of perturbed games is not necessarily an SPNE of the original game even for classical perturbations, we prove under non-restrictive convergence conditions how to produce an SPNE starting from a sequence of SPNEs of general perturbed games. This result allows to describe a procedure to find SPNEs that can accommodate various types of perturbations. More precisely, under mild assumptions on the data of the original game, we show that a large class of perturbed games (including, for example, perturbation approaches relying on the Tikhonov and entropic regularizations or motivated by altruistic and antagonistic behaviors) satisfies the convergence conditions for constructing an SPNE. The specific SPNE selections associated to such a class, together with their possible behavioral interpretations, are discussed and an illustrative example is provided.

Keywords: Subgame perfect Nash equilibrium; two-player Stackelberg game; bilevel optimization problem; constructive method; perturbation.

Acknowledgments: This research was carried out in the frame of Programme STAR Plus "Exchange of Indivisible Goods, Externalities and Groups" [21- UNINA-EPIG-075], financially supported by UniNA and Compagnia di San Paolo.

* Università di Napoli Federico II. E-mail: francesco.caruso@unina.it

† Università di Napoli Federico II. E-mail: mariacarmela.ceparano@unina.it

‡ Università di Napoli Federico II and CSEF. E-mail: morgan@unina.it

1 Introduction

Two-player Stackelberg games (originally introduced in an economic setting in [57]) describe the interactions between two players, called leader and follower, who make their decisions non-cooperatively in a sequential way. In the first stage, the leader chooses an action trying to forecast the future decision of the follower; in the second stage, the follower chooses an action after observing what the leader has played. This play timing places Stackelberg games into the class of perfect-information extensive-form games. So the solution notion we consider in this framework is the traditional game-theoretical solution for extensive-form games, that is the subgame perfect Nash equilibrium concept (introduced in [53], henceforth SPNE; see also for example [23, 47]), well-known refinement of the Nash equilibrium solution concept used to rule out possible irrational behaviours of the players off the equilibrium path. The situation originally investigated by von Stackelberg concerns a sequential duopoly where the follower has a unique optimal reaction to any choice of the leader and the game has a unique SPNE. However, in general Stackelberg games the follower may not have a unique optimal reaction to any choice of the leader, so it could be an hard task to detect SPNEs, especially when the players have a continuum of actions. Hence, a key-issue is how to construct an SPNE in such a framework. Literature on how to obtain an SPNE in Stackelberg games can be collected in two main strands.

SPNEs from bilevel optimization solution concepts. In the considered perfect information framework, two ways to select an SPNE are connected to the solutions of widely investigated problems reflecting the two "extreme" possible beliefs that the leader can have about how the follower chooses an action among his optimal reactions. More precisely, if the leader thinks that the follower will choose the worst action for her, then she faces the so-called weak Stackelberg problem (or pessimistic bilevel optimization problem; see [48, 27, 34, 39, 49, 32] and also [4] for first results on existence, well-posedness, approximation and algorithms), whose solution induces an SPNE selection motivated by the pessimistic behaviour of the leader. Instead, if the leader believes that the follower will choose the best action for her, then she deals with the so-called strong Stackelberg problem (or optimistic bilevel optimization problem; see [5, 8, 51, 45, 56, 58, 31] and also [18] for first investigations on existence, approximation, optimality conditions and algorithms), whose solution induces an SPNE selection justified by the optimistic attitude of the leader. Furthermore, when the leader does not have the above described extreme beliefs on the follower's behavior but she has a more general belief that attributes probabilities to all the optimal reactions, any solution of the so-called intermediate Stackelberg problem (introduced in [42, 40]; see also [41] for an application to oligopolistic markets) leads to an SPNE selection that reflects the leader's belief. For further discussion on how the SPNEs are "induced" by weak, strong and intermediate Stackelberg solutions see [12, Section 4.2].

Note that, although behaviourally motivated, these ways of selection bring along the crucial issues of the mathematical problems to which they are connected. For instance, the solutions and the values of the weak Stackelberg problem may not be stable with respect to perturbations of the payoff functions (see [34, 32]); and the same holds also for the strong Stackelberg problem (see [31, Example 4.1]). Moreover, the weak Stackelberg problem is not guaranteed to have a solution even under compactness of the action sets and continuity of the payoff functions (see [4, Remark 4.3]);

and the same holds also for the intermediate Stackelberg problem (see [40, Example 3.5]).

Such crucial issues have been faced by the introduction of "surrogate" solutions (in weak Stackelberg problems see [34, 36, 35], [32] for approximate solutions and [30] for viscosity solutions; in intermediate Stackelberg problems see [40] for approximate solutions; in strong Stackelberg problems see [31] for approximate solutions; see also [12, Section 4.4] for a comprehensive discussion). Clearly, as well as for the weak and the strong Stackelberg solutions, these surrogate solutions induce related approximate SPNEs.

SPNEs from perturbations of the game. Another way to select an SPNE is through the construction of suitable sequences of perturbed two-player Stackelberg games having a single-valued follower's best reply correspondence. Such an approach leads to perturbed games that are better-manageable from both theoretical and numerical points of view, due to the uniqueness of the solution of the second-stage problem (see [33] and also [12, Section 4.3.1] and [18, Chapter 6] for results on existence, stability and numerical approximation in Stackelberg problems with a unique lower-level solution). Two situations have been investigated when the follower's payoff function is convex with respect to his own variable. First, in [50] an SPNE has been obtained by perturbing the follower's payoff function via the Tikhonov regularization ([55]), one of the most used regularization techniques in optimization. Afterwards, a learning approach for SPNEs has been introduced in [11] by exploiting proximal point techniques (connected to the Moreau-Yosida regularization, see [46, 52]) applied to the payoff functions of both players. Besides the uniqueness of the solution to the perturbed second-stage problem, the method presents a behavioural interpretation reflecting the players' aversion to the costs of deviation from a current action to another one. Furthermore, for the mixed extension of finite-actions Stackelberg games, a perturbation relying on the Shannon entropy ([54]) has been presented in [43] to approach SPNEs.

In this paper, instead of considering a specific perturbation, we are interested in providing non-restrictive sufficient conditions on the original game and on the "convergence" of general perturbed games in order to construct an SPNE of the original game. Having in mind this goal, we set the most general framework as possible: we consider Stackelberg games where the feasible decisions of the follower depend on the leader's choices (meaning that the set of the follower's feasible actions is defined by a set-valued map) and we consider perturbed games involving the perturbation both of the players' feasible action sets and of the player's payoff functions (so convergence of sequences of functions, sets and set-valued maps are managed). This setting enlarges the ones presented in [50] (where the perturbation involves the follower's payoff function only), in [11] (where the set of follower's feasible actions does not depend on the leader's actions and the perturbation does not affect the action sets) and in [43] (where the payoff functions are linear and the perturbation affects the follower's payoff function only). Given such a sequence of general perturbed games, we take into account a related sequence of SPNEs (that is, one SPNE for each perturbed game). Note that such a sequence does not necessarily converge to an SPNE of the original game even for nice data of the game (meaning compact action sets and smooth payoff functions) and for traditional perturbations (see [10] for an investigation on the asymptotic behavior of SPNEs in perturbed Stackelberg games). Our first contribution consists in proving that, under non-restrictive convergence assumptions on the perturbed games, an SPNE can be generated from the limit of the sequence of SPNEs of the perturbed games even when that limit is not. Hence, such a key-result allows to pursue also a linked goal: the description of a general procedure for constructing an SPNE in Stackelberg games. Our second contribution consists in showing that a large class of perturbed games actually satisfies these convergence assumptions. Moreover, such a class allows to

- accommodate the methods proposed in [50] and in [43] relying on the Tikhonov regularization and on the Shannon entropy, respectively;
- produce new constructive approaches motivated by altruistic and antagonistic attitudes of the players;
- deal with perturbed games whose SPNEs are more manageable to find;
- possibly characterize the SPNE generated.

Our findings represent a first step. In a future research, we aim to investigate the construction of SPNEs also in a dynamic framework and, in particular, to provide SPNEs selection results in the class of differential games with hierarchical play and its applications (see, e.g., [20, 6, 24, 7]).

The paper is organized as follows. In Section 2, after the description of the framework and of the general perturbation scheme, we first state some notions of convergence for sets, functions and set-valued maps that will be used (Section 2.1), and then we show how to construct an SPNE starting from a sequence of SPNEs of perturbed games (Section 2.2). In Section 3 we present the class of perturbed games whereby the convergence assumptions of the general constructive method are satisfied. Finally, in Section 4, we discuss the possible behavioral interpretations of specific perturbations and the connections with existing selection results for SPNEs, and we provide an illustrative example of SPNE construction.

2 Construction of SPNEs

Throughout the paper we consider a two-player Stackelberg game Γ, namely a sequential perfect-information game developed as follows: the leader moves first and chooses an action u^L in her action set A, then the follower observes u^L and chooses an action u^F in his set of feasible actions $\mathcal{K}(u^L)$ which depends on the leader's decision. After the moves, the leader receives the payoff $J^L(u^L, u^F) \in \mathbb{R}$ and the follower receives the payoff $J^F(u^L, u^F) \in \mathbb{R}$. Let U^L and U^F be subsets of two Euclidean spaces $\mathbb{R}^{m_L}$ and $\mathbb{R}^{m_F}$, respectively, and assume that $A \subseteq U^L$, $\mathcal{K}(u^L) \subseteq U^F$ for any $u^L \in A$ (i.e., $\mathcal{K}$ is a set-valued map from A to U^F) and the payoff functions J^L and J^F are real-valued functions defined on $U^L \times U^F$. To better emphasize the relevant features of the game, Γ is also referred to as $\langle A, \mathcal{K}, J^L, J^F \rangle$.
A follower's strategy is a function assigning to each action of the leader a feasible action for the follower, so the set of follower's strategies is

$$\mathcal{W}_{\mathcal{K}}^{A} := \{\varphi\colon A \to U^F \;\mid \varphi(u^L) \in \mathcal{K}(u^L) \text{ for any } u^L \in A\}. \tag{1}$$

Supposing each player seeks to minimize its own payoff[1], a subgame perfect Nash equilibrium ([53]) of Γ is a strategy profile $(\bar{u}^L, \bar{\varphi}) \in A \times \mathcal{W}_{\mathcal{K}}^{A}$ that satisfies:

[1]In this paper, we chose to minimize intending the payoffs as costs, but obviously, the setting can be turned into a maximization being $\max\{f(\cdot)\} = -\min\{-f(\cdot)\}$.

(SG1) $J^F(u^L, \bar{\varphi}(u^L)) \leq J^F(u^L, u^F)$ for any $u^L \in A$ and $u^F \in \mathcal{K}(u^L)$

(SG2) $J^L(\bar{u}^L, \bar{\varphi}(\bar{u}^L)) \leq J^L(u^L, \bar{\varphi}(u^L))$ for any $u^L \in A$.

In this framework, the follower's best reply correspondence is the set-valued map defined on A by

$$\begin{aligned}\mathcal{B}(u^L) &:= \operatorname*{arg\,min}_{u^F \in \mathcal{K}(u^L)} J^F(u^L, u^F) \\ &= \left\{ u^F \in \mathcal{K}(u^L) \mid J^F(u^L, u^F) = \min_{z \in \mathcal{K}(u^L)} J^F(u^L, z) \right\} \subseteq U^F,\end{aligned} \tag{2}$$

which assigns to each action of the leader the set of the feasible optimal reactions of the follower. In light of (2), the condition (SG1) is equivalent to $\bar{\varphi}(u^L) \in \mathcal{B}(u^L)$ for any $u^L \in A$.

Our general constructive method for SPNEs relies on a key-result concerning the asymptotic behavior of SPNEs of perturbed games. Let $(\Gamma_n)_n$ be a sequence of general perturbed two-player Stackelberg games defined by

$$\Gamma_n := \langle A_n, \mathcal{K}_n, J_n^L, J_n^F \rangle \quad \text{for any } n \in \mathbb{N},$$

where $A_n \subseteq U^L$, $\mathcal{K}_n \colon A_n \rightrightarrows U^F$, J_n^L and J_n^F are real-valued functions defined on $U^L \times U^F$ and, analogously to the notation in (1), the set of follower's strategies in Γ_n is denoted with $\mathcal{W}_{\mathcal{K}_n}^{A_n}$. Before presenting such a key-result, it is worth to recall some notions of convergence for sets, functions and set-valued maps and to state new ones that we will use in the sequel.

2.1 Preliminaries on notions of convergence

Preliminarily, we recall the definitions of limits of sets and convergence of set-valued maps in the sense of Painlevé-Kuratowski.

According to [26, Ch. 2, §29] or [2, Sect. 1.1]:

- *Lower limit.* A point $a \in \mathbb{R}^{m_L}$ belongs to $\operatorname{Liminf}_{n\to+\infty} A_n$ if there exists a sequence $(a_n)_n \subseteq \mathbb{R}^{m_L}$ such that $(a_n)_n$ converges to a and $a_n \in A_n$ for any $n \in \mathbb{N}$;
- *Upper limit.* A point $a \in \mathbb{R}^{m_L}$ belongs to $\operatorname{Limsup}_{n\to+\infty} A_n$ if there exists a sequence $(a_k)_k \subseteq \mathbb{R}^{m_L}$ such that $(a_k)_k$ converges to a and $a_{n_k} \in A_{n_k}$ for a subsequence $(n_k)_k \subseteq \mathbb{N}$;
- *Limit.* If $\operatorname{Liminf}_n A_n = \operatorname{Limsup}_n A_n = A$, we write $\operatorname{Lim}_{n\to+\infty} A_n = A$ and we say that the sequence $(A_n)_n$ converges to the set A.

Let $(\mathcal{G}_n)_n$ be a sequence of set-valued maps with $\mathcal{G}_n \colon A \rightrightarrows U^F$ for any $n \in \mathbb{N}$ and $\mathcal{G} \colon A \rightrightarrows U^F$. According to [29, Sect. 2.2]:

- *Lower convergence.* The sequence $(\mathcal{G}_n)_n$ lower converges to $\mathcal{G}$ if for any $a \in A$, any sequence $(a_n)_n \subseteq A$ converging to a and any $b \in U^F$, there exists a sequence $(b_n)_n \subseteq U^F$ converging to b such that $b_n \in \mathcal{G}_n(a_n)$ for n large (that is, $\mathcal{G}(a) \subseteq \operatorname{Liminf}_n \mathcal{G}_n(a_n)$ for any $a \in A$ and any sequence $(a_n)_n \subseteq A$ converging to a);
- *Upper convergence.* The sequence $(\mathcal{G}_n)_n$ upper converges to $\mathcal{G}$ if for any $a \in A$, any sequence $(a_n)_n \subseteq A$ converging to a and any sequence $(b_n)_n \subseteq U^F$ such that $b_{n_k} \in \mathcal{G}_{n_k}(a_{n_k})$ and that $(b_{n_k})_k$ converges to y for a subsequence $(n_k)_k \subseteq \mathbb{N}$, we have $b \in \mathcal{G}(a)$ (that is, $\operatorname{Limsup}_n \mathcal{G}_n(a_n) \subseteq \mathcal{G}(a)$ for any $a \in A$ and any sequence $(a_n)_n \subseteq A$ converging to a);

- *Convergence (in the sense of Painlevé-Kuratowski).* The sequence $(\mathcal{G}_n)_n$ converges to $\mathcal{G}$ if $\operatorname{Limsup}_n \mathcal{G}_n(a_n) \subseteq \mathcal{G}(a) \subseteq \operatorname{Liminf}_n \mathcal{G}_n(a_n)$ (that is, $\operatorname{Lim}_n \mathcal{G}_n(a_n) = \mathcal{G}(a)$) for any sequence $(a_n)_n \subseteq A$ converging to $a \in A$.

Now we state a notion of convergence for set-valued maps with varying domain (as the case of the follower's constraints correspondence $\mathcal{K}_n$ in Γ_n) which can be seen as an "halfway" concept between lower convergence and convergence (in the sense of Painlevé-Kuratowski). Recall that $(\mathcal{K}_n)_n$ is a sequence of set-valued maps with $\mathcal{K}_n \colon A_n \rightrightarrows U^F$ for any $n \in \mathbb{N}$ and that $\mathcal{K} \colon A \rightrightarrows U^F$.

Definition 2.1 We say that the sequence $(A_n, \mathcal{K}_n)_n$ *inf-converges to* $(A, \mathcal{K})$ if

(a) $\operatorname{Liminf}_n A_n \subseteq A$

(b) $\operatorname{Liminf}_n \mathcal{K}_n(u^L_n) = \mathcal{K}(u^L)$, for any sequence $(u^L_n)_n$ converging to u^L, with $u^L_n \in A_n$ for any $n \in \mathbb{N}$.

If $(A_n, \mathcal{K}_n)_n$ *inf-converges to* $(A, \mathcal{K})$ and $A_n = A$ for any $n \in \mathbb{N}$, we simply say that $(\mathcal{K}_n)_n$ *inf-converges to* $\mathcal{K}$.

Note that the concept of inf-convergence of $(\mathcal{K}_n)_n$ to $\mathcal{K}$ is stronger than the lower convergence and weaker than the convergence (in the sense of Painlevé-Kuratowski), as shown in the following examples.

Example 2.1 Let $\mathcal{K}_n \colon \mathbb{R} \rightrightarrows \mathbb{R}$ and $\mathcal{K} \colon \mathbb{R} \rightrightarrows \mathbb{R}$ be defined by

$$\mathcal{K}_n(u^L) = \begin{cases} [0, 1 - 1/n], & \text{if } u^L \neq 1 \\ \{1/n\}, & \text{if } u^L = 1 \end{cases} \quad \text{and} \quad \mathcal{K}(u^L) = \begin{cases} [0, 1], & \text{if } u^L \neq 1 \\ \{0\}, & \text{if } u^L = 1. \end{cases}$$

The sequence $\mathcal{K}_n$ lower converges to $\mathcal{K}$ (since $\mathcal{K}(1) \subseteq [0, 1] = \operatorname{Liminf}_n \mathcal{K}_n(u^L_n)$ for any sequence $(u^L_n)_n$ converging to 1), but $\mathcal{K}_n$ is not inf-convergent since $\operatorname{Liminf}_n \mathcal{K}_n(1 - 2/n) = [0, 1] \neq \{0\} = \mathcal{K}(1)$.

Example 2.2 Let $\mathcal{K}_n \colon \mathbb{R} \rightrightarrows \mathbb{R}$ be defined for any $n \in \mathbb{N}$ by

$$\mathcal{K}_n(u^L) = \begin{cases} [0, 1], & \text{if } n \text{ even} \\ [-1, 0], & \text{if } n \text{ odd}. \end{cases}$$

The sequence $\mathcal{K}_n$ inf-converges to the map defined on $\mathbb{R}$ by $\mathcal{K}(u^L) = \{0\}$, but $\mathcal{K}_n$ does not converges to $\mathcal{K}$ as $\operatorname{Limsup}_n \mathcal{K}_n(u^L) = [-1, 1] \nsubseteq \{0\} = \mathcal{K}(u^L)$ for any $u^L \in \mathbb{R}$.

Concerning the convergence of functions, let us first recall the definition of continuous convergence (see, e.g., [26, Ch. 2, §20-VI]) for bivariate functions. Recall that $(J^L_n)_n$ is a sequence of real-valued functions defined on $U^L \times U^F$ for any $n \in \mathbb{N}$.

- *Continuous convergence.* The sequence $(J^L_n)_n$ continuously converges to the function $J^L \colon U^L \times U^F \to \mathbb{R}$ if $\lim_{n \to +\infty} J^L_n(u^L_n, u^F_n) = J^L(u^L, u^F)$, for any sequence $(u^L_n, u^F_n)_n \subseteq U^L \times U^F$ converging to $(u^L, u^F) \in U^L \times U^F$.

Moreover, we present another convergence notion that we will use in the sequel.

Definition 2.2 We say that the sequence $(J_n^L)_n$ *partially continuously converges* to J^L if

(a) for any $(u^L, u^F) \in U^L \times U^F$ and any sequence $(u_n^L, u_n^F)_n \subseteq U^L \times U^F$ converging to (u^L, u^F), we have $\liminf_{n\to+\infty} J_n^L(u_n^L, u_n^F) \geq J^L(u^L, u^F)$;

(b) for any $(u^L, u^F) \in U^L \times U^F$ and any sequence $(u_n^F)_n \subseteq U^F$ converging to u^F, we have $\limsup_{n\to+\infty} J_n^L(u^L, u_n^F) \leq J^L(u^L, u^F)$;

The partial continuous convergence involves the continuous convergence with respect to only one variable. More precisely, given a sequence $(u_n^L)_n \subseteq U^L$ converging to $u^L \in U^L$ and defining h_n and h by $h_n(\cdot) = J_n^L(u_n^L, \cdot)$ and $h(\cdot) = J^L(u^L, \cdot)$, the partial continuous convergence of $(J_n^L)_n$ to J^L implies that the sequence $(h_n)_n$ continuously converges to h. However, the partial continuous convergence is weaker than the continuous convergence of $(J_n^L)_n$ to J^L, as illustrated in the following example.

Example 2.3 Let $J_n^L : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $J^L : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined by

$$J_n^L(u^L, u^F) = \begin{cases} nu^L u^F, & \text{if } 0 < u^L \leq 1/n \text{ and } u^F > 0 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad J^L(u^L, u^F) = 0$$

One can show that the sequence $(J_n^L)_n$ is partially continuously converges to J^L, but there is not continuously convergence as $\lim_n J_n^L(1/n, 1) = 1 \neq 0 = J^L(0, 1)$.

Note that the notion of partial continuous convergence is stronger than the convergence conditions defined in [33, p. 97] for the approximation of Stackelberg problems with single-valued follower's best reply correspondence (see [33, Remarks 2.2 and 2.3] for examples and discussion on connections with existing notions of convergence).

The notions of convergence considered in this section aim at achieving a compromise between reasons of readability and of sufficient generality in order to prove all the results presented throughout the paper. In fact, on the one hand, Definitions 2.1 and 2.2 involve conditions stronger than the minimal ones used in the frameworks of approximation and stability for Stackelberg and bilevel optimization problems (see, e.g., [33, 32, 31]). Nevertheless, on the other hand, such conditions are not so demanding since they are generally weaker than usual continuity requirements and accommodate many types of perturbation (as will be shown in Section 3).

2.2 Main result and constructive method

The following fundamental result shows the convergence assumptions on the data of the perturbed Stackelberg games $(\Gamma_n)_n$ which allow to obtain an SPNE of the original game Γ and how such an SPNE is actually generated from the "limit" of a sequence of SPNEs related to $(\Gamma_n)_n$.

Theorem 2.1. *Assume*

(H) *$A \subseteq A_n$ for any $n \in \mathbb{N}$, $(A_n, \mathcal{K}_n)_n$ inf-converges to $(A, \mathcal{K})$, $(J_n^L)_n$ partially continuously converges to J^L and $(J_n^F)_n$ continuously converges to J^F.*

Let $(\bar{u}_n^L, \bar{\varphi}_n) \in A_n \times \mathcal{W}_{\mathcal{K}_n}^{A_n}$ be an SPNE of Γ_n for any $n \in \mathbb{N}$.
If $(\bar{u}_n^L, \bar{\varphi}_n(\bar{u}_n^L))_n \subseteq U^L \times U^F$ converges to $(\bar{u}^L, \bar{u}^F)$ and if $(\bar{\varphi}_n(u^L))_n \subseteq U^F$ converges to $\bar{\varphi}(u^L)$ for any $u^L \in A$, then the pair $(\bar{u}^L, \widehat{\varphi})$ where

$$\widehat{\varphi}(u^L) := \begin{cases} \bar{\varphi}(u^L) & \text{if } u^L \neq \bar{u}^L \\ \bar{u}^F, & \text{if } u^L = \bar{u}^L, \end{cases} \tag{3}$$

is an SPNE of Γ.

Proof. Firstly note that $\bar{\varphi}_n(u^L)$ is well-defined for any $u^L \in A$ (as $A \subseteq A_n$). Moreover, in light of the inf-convergence of $(A_n, \mathcal{K}_n)_n$ to $(A, \mathcal{K})$ in Definition 2.1, we have $\bar{u}^L \in A$, $\bar{\varphi}(u^L) \in \mathcal{K}(u^L)$ for any $u^L \in A$ and $\bar{u}^F \in \mathcal{K}(\bar{u}^L)$. Hence $(\bar{u}^L, \widehat{\varphi}) \in A \times \mathcal{W}_{\mathcal{K}}^A$ is a strategy profile in Γ. We need to prove that $(\bar{u}^L, \widehat{\varphi})$ satisfies (SG1) and (SG2). The proof is divided in three steps.
Step 1. Let $u^L \in A \setminus \{\bar{u}^L\}$ and $u^F \in \mathcal{K}(u^L)$. By Definition 2.1(b), there exists a sequence $(\widetilde{u}_n^F)_n$ converging to u^F such that $\widetilde{u}_n^F \in \mathcal{K}_n(u^L)$ for any $n \in \mathbb{N}$. Hence, in light of the continuous convergence of $(J_n^F)_n$ to J^F and since $(\bar{\varphi}_n(u^L))_n$ converges to $\bar{\varphi}(u^L)$ and $\bar{\varphi}_n$ is an SPNE follower's strategy of Γ_n, we have

$$\begin{aligned} J^F(u^L, \widehat{\varphi}(u^L)) = J^F(u^L, \bar{\varphi}(u^L)) &= \lim_{n \to +\infty} J_n^F(u^L, \bar{\varphi}_n(u^L)) \\ &\leq \lim_{n \to +\infty} J_n^F(u^L, \widetilde{u}_n^F) = J^F(u^L, u^F). \end{aligned}$$

Step 2. Let $u^F \in \mathcal{K}(\bar{u}^L)$. In light of Definition 2.1(b), there exists a sequence $(\widetilde{u}_n^F)_n$ converging to u^F such that $\widetilde{u}_n^F \in \mathcal{K}_n(\bar{u}_n^L)$ for any $n \in \mathbb{N}$. So, since $(\bar{u}_n^L, \bar{\varphi}_n(\bar{u}_n^L))_n$ converges to $(\bar{u}^L, \bar{u}^F)$, by the same arguments of Step 1 it follows that

$$\begin{aligned} J^F(\bar{u}^L, \widehat{\varphi}(\bar{u}^L)) = J^F(\bar{u}^L, \bar{u}^F) &= \lim_{n \to +\infty} J_n^F(\bar{u}_n^L, \bar{\varphi}_n(\bar{u}_n^L)) \\ &\leq \lim_{n \to +\infty} J_n^F(\bar{u}_n^L, \widetilde{u}_n^F) = J^F(\bar{u}^L, u^F). \end{aligned}$$

Therefore, steps 1 and 2 shows that condition (SG1) is satisfied.
Step 3. Let $u^L \in A \setminus \{\bar{u}^L\}$. By the convergence of $(\bar{u}_n^L, \bar{\varphi}_n(\bar{u}_n^L))_n$ and of $(\bar{\varphi}_n(u^L))_n$ to $(\bar{u}^L, \bar{u}^F)$ and $\bar{\varphi}(u^L)$, respectively, and the partial continuous convergence of $(J_n^L)_n$ to J^L (Definition 2.2), we get

$$\begin{aligned} J^L(\bar{u}^L, \bar{u}^F) \leq \liminf_{n \to +\infty} J_n^L(\bar{u}_n^L, \bar{\varphi}_n(\bar{u}_n^L)) &\leq \limsup_{n \to +\infty} J_n^L(\bar{u}_n^L, \bar{\varphi}_n(\bar{u}_n^L)) \\ &\leq \limsup_{n \to +\infty} J_n^L(u^L, \bar{\varphi}_n(u^L)) \leq J^L(u^L, \bar{\varphi}(u^L)), \end{aligned}$$

where the third inequality holds in light of condition (SG2) applied to Γ_n observing that $u^L \in A_n$ for any $n \in \mathbb{N}$ (as $A \subseteq A_n$ for any $n \in \mathbb{N}$). This is sufficient to prove that $J^L(\bar{u}^L, \widehat{\varphi}(\bar{u}^L)) \leq J^L(u^L, \widehat{\varphi}(u^L))$ for any $u^L \in A$ (as $\widehat{\varphi}(\bar{u}^L) = \bar{u}^F$). Therefore, condition (SG2) is satisfied. □

In Theorem 2.1, the limit of the sequence $(\bar{\varphi}_n(\bar{u}_n^L))_n$ plays a crucial role for the definition of an SPNE of Γ. More precisely, two situations can happen.
⋄ *Case A:* $\lim_{n \to +\infty} \bar{\varphi}_n(\bar{u}_n^L) = \bar{\varphi}(\bar{u}^L)$. Then $(\bar{u}^L, \bar{\varphi})$ is an SPNE of Γ since $\widehat{\varphi} \equiv \bar{\varphi}$, where $\widehat{\varphi}$ is defined in (3). We point out that such an equality is guaranteed (and so $(\bar{u}^L, \bar{\varphi})$ is an SPNE of Γ) if we replace the assumption on convergence of $(\bar{\varphi}_n(u^L))_n$ to $\bar{\varphi}(u^L)$ for any $u^L \in A$ with the stronger condition

(∗) for any sequence $(u^L_n)_n$ converging to $u^L \in A$, with $u^L_n \in A_n$ for any $n \in \mathbb{N}$, we have $\lim_{n\to+\infty} \bar{\varphi}_n(u^L_n) = \bar{\varphi}(u^L)$.

Note that such an assumption is equivalent to the continuous convergence of $(\bar{\varphi}_n)_n$ to $\bar{\varphi}$ when $A_n = A$ for any $n \in \mathbb{N}$ and it is satisfied when the follower's best reply correspondences in Γ and in each Γ_n are single-valued (see [33, Proposition 3.1]). However, (∗) can be a very demanding requirement, since it may be not satisfied even under nice data of the game Γ and for nice types of perturbations. One can refer to [37, Remark 3.1] for a first counterexample (where the action sets are not perturbed and the Tikhonov regularization is used for the perturbation of just the follower's payoff function); see also [50] and [10, Remark 6] for further examples and discussion.

⋄ *Case B:* $\lim_{n\to+\infty} \bar{\varphi}_n(\bar{u}^L_n) \neq \bar{\varphi}(\bar{u}^L)$. Then $(\bar{u}^L, \bar{\varphi})$ is not necessarily an SPNE of Γ (see [11, Example 3] and [10, Section 3.2] for examples and discussion). Instead, the pair $(\bar{u}^L, \widehat{\varphi})$, where $\widehat{\varphi}$ and $\bar{\varphi}$ differ at $\bar{u}^L$, actually is an SPNE.

In summary, our key-result shows when and how a general sequence of SPNEs of perturbed games produces an SPNE of the original game, where "when" means assuming the hypotheses of Theorem 2.1 and "how" means taking into account the adjusted function $\widehat{\varphi}$. By considering different kinds of perturbations, such a result paves the way to define a procedure for constructing an SPNE in a two-player Stackelberg game. The procedure consists in:

1) choose a perturbation of the original game, possibly justified by specific players' attitudes, such that assumption **(H)** in Theorem 2.1 hold;

2) find an SPNE $(\bar{u}^L_n, \bar{\varphi}_n)$ for each perturbed Stackelberg game Γ_n;

3) take the limit $(\bar{u}^L, \bar{\varphi})$ of the sequence $(\bar{u}^L_n, \bar{\varphi}_n)_n$
(i.e., $\bar{u}^L$ is the limit of $(\bar{u}^L_n)_n$ and $\bar{\varphi}(u^L)$ is the limit of $(\bar{\varphi}_n(u^L))_n$ for any $u^L \in A$);

4) compute $\bar{u}^F := \lim_{n\to+\infty} \bar{\varphi}_n(\bar{u}^L_n)$ and

- if $\bar{u}^F = \bar{\varphi}(\bar{u}^L)$, then $(\bar{u}^L, \bar{\varphi})$ is an SPNE of Γ (*Case A*),
- if $\bar{u}^F \neq \bar{\varphi}(\bar{u}^L)$, then replace the follower's action $\bar{\varphi}(\bar{u}^L)$ with $\bar{u}^F$ and get $(\bar{u}^L, \widehat{\varphi})$ as an SPNE of Γ (*Case B*).

In order to be effective and most desirable, such a procedure for finding an SPNE should satisfy three key features: hypothesis **(H)** should hold for non-restrictive types of perturbations, the selected SPNE should be characterized in terms of properties connected to the perturbation chosen (and/or possibly related to players' behavioral attitudes), and the SPNEs of the perturbed games should not be hard to obtain. In the next section, we show a class of perturbations that accommodates all the features just described and such that **(H)** is satisfied.

3 A suitable class of perturbed games

In this section, assume that the sets of feasible actions of both players correspond to the solution sets of a finite number of inequalities, i.e.

$$A = \bigcap_{i=1}^{m_l} \{u^L \in U^L \mid g_i(u^L) \leq 0\} \text{ and } \mathcal{K}(u^L) = \bigcap_{j=1}^{m_f} \{u^F \in U^F \mid h_j(u^L, u^F) \leq 0\} \tag{4}$$

where $g_i\colon U^L \to \mathbb{R}$ for any $i \in \{1, \dots, m_l\}$ and $h_j\colon U^L \times U^F \to \mathbb{R}$ for any $j \in \{1, \dots, m_f\}$.
Let $(\epsilon_n, \nu_n, \alpha_n, \beta_n)_n \subseteq \mathbb{R}_+^{m_l} \times \mathbb{R}_+^{m_f} \times \mathbb{R}_+ \times \mathbb{R}_{++}$ be a sequence converging to $(\mathbf{0}, \mathbf{0}, 0, 0)$ and, for each $n \in \mathbb{N}$, consider the associate perturbed game $\Gamma_n = \langle A_n, \mathcal{K}_n, J_n^L, J_n^F \rangle$ defined by

$$\begin{aligned} A_n &= \bigcap_{i=1}^{m_l} \{u^L \in U^L \mid g_i(u^L) - \epsilon_{n,i} \leq 0\}, \\ \mathcal{K}_n(u^L) &= \bigcap_{j=1}^{m_f} \{u^F \in U^F \mid h_j(u^L, u^F) - \nu_{n,j} \leq 0\}, \end{aligned} \tag{5}$$

$$\begin{aligned} J_n^L(u^L, u^F) &= J^L(u^L, u^F) + \alpha_n C^L(u^L, u^F) \\ J_n^F(u^L, u^F) &= J^F(u^L, u^F) + \beta_n C^F(u^L, u^F), \end{aligned} \tag{6}$$

where $C^L\colon U^L \times U^F \to \mathbb{R}$ and $C^F\colon U^L \times U^F \to \mathbb{R}$.
The perturbation applied to the payoff functions of Γ, namely the definitions of J_n^L and J_n^F in (6), includes various types of perturbation approaches used both in bilevel optimization and in game theory frameworks, as for example:

- *Tikhonov regularization*: $C^L(u^L, u^F) = \|u^L\|^2$ and $C^F(u^L, u^F) = \|u^F\|^2$. Refer to [3, 28] for zero-sum differential games, [37, 17] for bilevel optimization problems, [50] for Stackelberg games and [25, 14] for normal-form games.
- *Entropic penalties*: $U^L = [0,1]^{m_L}$, $U^F = [0,1]^{m_F}$, $C^L(u^L, u^F) = \sum_{i=1}^{m_L} (u^L)_i \log{(u^L)_i}$ and $C^F(u^L, u^F) = \sum_{i=1}^{m_F} (u^F)_i \log{(u^F)_i}$ (setting $t \log t = 0$ when $t = 0$). Refer to [13] for normal-form games and [43] for Stackelberg games.
- *Altruism*: $C^L(u^L, u^F) = J^F(u^L, u^F)$ and $C^F(u^L, u^F) = J^L(u^L, u^F)$. Refer to [16, 15] for normal-form games and [19] for bilevel optimization problems.
- *Antagonism*: $C^L(u^L, u^F) = -J^F(u^L, u^F)$ and $C^F(u^L, u^F) = -J^L(u^L, u^F)$. Refer to [48, 38] for bilevel optimization problems.

Let $(\boldsymbol{\mathcal{H}})$ be the following set of assumptions on the original game Γ and on the perturbed games Γ_n:

$(\mathcal{H}_1)$ U^L is compact and g_i is lower semicontinuous for any $i \in \{1, \dots, m_l\}$;

$(\mathcal{H}_2)$ U^F is compact and convex, h_j is continuous and $h_j(u^L, \cdot)$ is semistrictly quasiconvex[2] (or, following [44, Ch. 9], strictly quasiconvex) on U^F for any $u^L \in U^L$ and $j \in \{1, \dots, m_f\}$;

$(\mathcal{H}_3)$ for each $u^L \in U^L$ there exists $u^F \in U^F$ such that $h_j(u^L, u^F) < 0$ for any $j \in \{1, \dots, m_f\}$;

$(\mathcal{H}_4)$ J^L is lower semicontinuous and $J^L(u^L, \cdot)$ is upper semicontinuous on U^F for any $u^L \in U^L$;

$(\mathcal{H}_5)$ J^F is continuous and $J^F(u^L, \cdot)$ is convex on U^F for any $u^L \in U^L$;

$(\mathcal{H}_6)$ C^L is continuous;

[2] That is, $h_j(u^L, \lambda z' + (1-\lambda) z'') < \max\{h_j(u^L, z'), h_j(u^L, z'')\}$ for any $\lambda \in]0,1[$ and any $z', z'' \in U^F$ such that $h_j(u^L, z') \neq h_j(u^L, z'')$; see, e.g., [9, Definition 2.3.1].

($\mathcal{H}_7$) C^F is continuous and $C^F(u^L, \cdot)$ is strictly convex on U^F for any $u^L \in U^L$.

In the next result, we prove that under the set of assumptions ($\boldsymbol{\mathcal{H}}$) there exists an SPNE in each perturbed Stackelberg game Γ_n and that the sequence $(\Gamma_n)_n$ satisfies the hypothesis (**H**). So, by applying Theorem 2.1, we can generate an SPNE of Γ.

Proposition 3.1. *Assume that ($\boldsymbol{\mathcal{H}}$) holds. Then*

(i) for any $n \in \mathbb{N}$, there exists an SPNE $(\bar{v}_n^L, \bar{\chi}_n)$ of Γ_n which satisfies

$$\{\bar{\chi}_n(u^L)\} = \arg\min_{u^F \in \mathcal{K}_n(u^F)} J_n^F(u^L, u^F) \quad \text{and} \quad \bar{v}_n^L \in \arg\min_{u^L \in A_n} J_n^L(u^L, \bar{\chi}_n(u^L)); \tag{7}$$

*(ii) hypothesis (**H**) is satisfied;*

(iii) if $(\bar{v}_n^L, \bar{\chi}_n(\bar{v}_n^L))_n$ converges to $(\bar{v}^L, \bar{v}^F)$ and if $(\bar{\chi}_n(u^L))_n$ converges to $\bar{\chi}(u^L)$ for any $u^L \in A$, then the strategy profile $(\bar{v}^L, \widehat{\chi})$ where

$$\widehat{\chi}(u^L) := \begin{cases} \bar{\chi}(u^L), & \text{if } u^L \neq \bar{v}^L \\ \bar{v}^F, & \text{if } u^L = \bar{v}^L \end{cases}$$

is an SPNE of Γ.

Proof. Firstly, assumptions ($\mathcal{H}_1$)-($\mathcal{H}_2$) guarantee that the sets A and A_n are compact for any $n \in \mathbb{N}$, $\mathcal{K}(u^L)$ is compact for any $u^L \in A$, and $\mathcal{K}_n(u^L)$ is compact for any $u^L \in A_n$ and $n \in \mathbb{N}$. Moreover, ($\mathcal{H}_2$) implies also that $h_j(u^L, \cdot)$ is quasiconvex for any $u^L \in U^L$ and $j \in \{1, \dots, m_f\}$ (see, for example, [9, Theorem 2.3.2]), so $\mathcal{K}$ and $\mathcal{K}_n$ are convex-valued for any $n \in \mathbb{N}$. Furthermore, assumption ($\mathcal{H}_5$) ensures that the follower's best reply correspondence $\mathcal{B}$ defined in (2) has compact and convex values.

Proof of (i). Fix $n \in \mathbb{N}$. The follower's reaction $\bar{\chi}_n(u^L)$ in (7) is well-defined for any $u^L \in U^L$ as, by assumptions ($\mathcal{H}_5$) and ($\mathcal{H}_7$), the function $J_n^F(u^L, \cdot)$ is continuous and strictly convex over the compact and convex set $\mathcal{K}_n(u^L)$. Moreover, in light of the continuity of J_n^F and assumptions ($\mathcal{H}_2$)-($\mathcal{H}_3$), the set-valued map $u^L \mapsto \arg\min_{u^F \in \mathcal{K}_n(u^L)} J_n^F(u^L, u^F)$ is closed by Proposition 2.1 in [31] and, being U^F compact, it is upper semicontinuous (see, e.g., [2, Proposition 1.4.8]). Furthermore, since $\arg\min_{u^F \in \mathcal{K}_n(u^L)} J_n^F(u^L, u^F) = \{\bar{\chi}_n(u^L)\}$, we have that $\bar{\chi}_n$ is a continuous function. Hence, by assumptions ($\mathcal{H}_4$) and ($\mathcal{H}_6$), the function $J_n^L(\cdot, \bar{\chi}_n(\cdot))$ is lower semicontinuous over the compact set A_n, and so $\bar{v}_n^L$ in (7) is well-defined. Then, $(\bar{v}_n^L, \bar{\varphi}_n)$ is an SPNE of Γ_n by definition.

Proof of (ii). Clearly, $A \subseteq A_n$ for any $n \in \mathbb{N}$ by (4)-(5) and since $\epsilon_n \in \mathbb{R}_+^{m_l}$.

Let us prove that $(A_n, \mathcal{K}_n)_n$ inf-converges to $(A, \mathcal{K})$. Picking $u^L \in \operatorname{Liminf}_n A_n$, there exists a sequence $(u_n^L)_n$ converging to u^L with $u_n^L \in A_n$ for any $n \in \mathbb{N}$ (by definition of lower limit). So $g_i(u_n^L) \leq \epsilon_{n,i}$ for any $i \in \{1, \dots, m_l\}$ and any $n \in \mathbb{N}$ and, in light of assumptions ($\mathcal{H}_1$), we have $g_i(u^L) \leq \liminf_n g_i(u_n^L) \leq 0$ for any $i \in \{1, \dots, m_l\}$; hence $u^L \in A$ and Definition 2.1(a) holds. From assumptions ($\mathcal{H}_2$),($\mathcal{H}_3$) and by applying Propositions 3.3.2 and 3.3.3 in [29], we get that Definition 2.1(b) is also satisfied.

Assumptions ($\mathcal{H}_4$) and ($\mathcal{H}_6$) guarantee that $(J_n^L)_n$ partially continuously converges to J^L. In fact,

on the one hand, taking a sequence $(u_n^L, u_n^F)_n \subseteq U^L \times U^F$ converging to (u^L, u^F), since $(\alpha_n)_n$ converges to 0, J^L is lower semicontinuous and C^L is continuous we have $\liminf_n J_n^L(u_n^L, u_n^F) = \liminf_n J^L(u_n^L, u_n^F) + \lim_n \alpha_n C^L(u_n^L, u_n^F) = \liminf_n J^L(u_n^L, u_n^F) \geq J^L(u^L, u^F)$. So Definition 2.2(a) is satisfied. On the other hand, by arguing as above and using the upper semicontinuity of $J^L(u^L, \cdot)$, also Definition 2.2(b) holds.
In light of assumptions $(\mathcal{H}_5)$,$(\mathcal{H}_7)$ and by exploiting analogous arguments, one can show that $(J_n^F)_n$ continuously converges to J^F.
Proof of (iii). It follows immediately by Theorem 2.1. □

If the set of follower's feasible actions is not perturbed (that is, $\Gamma_n = \langle A_n, \mathcal{K}, J_n^L, J_n^F \rangle$), we can even prove the convergence of sequence $(\bar{\chi}_n(u^L))_n$ and we can provide a (partial) characterization of the SPNE of Γ in terms of the payoffs' perturbation.

Proposition 3.2. *Assume that $(\mathcal{H})$ holds and that $\nu_n = \mathbf{0}$ for any $n \in \mathbb{N}$. Recall that $(\bar{v}_n^L, \bar{\chi}_n)_n$ is the sequence of SPNEs of Γ_n defined in (7). Then*

(i) the sequence $(\bar{\chi}_n(u^L))_n$ converges to $\bar{\chi}(u^L)$ for any $u^L \in A$, where $\bar{\chi}(u^L)$ is defined by

$$\{\bar{\chi}(u^L)\} := \arg\min_{u^F \in \mathcal{B}(u^L)} C^F(u^L, u^F), \tag{8}$$

and $\mathcal{B}$ is the follower's best reply correspondence, as defined in (2).

(ii) Proposition 3.1(iii) turns into:
if $(\bar{v}_n^L, \bar{\chi}_n(\bar{v}_n^L))_n$ converges to $(\bar{v}^L, \bar{v}^F)$, then the strategy profile $(\bar{v}^L, \widehat{\chi})$ where

$$\{\widehat{\chi}(u^L)\} := \begin{cases} \arg\min_{u^F \in \mathcal{B}(u^L)} C^F(u^L, u^F), & \text{if } u^L \neq \bar{v}^L \\ \{\bar{v}^F\}, & \text{if } u^L = \bar{v}^L \end{cases}$$

is an SPNE of Γ.

Proof. Note preliminarily that, since $\nu_n = \mathbf{0}$, then $\mathcal{K}_n \equiv \mathcal{K}$ for any $n \in \mathbb{N}$ and the sequence $(\bar{v}_n^L, \bar{\chi}_n)_n$ in (7) is rewritten as

$$\{\bar{\chi}_n(u^L)\} = \arg\min_{u^F \in \mathcal{K}(u^F)} J_n^F(u^L, u^F) \quad \text{and} \quad \bar{v}_n^L \in \arg\min_{u^L \in A_n} J_n^L(u^L, \bar{\chi}_n(u^L)). \tag{9}$$

Proof of (i). The convergence of the sequence $(\bar{\chi}_n(u^L))_n$ toward $\bar{\chi}(u^L)$ defined in (8) follows by generalizing the proofs of well-known results on the convergence of the Tikhonov regularization (see, for example, [21, Theorem 44]); for the sake of completeness, we provide a direct proof.
Fix $u^L \in A$. Firstly, by assumption $(\mathcal{H}_7)$ and since the set-valued map $\mathcal{B}$ has compact and convex values, the function $C^F(u^L, \cdot)$ has a unique minimizer over $\mathcal{B}(u^L)$. So $\bar{\chi}(u^L)$ defined in (8) is well-defined. Consider a subsequence $(\bar{\chi}_{n_k}(u^L))_k$ of $(\bar{\chi}_n(u^L))_n$ converging to $\widetilde{u}^F \in U^F$, whose existence is guaranteed by the compactness of U^F. Since $\mathcal{K}$ has compact values, in particular $\widetilde{u}^F \in \mathcal{K}(u^L)$. On the one hand, in light of the continuous convergence of $(J_n^F)_n$ towards J^F (see Proposition 3.1*(ii)*), the definition of $\bar{\chi}_n(u^L)$ in (9) and being $\lim_n \beta_n = 0$, we get

$$J^F(u^L, \widetilde{u}^F) = \lim_{k \to +\infty} J_{n_k}^F(u^L, \bar{\chi}_{n_k}(u^L)) \leq \lim_{k \to +\infty} J_{n_k}^F(u^L, u^F) = J^F(u^L, u^F),$$

for any $u^F \in U^F$. So $\widetilde{u}^F \in \mathcal{B}(u^L)$ and the definition of $\bar{\chi}$ implies that

$$C^F(u^L, \widetilde{u}^F) \geq C^F(u^L, \bar{\chi}(u^L)). \tag{10}$$

On the other hand, for any $k \in \mathbb{N}$ we have

$$\begin{aligned} J^F(u^L, \bar{\chi}_{n_k}(u^L)) + \beta_{n_k} C^F(u^L, \bar{\chi}_{n_k}(u^L)) &= J^F_{n_k}(u^L, \bar{\chi}_{n_k}(u^L)) \\ &\leq J^F_{n_k}(u^L, \bar{\chi}(u^L)) = J^F(u^L, \bar{\chi}(u^L)) + \beta_{n_k} C^F(u^L, \bar{\chi}(u^L)) \\ &\leq J^F(u^L, \bar{\chi}_{n_k}(u^L)) + \beta_{n_k} C^F(u^L, \bar{\chi}(u^L)), \end{aligned}$$

where the inequalities follows by the definition of $\bar{\chi}_{n_k}$ and $\bar{\chi}$. Since $(\beta_n)_n \subseteq \mathbb{R}_{++}$, then $C^F(u^L, \bar{\chi}_{n_k}(u^L)) \leq C^F(u^L, \bar{\chi}(u^L))$ which implies, taking the limit as k goes to $+\infty$,

$$C^F(u^L, \widetilde{u}^F) \leq C^F(u^L, \bar{\chi}(u^L)). \tag{11}$$

In light of inequalities (10)-(11), then $C^F(u^L, \widetilde{u}^F) = C^F(u^L, \bar{\chi}(u^L))$. So, reminding that $\bar{\chi}(u^L)$ is the unique minimizer of $C^F(u^L, \cdot)$ over $\mathcal{B}(u^L)$ (see (8) and the first part of the proof), it follows that $\widetilde{u}^F = \bar{\chi}(u^L)$. Since this holds for any convergent subsequence of $(\bar{\chi}_n(u^L))_n$ in the compact set U^F, then $\lim_{n\to+\infty} \bar{\chi}_n(u^L) = \bar{\chi}(u^L)$.
Proof of (ii). It follows by Proposition 3.2*(i)* and Proposition 3.1*(iii)*. □

Remark 3.1 Propositions 3.1 and 3.2 still hold if we assume continuity-like hypotheses weaker than $(\mathcal{H}_1)$-$(\mathcal{H}_2)$, and if we consider perturbed action sets (in Proposition 3.1) defined by the solutions of more general inequalities, instead of the specific ones in (5). We preferred to not state the minimal assumptions for the sake of readability; see [29, Proposition 3.3.1] for the detailed conditions. Furthermore, we point out that the strict convexity in assumption $(\mathcal{H}_7)$ can be weakened by inserting an additional quadratic term à la Tikhonov in the definition of J^F_n. More precisely, if we define $J^F_n(u^L, u^F) = J^F(u^L, u^F) + \beta_n[C^F(u^L, u^F) + \|u^F\|^2]$ and assume $C^F(u^L, \cdot)$ just convex, results analogous to Propositions 3.1 and 3.2 can be shown.

We highlight that the way to construct an SPNE shown in Propositions 3.1 and 3.2, starting from the sequence of perturbed games $(\Gamma_n)_n$ defined by (5)-(6), has two relevant characteristics. First, the follower's best reply correspondence of Γ_n is shown to be single-valued for any $n \in \mathbb{N}$ (see Proposition 3.1*(i)*), so the SPNE of each perturbed game can be easier to find (since it is unique from the follower's side) and the method bypasses the hard-to-control possible non-uniqueness of the follower's optimal reaction of the original game. Moreover, the follower's strategy in the selected SPNE is characterized in terms of the perturbation chosen (at least when the perturbation does not affect the players' action sets, see Proposition 3.2*(ii)*).

4 Concluding discussion and illustrative example

The choice of specific perturbations in the constructive procedure for SPNEs can lead to specific SPNE selections embodying behavioral interpretations. For example, focusing on the class defined in Section 3 and on the perturbation terms C^L and C^F in (6), we have:

- If $C^L(u^L, u^F) = \|u^L\|^2$ and $C^F(u^L, u^F) = \|u^F\|^2$, then these functions represent penalty terms that have a greater impact on the payoffs as the size of the player's decision variable increases and they can be interpreted as a punishment for pollution or a cost for high consumption levels. This explicitly affects the SPNE selected according to Proposition 3.2 *(ii)*, as $\arg\min_{u^F \in \mathcal{B}(u^L)} C^F(u^L, u^F) = \arg\min_{u^F \in \mathcal{B}(u^L)} \|u^F\|$ and so the follower would choose the minimum norm element in the set of his optimal reaction at least for all $u^L \neq \bar{v}^L$ (see [50] for the specific analysis of this Tikhonov-perturbed-approach when the perturbation involves the follower's payoff function only and also for the case of Stackelberg games with two followers).
- If $U^L = [0,1]^{m_L}$, $U^F = [0,1]^{m_F}$, $C^L(u^L, u^F) = \sum_{i=1}^{m_L} (u^L)_i \log (u^L)_i$ and $C^F(u^L, u^F) = \sum_{i=1}^{m_F} (u^F)_i \log (u^F)_i$, then we can interpret the game as affected by an entropic perturbation that measures the levels of uncertainty or disorder concerning the players. In this case, in the SPNE selected via Proposition 3.2 *(ii)*, the follower would choose (for $u^L \neq \bar{v}^L$) the action that has the maximal entropy in the set of his optimal reactions (see [43] for the investigation on perturbations of the follower's payoff function via the Shannon entropy in mixed extension of finite-action Stackelberg games with one and two followers).
- If $C^L(u^L, u^F) = J^F(u^L, u^F)$ and $C^F(u^L, u^F) = J^L(u^L, u^F)$, an idea of altruistic behavior can be understood by looking at (6)-(7), since both the leader and the follower make a compromise between minimizing their original payoff functions and slightly minimizing the payoff function of the other player (see [16], where a refinement concept of the Nash equilibrium based on this kind of slightly altruistic behavior has been introduced in normal-form games). The altruistic attitude of the follower can be observed also in the SPNE constructed according to Proposition 3.2 *(ii)*, as $\{\bar{\chi}(u^L)\} = \arg\min_{u^F \in \mathcal{B}(u^L)} J^L(u^L, u^F)$; so the follower would choose the best action for the leader among his optimal reactions (at least for $u^L \neq \bar{v}^L$).
- Contrary to the previous case, if $C^L(u^L, u^F) = -J^F(u^L, u^F)$ and $C^F(u^L, u^F) = -J^L(u^L, u^F)$, then (6)-(7) emphasize a sort of antagonistic attitude of the leader with respect to the follower and the other way round. In particular, the antagonistic feeling of the follower appears also in the SPNE found via Proposition 3.2 *(ii)*, since $\{\bar{\chi}(u^L)\} = \arg\min_{u^F \in \mathcal{B}(u^L)} \{-J^L(u^L, u^F)\}$.

Furthermore, we point out that the procedure to find an SPNE illustrated in Section 2.2 can accommodate also perturbations non-belonging to the class defined in Section 3. For example, the perturbation of the payoff functions relying on the proximal point method ([46, 52]) proposed in [11] satisfies the assumptions of Theorem 2.1, even if it does not fall into (6). Moreover, the SPNE constructed via such an approach has a behavioral interpretation connected to the willingness of the players to be close to their past decisions (and the dislike to pay the costs for moving from their current actions). We refer to [11] for the investigation on a learning method to select SPNEs in Stackelberg games motivated by *adverse-to-move* behaviors and to [1, 22] for discussion on *worthwhile-to-move* dynamics, *costs-of-change* and proximal algorithms.

Examples on the construction of an SPNE in Stackelberg games via perturbations of the payoff functions based on the Tikhonov regularization, the Shannon entropy and motivated by adverse-to-move behavior can be found in [50], [43] and [11], respectively. In such examples, the players' action sets are not perturbed, the follower's action set does not depend on the leader's decisions and the *Case B* of Section 2.2 occurs.

To complete the picture, we conclude the paper by showing an example of application of our general constructive procedure where the *Case A* of Section 2.2 appears, the sets of feasible actions of both players are perturbed, the follower's feasible action set depends on the leader's choices and the follower's payoff function is perturbed according to a slightly altruistic behavior (a perturbation not considered in the previous works). The perturbed games belong to the class defined in Section 3.

Example 4.1 Let $U^L = U^F = [-2, 2]$ and $\Gamma = \langle A, \mathcal{K}, J^L, J^F \rangle$ be the game where

$$A = \left[-\frac{1}{2}, \frac{1}{2}\right], \quad \mathcal{K}(u^L) = \left[-1 + |u^L|, 1 - |u^L|\right]$$
$$J^L(u^L, u^F) = -(u^L)^2 + (u^F)^2, \quad J^F(u^L, u^F) = -u^L u^F.$$

The sets A and $\mathcal{K}(u^L)$ are representable as in (4) (with $m_l = m_f = 2$, $g_1(u^L) = -u^L - \frac{1}{2}$, $g_2(u^L) = u^L - \frac{1}{2}$, $h_1(u^L, u^F) = |u^L| - 1 - u^F$ and $h_2(u^L, u^F) = |u^L| - 1 + u^F$) and assumptions $(\mathcal{H}_1)$-$(\mathcal{H}_5)$ are satisfied. Note that $\mathcal{K}(u^L)$ is well-defined (as $-1 + |u^L| \leq 1 - |u^L|$ if and only if $u^L \in [-1, 1] \supseteq A$).
The follower's best reply correspondence $\mathcal{B}$ is not single-valued and it is defined on $\left[-\frac{1}{2}, \frac{1}{2}\right]$ by

$$\mathcal{B}(u^L) = \begin{cases} \{-1 - u^L\}, & \text{if } u^L \in \left[-\frac{1}{2}, 0\right[\\ [-1, 1], & \text{if } u^L = 0 \\ \{1 - u^L\}, & \text{if } u^L \in \left]0, \frac{1}{2}\right]. \end{cases}$$

For any $n \in \mathbb{N}$, we consider the perturbed game $\Gamma_n = \langle A_n, \mathcal{K}_n, J_n^L, J_n^F \rangle$, where

$$A_n = \left[-\frac{1}{2}, \frac{1}{2} + \frac{1}{n}\right], \quad \mathcal{K}_n(u^L) = \left[-1 + |u^L| - \frac{1}{n}, 1 - |u^L| + \frac{1}{n}\right]$$
$$J_n^L(u^L, u^F) = -(u^L)^2 + (u^F)^2, \quad J_n^F(u^L, u^F) = -u^L u^F + \frac{1}{2n}\left[-(u^L)^2 + (u^F)^2\right].$$

So Γ_n fits (5)-(6) (with $\epsilon_{n,1} = 0$, $\epsilon_{n,2} = \frac{1}{n}$, $\nu_{n,1} = \nu_{n,2} = \frac{1}{n}$, $\alpha_n = 0$, $\beta_n = \frac{1}{2n}$, $C^L \equiv 0$ and $C^F(u^L, u^F) = -(u^L)^2 + (u^F)^2$) and assumptions $(\mathcal{H}_6)$-$(\mathcal{H}_7)$ are satisfied. Note that $\mathcal{K}_n(u^L)$ is well-defined (since $-1 + |u^L| - \frac{1}{n} \leq 1 - |u^L| + \frac{1}{n}$ if and only if $u^L \in \left[-1 - \frac{1}{n}, 1 + \frac{1}{n}\right] \supseteq A_n$).
The follower's optimal reaction in each perturbed game is unique and, moreover, Γ_n has a unique SPNE, that is the strategy profile $(\bar{v}_n^L, \bar{\chi}_n)$ defined by

$$\bar{v}_n^L = \frac{1}{2} + \frac{1}{n} \quad \text{and} \quad \bar{\chi}_n(u^L) = \begin{cases} -1 - u^L - \frac{1}{n}, & \text{if } u^L \in \left[-\frac{1}{2}, -\frac{1}{n}\right[\\ nu^L, & \text{if } u^L \in \left[-\frac{1}{n}, \frac{1}{n}\right] \\ 1 - u^L + \frac{1}{n}, & \text{if } u^L \in \left]\frac{1}{n}, \frac{1}{2} + \frac{1}{n}\right], \end{cases}$$

for n sufficiently large, in line with Proposition 3.1 *(i)*.
Both the sequence of action profiles $(\bar{v}_n^L, \bar{\chi}_n(\bar{v}_n^L))_n$ and the sequence of actions $(\bar{\chi}_n(u^L))_n$ (for any $u^L \in A$) are convergent. Denoting by $(\bar{v}^L, \bar{v}^F)$ and by $\bar{\chi}(u^L)$ the respective limits, we have

$$(\bar{v}^L, \bar{v}^F) = \left(\frac{1}{2}, \frac{1}{2}\right) \quad \text{and} \quad \bar{\chi}(u^L) = \begin{cases} -1 - u^L, & \text{if } u^L \in \left[-\frac{1}{2}, 0\right[\\ 0, & \text{if } u^L = 0 \\ 1 - u^L, & \text{if } u^L \in \left]0, \frac{1}{2}\right]. \end{cases}$$

Being $\lim_{n\to+\infty} \bar{\chi}_n(\bar{v}_n^L) = \bar{\chi}(\bar{v}^L)$, we fall into *Case A* of Section 2.2, so the SPNE of Γ obtained according to Proposition 3.1 *(iii)* is the strategy profile $(\bar{v}^L, \widehat{\chi})$ defined by

$$\bar{v}^L = \frac{1}{2} \quad \text{and} \quad \widehat{\chi}(u^L) = \bar{\chi}(u^L)$$

The perturbation term $C^F(u^L, u^F) = -(u^L)^2 + (u^F)^2 = J^L(u^L, u^F)$ in the follower's payoff function J_n^F of the perturbed game Γ_n leads to think that the follower has an altruistic attitude with respect to the leader (as discussed at the beginning of the section). Such a behavior actually appears in the SPNE selected. In fact, focusing on the point where the follower's best reply correspondence $\mathcal{B}$ is not single-valued, we have $\{\widehat{\chi}(0)\} = \arg\min_{u^F \in \mathcal{B}(0)} J^L(0, u^F)$, hence the follower would choose the best action for the leader among his optimal reactions.

Finally, note that the game Γ has infinitely many SPNEs: in particular each pair $(\bar{u}^L, \bar{\varphi}_\rho)$ where

$$\bar{u}^L \in \left\{-\frac{1}{2}, \frac{1}{2}\right\} \quad \text{and} \quad \bar{\varphi}_\rho(u^L) = \begin{cases} -1 - u^L, & \text{if } u^L \in \left[-\frac{1}{2}, 0\right[\\ \rho, & \text{if } u^L = 0 \\ 1 - u^L, & \text{if } u^L \in \left]0, \frac{1}{2}\right]. \end{cases}$$

at varying $\rho \in [0, 1]$, is an SPNE of Γ, among which the one selected by our procedure is $(\frac{1}{2}, \bar{\varphi}_0)$.

References

[1] H. Attouch and A. Soubeyran. "Local search proximal algorithms as decision dynamics with costs to move". In: *Set-Valued Anal* 19 (2011), pp. 157–177. DOI: 10.1007/s11228-010-0139-7.

[2] J.-P. Aubin and H. Frankowska. *Set-valued analysis*. Boston, MA: Birkhäuser, 2009.

[3] A. Auslender. "Recherche des points de selle d'une fonction". In: *Symposium on Optimization*. Ed. by A.V. Balakrishnan et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 1970, pp. 37–52.

[4] T. Başar and G. Olsder. *Dynamic Noncooperative Game Theory*. Society for Industrial and Applied Mathematics, 1998.

[5] 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.

[6] T. Başar and P. Bernhard. *H^∞ optimal control and related minimax design problems: a dynamic game approach*. Birkhäuser Boston, MA, 2008.

[7] T. Basar and G. Zaccour. *Handbook of Dynamic Game Theory*. Springer Cham, 2018.

[8] 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.

[9] A. Cambini and L. Martein. *Generalized Convexity and Optimization*. Berlin Heidelberg 2009: Springer, 2009. ISBN: 978-3-540-70875-9. DOI: 10.1007/978-3-540-70876-6.

[10] F. Caruso, M. C. Ceparano, and J. Morgan. "Asymptotic behavior of subgame perfect Nash equilibria in Stackelberg games". In: *Ann. Oper. Res.* Forthcoming (2023). Working paper available from: https://www.csef.it/WP/wp661.pdf.

[11] 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.

[12] F. Caruso, M. B. Lignola, and J. Morgan. "Regularization and approximation methods in Stackelberg games and bilevel optimization". In: *Bilevel Optimization: Advances and Next Challenges*. Ed. by S. Dempe and A. Zemkoho. Cham: Springer International Publishing, 2020, pp. 77–138. DOI: 10.1007/978-3-030-52119-6_4.

[13] E. Cavazzuti and S. D. Flåm. "Entropic penalties in finite games". In: *Ann. Oper. Res.* 137 (2005), pp. 331–348. DOI: 10.1007/s10479-005-2264-5.

[14] 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.

[15] G. De Marco and J. Morgan. "Friendliness and reciprocity in equilibrium selection". In: *International Game Theory Review* 10.1 (2008), pp. 53–72. DOI: 10.1142/S0219198908001777.

[16] G. De Marco and J. Morgan. "Slightly Altruistic Equilibria". In: *J. Optim. Theory Appl.* 137 (2008), pp. 347–362. DOI: 10.1007/s10957-008-9353-y.

[17] 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.

[18] S. Dempe. *Foundations of bilevel programming*. Springer US, 2002.

[19] 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.

[20] E. J. Dockner, S. Jorgensen, N. Van Long, and G. Sorger. *Differential Games in Economics and Management Science*. Cambridge University Press, 2000.

[21] A. L. Dontchev and T. Zolezzi. *Well-Posed Optimization Problems*. Springer, Berlin, Heidelberg, 1993.

[22] S.D. Flåm. "On games and cost of change". In: *Ann. Oper. Res.* 301 (2021), pp. 107–119. DOI: 10.1007/s10479-020-03585-w.

[23] D. Fudenberg and J. Tirole. *Game Theory*. Cambridge, MA: MIT Press, 1991.

[24] A. Haurie, J. B. Krawczyk, and G. Zaccour. *Games and Dynamic Games*. World Scientific, 2012.

[25] 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.

[26] K. Kuratowski. *Topology*. Vol. 1. New York: Academic Press, 1966.

[27] G. Leitmann. "On generalized Stackelberg strategies". In: *J. Optim. Theory Appl.* 26 (1978), pp. 637–643. DOI: 10.1007/BF00933155.

[28] B. Lemaire. "Jeux dans les equations aux derivees partielles". In: *Symposium on Optimization*. Ed. by A. V. Balakrishnan et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 1970, pp. 181–195.

[29] M.B. Lignola and J. Morgan. "Convergences of marginal functions with dependent constraints". In: *Optimization* 23 (1992), pp. 189–213. DOI: 10.1080/02331939208843758.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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-Novel (Varetz), France, Oct. 3–8, 1988*. Ed. by S. Dolecki. Berlin, Heidelberg: Springer Berlin Heidelberg, 1989, pp. 99–113. DOI: 10.1007/BFb0083589.

[36] P. Loridan and J. Morgan. "New results on approximate solution in two-level optimization". In: *Optimization* 20.6 (1989), pp. 819–836. DOI: 10.1080/02331938908843503.

[37] 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.

[38] P. Loridan and J. Morgan. "Weak via strong Stackelberg problem: new results". In: *J. Glob. Optim.* (1996), pp. 263–287. DOI: 10.1007/BF00121269.

[39] 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.

[40] L. Mallozzi and J. Morgan. "Hierarchical systems with weighted reaction set". In: *Nonlinear 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.

[41] 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.

[42] L. Mallozzi and J. Morgan. "Problema di Stackelberg con risposte pesate". In: *Atti del XIX Convegno AMASES* (1995), pp. 416–425.

[43] L. Mallozzi and P. M. Pardalos. "Entropic regularization in hierarchical games". In: *Operations Research Forum* 3 (2022). DOI: 10.1007/s43069-022-00130-2.

[44] O. L. Mangasarian. *Nonlinear programming*. Society for Industrial and Applied Mathematics, 1994. DOI: 10.1137/1.9781611971255.

[45] 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.

[46] B. Martinet. "Brève communication. Régularisation d'inéquations variationnelles par approximations successives". In: *Rev. Fr. Inform. Rech. Opér. Sér. Rouge* 4 (1970), pp. 154–158.

[47] M. Maschler, E. Solan, and S. Zamir. *Game theory*. Cambridge: Cambridge University Press, 2013.

[48] 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.

[49] 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.

[50] 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. Petrosyan, and T.E.S. Raghavan. Boston, MA: Birkhäuser Boston, 2006, pp. 209–221. DOI: 10.1007/0-8176-4501-2_12.

[51] 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.

[52] R.T. Rockafellar. "Monotone operators and the proximal point algorithm". In: *SIAM J. Control Optim.* 14 (1976), pp. 877–898. DOI: 10.1137/0314056.

[53] R. Selten. "Spieltheoretische behandlung eines oligopolmodells mit nachfrageträgheit: Teil i: Bestimmung des dynamischen preisgleichgewichts". In: *Z. Gesamte Staatswiss.* (1965), pp. 301–324.

[54] C. E. Shannon. "A mathematical theory of communication". In: *Bell Syst. Tech. J.* 27 (1948), pp. 379–423.

[55] A.N. Tikhonov. "Methods for the regularization of optimal control problems". In: *Sov. Math. Dokl.* 6 (1965), pp. 761–763.

[56] 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.

[57] 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.

[58] J.J. Ye and D.L. Zhu. "Optimality conditions for bilevel programming problems". In: *Optimization* 33 (1995), pp. 9–27. DOI: 10.1080/02331939508844060.