KYBERNETIKA — VOLUME 52 (2016), NUMBER 2, PAGES 258–279

AN OPTIMAL STRONG EQUILIBRIUM SOLUTION FOR COOPERATIVE MULTI-LEADER-FOLLOWER STACKELBERG MARKOV CHAINS GAMES

K. K. Trejo, J. B. Clempner and A. S. Poznyak

This paper presents a novel approach for computing the strong Stackelberg/Nash equilibrium for Markov chains games. For solving the cooperative n-leaders and m-followers Markov game we consider the minimization of the L_p-norm that reduces the distance to the utopian point in the Euclidian space. Then, we reduce the optimization problem to find a Pareto optimal solution. We employ a bi-level programming method implemented by the extraproximal optimization approach for computing the strong L_p-Stackelberg/Nash equilibrium. We validate the proposed method theoretically and by a numerical experiment related to marketing strategies for supermarkets.

Keywords: strong equilibrium, Stackelberg and Nash, L_p-norm, Markov chains

Classification: 35K20, 93B05

1. INTRODUCTION

Stackelberg games are usually represented by a leader-follower problem which corresponds to a bi-level programming problem. In bi-level programming problems there are two competing decision-making parties [4]: a) one is upper level decision makers and, b) the other is lower level decision makers. The two levels interact with each other as follows. The lower level is completely restricted by the upper level's decision and for each decision made by the upper level, lower level will choose the best option according to their objectives. Instead the upper level objectives are restricted from below by the lower level: the upper level control the lower level's decision in the way that lower level will react by choosing the best option.

In a Stackelberg game the leader's optimization problem is represented by the upper level, restricted by the follower's optimization mission at the lower level. The dynamics of a Stackelberg game is as follows ([30]): The leader considers the best-reply of the follower. Then, he/she commits to a mixed strategy (a probability distribution over deterministic schedules) that minimizes the cost, anticipating the predicted best-reply of the follower. Then, taking into the account the adversary's mixed strategy selection, the follower in equilibrium selects the expected best-reply that minimizes the cost.

DOI: 10.14736/kyb-2016-2-0258

Bi-level programming models have vast theoretical studies and applications in the real world. The traditional methods employed to solve these problems include penalty functions ([1]), the Karush–Kuhn–Tucker method ([6, 19]) and branch-and-bound procedures [5]. Relevant literature related to bi-level programming models is presented in ([14, 12, 20]). Alternative techniques using evolutionary algorithms have also been used to solve bi-level programming problems ([11, 24]). Applications were presented into the security domain by ([9, 27, 31]) suggesting a upper level that represents defenders trying to minimize risk, and a lower level that represents attackers trying maximizing destruction for a given target. Additionally, an application into energy area was suggested by ([15]) where the upper level represents the energy provider that minimizes total cost, and the lower level represents the energy consumer that determines the pattern of consumption. There are several applications implemented into different areas: transportation ([7, 10, 21]), agriculture ([16]), network ([23, 22]), management ([3]), gas ([13]).

This paper presents a novel approach for computing the strong Stackelberg/Nash equilibrium for Markov chains games. We solve the cooperative n-leaders and m-followers Markov game considering the minimization of the L_p-norm. The existence of the L_p-Stackelberg/Nash equilibrium is characterized as a strong Pareto policy, which is the closest in the Euclidean norm to the virtual minimum (utopia point). Then, we reduce the optimization problem to find a Pareto optimal solution. We employ a bi-level programming model implemented by the extraproximal optimization approach [2, 30] for computing the strong Stackelberg/Nash equilibrium. The extraproximal approach is a natural extension of the proximal and the gradient optimization methods used for solving the more difficult problems for finding an equilibrium point in game theory. It is defined by a two-step iterated procedure consisting of a prediction step that calculates the preliminary position approximation to the equilibrium point, and a basic adjustment of the previous step. We design the method for the static strong Stackelberg/Nash game in terms of nonlinear programming problems implementing the Lagrange principle. In addition, we make use of the Tikhonov's regularization method to ensure the convergence of the cost-functions to a unique Strong L_p-Stackelberg/Nash equilibrium. We formulate the nonlinear programming problem considering several linear constraints employing the c-variable method for making the problem computationally tractable. For solving each equation of the extraproximal optimization approach we use the projectional gradient method. The proposed method approaches in exponential time to a unique Strong L_p-Stackelberg/Nash equilibrium. The usefulness of the proposed solution is proved theoretically, and by an application example related to the effectiveness of relationship marketing strategies within the department store sector of the retail industry (supermarkets).

The remainder of the paper is structured as follows. The next Section presents the preliminaries needed to understand the rest of the paper. Section 3 establishes the definitions of the strong L_p-Stackelberg/Nash equilibrium. The extraproximal approach for the conditional optimization problems is described in Section 4. Section 5 proves that the proposed method approaches in exponential time to a unique strong Stackelberg/Nash equilibrium. A numerical example related to marketing strategies for supermarkets validates the proposed method in Section 6. We close with final comments in Section 7.

2. MARKOV GAMES

We consider the usual partial order for n-vectors x and y, the inequality $x \leq y$ means that $x_i \leq y_i$ for all $l = 1, \ldots, \mathcal{N}$. We have that

$$x < y \Leftrightarrow x \leq y \text{ and } x \neq y$$
$$x \ll y \Leftrightarrow x_l < y_l \text{ for all } l = 1, \ldots, \mathcal{N}.$$

A sequence $\{x^n\} \subset \mathbb{R}^n$ converging to x is said to converge in the direction $y \in \mathbb{R}^n$ if there is a sequence of positive numbers i_n such that $i_n \to 0$ and

$$\lim_{n\to\infty} (x^n - x)/i_n = y.$$

Let S be a finite set consisting of states $\{s_1, \ldots, s_N\}$, $N \in \mathbb{N}$, called the *state space.* A *Stationary Markov chain* ([8]) is a sequence of S-valued random variables $s(n)$, $n \in \mathbb{N}$, satisfying the *Markov condition*:

$$\begin{gathered} P(s(n+1) = s_{(j)} | s(n) = s_{(i)}, s(n-1) = s_{(i_{n-1})}, \ldots, s(1) = s_{(i_1)}) \\ P\left(s(n+1) = s_{(j)} | s(n) = s_{(i)}\right) =: \pi_{(ij)}. \end{gathered} \tag{1}$$

The Markov chain can be represented by a complete graph whose nodes are the states, where each edge $(s_{(i)}, s_{(j)}) \in S^2$ is labeled by the transition probability (1). The matrix $\Pi = (\pi_{(ij)})_{(s(i),s(j))\in S} \in [0,1]^{N\times N}$ determines the evolution of the chain: for each $n \in \mathbb{N}$, the power Π^n has in each entry $(s_{(i)}, s_{(j)})$ the probability of going from state $s_{(i)}$ to state $s_{(j)}$ in exactly n steps.

Definition 2.1. A controllable Markov chain ([26]) is a 4-tuple

$$MC = \{S, A, \mathbb{K}, \Pi\} \tag{2}$$

where:

- S is a finite set of states, $S \subset \mathbb{N}$.
- A is the set of actions, which is a metric space. For each $s \in S$, $A(s) \subset A$ is the non-empty set of admissible actions at state $s \in S$. Without loss of generality we may take $A = \cup_{s\in S} A(s)$;
- $\mathbb{K} = \{(s, a) | s \in S, a \in A(s)\}$ is the set of admissible state-action pairs, which is a finite subset of $S \times A$;
- $\Pi = \left[\pi_{(i,j|k)}\right]$ is a stationary controlled transition matrix, where

  $$\pi_{(ij|k)} \equiv P(s(n+1) = s_{(j)} | s(n) = s_{(i)}, a(n) = a_{(k)})$$

  represents the probability associated with the transition from state $s_{(i)}$ to state $s_{(j)}$ under an action $a_{(k)} \in A(s_{(i)})$, $k = 1, \ldots, M$, $M \in \mathbb{N}$.

Definition 2.2. A Markov Decision Process is a pair

$$MDP = \{MC, J\} \tag{3}$$

where:

- MC is a controllable Markov chain (2)
- $J : S \times \mathbb{K} \to \mathbb{R}$ is a cost function, associating to each state a real value.

The *Markov property* of the decision process in (3) is said to be fulfilled if

$$P(s(n+1)|(s(1), s(2), \ldots, s(n-1)), s(n) = s_{(i)}, a(n) = a_{(k)}) \\ = P(s(n+1)|s(n) = s_{(i)}, a(n) = a_{(k)}).$$

The strategy (policy)

$$d_{(k|i)}(n) \equiv P(a(n) = a_{(k)}|s(n) = s_{(i)})$$

represents the probability measure associated with the occurrence of an action $a(n)$ from state $s(n) = s_{(i)}$.

The elements of the transition matrix for the controllable Markov chain can be expressed as

$$P\left(s(n+1) = s_{(j)}|s(n) = s_{(i)}\right) \\ \sum_{k=1}^{M} P\left(s(n+1) = s_{(j)}|s(n) = s_{(i)}, a(n) = a_{(k)}\right) d_{(k|i)}(n).$$

Let us denote the collection $\left\{d_{(k|i)}(n)\right\}$ by D_n as follows

$$D_n = \left\{d_{(k|i)}(n)\right\}_{k=\overline{1,M},\ i=\overline{1,N}}.$$

A policy $\left\{d^{loc}(n)\right\}_{n\geq 0}$ is said to be *local optimal* if for each $n \geq 0$ it maximizes the conditional mathematical expectation of the utility function $J(d_{(k|i)}(n))$ under the condition that the history of the process

$$F_n := \{D_0, P\{s_0 = s(j)\}_{j=\overline{1,N}}; \ldots ; D_{n-1}, P\{s_n = s(j)\}_{j=\overline{1,N}}\}$$

is fixed and can not be changed hereafter, i. e., it realizes the "one-step ahead" conditional optimization rule

$$d^{loc}(n) := \arg \min_{d(n)\in D_n} \mathsf{E}\left\{J(d_{(k|i)}(n)) \mid F_n\right\} \tag{4}$$

where $J(d_{(k|i)}(n))$ is the utility function at the state s_{n+1}.

The dynamic of the game for Markov chains is described as follows. The game consists of $\mathcal{N}$ players (denoted by $l = \overline{1,\mathcal{N}}$) and begins at the initial state $s^l(0)$ which (as well as the states further realized by the process) is assumed to be completely measurable. Each of the players l is allowed to randomize, with distribution $d^l_{(k|i)}(n)$, over the pure action choices $a^l_{(k)} \in A^l\left(s^l_{(i)}\right)$, $i = \overline{1, N_l}$ and $k = \overline{1, M_l}$. From now on, we will consider only stationary strategies $d^l_{(k|i)}(n) = d^l_{(k|i)}$. These choices induce the state distribution dynamics

$$P^l\left(s^l(n+1){=}s_{(j_l)}\right) = \sum_{i_l=1}^{N_l}\left(\sum_{k_l=1}^{M_l} \pi^l_{(i_l,j_l|k_l)} d^l_{(k_l|i_l)}\right) P^l\left(s^{(l)}(n){=}s_{(i_l)}\right).$$

In the ergodic case when the Markov chain is ergodic for any stationary strategy $d^l_{(k|i)}$ the distributions $P^l\left(s^l(n+1)=s_{(j_l)}\right)$ exponentially fast converge to their limits $P^l\left(s=s_{(i)}\right)$ satisfying

$$P^l\left(s^{(l)}=s_{(j_l)}\right)=\sum_{i_l=1}^{N_l}\left(\sum_{k_l=1}^{M_l}\pi^l_{(i_l,j_l|k_l)}d^l_{(k_l|i_l)}\right)P^l\left(s^{(l)}=s_{(i_l)}\right). \tag{5}$$

The cost function of each player, depending on the states and actions of all the other players, is given by the values $W^l_{(i_1,k_1;\ldots;i_\mathcal{N},k_\mathcal{N})}$, so that the "average cost function" $\mathbf{J}^l$ in the stationary regime can be expressed as

$$\mathbf{J}^l\left(c^1,\ldots,c^\mathcal{N}\right):=\sum_{i_1,k_1}\cdots\sum_{i_\mathcal{N},k_\mathcal{N}}W^l_{(i_1,k_1,\ldots,i_\mathcal{N},k_\mathcal{N})}\prod_{l=1}^{\mathcal{N}}c^l_{(i_l,k_l)} \tag{6}$$

where

$$W^l_{(i_1,k_1,\ldots,i_\mathcal{N},k_\mathcal{N})}=\sum_{j_1}\cdots\sum_{j_\mathcal{N}}J^l_{(i_1,j_1,k_1,\ldots,i_\mathcal{N},j_\mathcal{N},k_\mathcal{N})}\prod_{l=1}^{\mathcal{N}}\pi^l_{(i_l,j_l|k_l)}$$

and $c^l:=\left[c^l_{(i_l,k_l)}\right]_{i_l=\overline{1,N_l};k_l=\overline{1,M_l}}$ is a matrix with elements

$$c^l_{(i_l,k_l)}=d^l_{(k_l|i_l)}P^l\left(s^{(l)}=s_{(i_l)}\right) \tag{7}$$

satisfying

$$c^l\in C^l_{adm}=\begin{cases} c^l:\sum_{i_l,k_l}c^l_{(i_l,k_l)}=1,\ c^l_{(i_l,k_l)}\geq 0, \\ \\ \sum_{k_l}c^l_{(j_l,k_l)}=\sum_{i_l,k_l}\pi^l_{(i_l,j_l|k_l)}c^l_{(i_l,k_l)} \end{cases} \tag{8}$$

where C^l_{adm} (C admissible). Notice that by (7) it follows that

$$P^l\left(s^{(l)}=s_{(i_l)}\right)=\sum_{k_l}c^l_{(i_l,k_l)} \qquad d^l_{(k_l|i_l)}=\frac{c^l_{(i_l,k_l)}}{\sum_{k_l}c^l_{(i_l,k_l)}}. \tag{9}$$

In the ergodic case $\sum_{k_l}c^l_{(i_l,k_l)}>0$ for all $l=\overline{1,\mathcal{N}}$. The *individual aim* of each player is $\mathbf{J}^l(c^l)\to\min_{c^{(l)}\in C^{(l)}_{adm}}$.

To study the existence of Pareto policies we shall first follow the well-known "scalarization" approach. Thus, given a n-vector $\lambda>0$ we consider the scalar (or real-valued) cost-function $\mathbf{J}$.

Let

$$u^l:=col\ \left(c^l_{(i_l,k_l)}\right),\ U^l:=C^{(l)}_{adm}\ \left(l=\overline{1,\mathcal{N}}\right),\ U:=\bigotimes_{l=1}^{\mathcal{N}}U^l$$

where *col* is the column operator which transforms the matrix $c^l_{(i_l,k_l)}$ in a column.

The *Pareto set* can be defined as ([17, 18])

$$\mathcal{P}:=\left\{u^*\left(\lambda\right):=\arg\min_{u\in U}\left[\sum_{l=1}^{\mathcal{N}}\lambda_l J^l\left(u\right)\right],\ \lambda\in\mathcal{S}^N\right\} \tag{10}$$

such that

$$\mathcal{S}^{\mathcal{N}} := \left\{ \lambda \in \mathbb{R}^{\mathcal{N}} : \lambda \in [0,1], \ \sum_{l=1}^{\mathcal{N}} \lambda_l = 1 \right\}.$$

The *Pareto front* is defined as the image of $\mathcal{P}$ under $\mathbf{J}$ as follows

$$\mathbf{J}(\mathcal{P}) := \left\{ \left(J^1 \left(u^* \left(\lambda\right)\right), J^2 \left(u^* \left(\lambda\right)\right), \ldots, J^{\mathcal{N}} \left(u^* \left(\lambda\right)\right)\right) | u^* \in \mathcal{P} \right\}.$$

The vector u^* is called a *Pareto optimal solution* for $\mathcal{P}$.

A *Nash equilibrium* is a strategy $u^* = \left(u^{0*}, \ldots, u^{\mathcal{N}*}\right)$ such that

$$\mathbf{J}\left(u^{0*}, \ldots, u^{\mathcal{N}*}\right) \leq \mathbf{J}\left(u^{0*}, \ldots, u^{l}, \ldots, u^{\mathcal{N}*}\right) \text{ for } u^l \in U^l, \ l = \overline{1, \mathcal{N}}.$$

A *strong Nash equilibrium* is a strategy $u^{**} = \left(u^{0**}, \ldots, u^{\mathcal{N}**}\right)$ such that there does not exist any $u^l \in U^l$

$$\mathbf{J}\left(u^{0**}, \ldots, u^{l}, \ldots, u^{\mathcal{N}**}\right) \leq \mathbf{J}\left(u^{0**}, \ldots, u^{\mathcal{N}**}\right) \text{ for } u^l \in U^l, \ l = \overline{1, \mathcal{N}}.$$

Formulation of the problem: The game problem is to find a policy u^* that minimizes $\mathbf{J}(u^1, \ldots, u^{\mathcal{N}})$ in the sense of Pareto.

Let $\mathcal{P}$ be a subset of $\mathbb{R}^n$. The *tangent cone* to $\mathcal{P}$ at $u \in \mathcal{P}$ is the set of all the directions $u' \in \mathbb{R}^n$ in which some sequence in $\mathcal{P}$ converges to u.

A vector $u^* \in \mathcal{P}$ in $\mathbb{R}^n$ is said to be

1. a *Pareto point* of $\mathcal{P}$ if there is no $u \in \mathcal{P}$ such that $u < u^*$;
2. a *weak Pareto* point of $\mathcal{P}$ if there is no $u \in \mathcal{P}$ such that $u \ll u^*$;
3. a *proper Pareto* point of $\mathcal{P}$ if u^* is a Pareto point and, in addition, the tangent cone to $\mathcal{P}$ at u^* does not contain vectors $u' < 0$.

A policy u^* is said to be a *Pareto policy* (or Pareto optimal) if there is no policy u such that $\mathbf{J}(u) < \mathbf{J}(u^*)$, and similarly for weak or proper Pareto policies.

Let $\|\cdot\|$ be the Euclidean norm in $\mathbb{R}^n$ and let $\varrho : \varDelta \to \mathbb{R}_+$ be the map defined as

$$\varrho(d) := \left\|\mathbf{J}(u) - \mathbf{J}(u^*)\right\|.$$

This is a utility function for the Markov Chain game in the sense that if u and u' are such that $\mathbf{J}(u) < \mathbf{J}(u')$, then $\varrho(u) < \varrho(u')$.

A policy u^* is said to be *strong Pareto optimal* (or a strong Pareto policy) if it minimizes the function ϱ that is,

$$\varrho(u^*) = \inf\left\{\varrho(u) | u^* \in \varDelta\right\} =: \varrho^*. \tag{11}$$

As ϱ is a utility function, it is clear that a strong Pareto policy is Pareto optimal, but of course the converse is not true.

3. THE STACKELBERG/NASH GAME

Let us introduce the variables

$$v^m := col\ c^{(m)},\ V^m := C_{adm}^{(m)}\ \left(m = \overline{1, \mathcal{M}}\right). \tag{12}$$

Let us consider a Stackelberg game *with $\mathcal{N}$ leaders* whose strategies are denoted by $u^l \in U^l\ \left(l = \overline{1, \mathcal{N}}\right)$ where U is a convex an compact set. Denote by $u = (u^1, \dots, u^{\mathcal{N}})^\top \in U$, the joint strategy of the players and $u^{\hat{l}}$ is a strategy of the rest of the players adjoint to u^l, namely,

$$u^{\hat{l}} := \left(u^1, \dots, u^{l-1}, u^{l+1}, \dots, u^{\mathcal{N}}\right)^\top \in U^{\hat{l}} := \bigotimes_{h=1,\ h\neq l}^{\mathcal{N}} U^h$$

such that $u = (u^l, u^{\hat{l}})\ \left(l = \overline{1, \mathcal{N}}\right)$. As well, let us consider $\mathcal{M}$ followers with strategies $v^m \in V^m\ \left(m = \overline{1, \mathcal{M}}\right)$ and V is also a convex an compact set. Denote by $v = (v^1, \dots, v^{\mathcal{M}}) \in V := \bigotimes_{m=1}^{\mathcal{M}} V^l$ the joint strategy of the followers and $v^{\hat{m}}$ is a strategy of the rest of the players adjoint to v^m, namely,

$$v^{\hat{m}} := \left(v^1, \dots, v^{m-1}, v^{m+1}, \dots, v^{\mathcal{M}}\right)^\top \in V^{\hat{m}} := \bigotimes_{q=1,\ q\neq m}^{\mathcal{M}} V^q$$

such that $v = (v^m, v^{\hat{m}})\ \left(m = \overline{1, \mathcal{M}}\right)$.

3.1. The Nash and Strong Nash equilibrium

The dynamics of the game is as follows. The *leaders play cooperatively* and they are assumed to anticipate the reactions of the followers trying to reach the strong Nash equilibria. For reaching the goal of the game leaders first try to find a joint strategy $u^* = \left(u^{1*}, \dots, u^{\mathcal{N}*}\right) \in U$ satisfying for any admissible $u^l \in U^l$ and any $l = \overline{1, \mathcal{N}}$

$$G_{L_p}(u, \hat{u}(u)) := \sum_{l=1}^{\mathcal{N}} \left[\left|\left(\min_{u^l \in U^l} \varphi_l\left(u^l, u^{\hat{l}}\right)\right) - \varphi_l\left(u^l, u^{\hat{l}}\right)\right|^p\right]^{1/p} \tag{13}$$

where $\hat{u}(u) = (u^{\hat{1}\top}, \dots, u^{\hat{\mathcal{N}}\top})^\top \in \hat{U} \subseteq \mathbb{R}^{\mathcal{N}(\mathcal{N}-1)}$ ([29, 28]). Here $\varphi_l\left(u^l, u^{\hat{l}}\right)$ is the cost-function of the leader l which plays the strategy $u^l \in U^l$ and the rest of the leaders play the strategy $u^{\hat{l}} \in U^{\hat{l}}$.

If we consider the utopia point

$$\bar{u}^l := \arg\min_{u^l \in U^l} \varphi_l\left(u^l, u^{\hat{l}}\right) \tag{14}$$

then, we can rewrite Eq. (13) as follows

$$G_{L_p}(u, \hat{u}(u)) := \sum_{l=1}^{\mathcal{N}} \left[\left|\varphi_l\left(\bar{u}^l, u^{\hat{l}}\right) - \varphi_l\left(u^l, u^{\hat{l}}\right)\right|^p\right]^{1/p}. \tag{15}$$

The functions $\varphi_l\left(u^l, u^{\hat{l}}\right)\ \left(l = \overline{1, \mathcal{N}}\right)$ are assumed to be convex in all their arguments.

Condition 3.1. The function $G_{L_p}(u,\hat{u}(u))$ satisfies *the* **Nash condition**

$$\max_{\hat{u}(u)\in\hat{U}} g(u,\hat{u}(u)) = \sum_{l=1}^{\mathcal{N}} \varphi_l\left(\bar{u}^l,u^{\hat{l}}\right) - \varphi_l\left(u^l,u^{\hat{l}}\right) \leq 0 \tag{16}$$

for any $u^l \in U^l$ and all $l=\overline{1,\mathcal{N}}$

Definition 3.2. A strategy $u^* \in U_{adm}$ is said to be a L_p-**Nash equilibrium** if

$$u^*_{L_p} \in Arg \min_{u\in U_{adm}} \left\{G_{L_p}(u,\hat{u}(u))\right\}. \tag{17}$$

Remark 3.3. If $G_{L_p}(u,\hat{u}(u))$ is strictly convex then

$$u^*_{L_p} = \arg\min_{u\in U_{adm}} \left\{G_{L_p}(u,\hat{u}(u))\right\}.$$

Definition 3.4. A strategy $u^{**} \in U$ is said to be a **Strong** L_p-**Nash equilibrium** if

$$u^{**}_{L_p} \in Arg \min_{u\in U_{adm},\lambda\in\mathcal{S}^N} \left\{G_{L_p}(u(\lambda),\hat{u}(u,\lambda))\right\}. \tag{18}$$

Remark 3.5. If $G_{L_p}(u(\lambda),\hat{u}(u,\lambda))$ is strictly convex then

$$u^{**}_{L_p} = \arg\min_{u\in U_{adm},\lambda\in\mathcal{S}^N} \left\{G_{L_p}(u(\lambda),\hat{u}(u,\lambda))\right\}.$$

As well, in this process the followers try to reach one of the Nash equilibria trying to find a joint strategy $v^* = \left(v^{1*},\ldots,v^{\mathcal{M}*}\right) \in V$ satisfying for any admissible $v^m \in V^m$ and any $m=\overline{1,\mathcal{M}}$

$$F_{L_p}(v,\hat{v}(v)) := \sum_{m=1}^{\mathcal{M}} \left[\left|\left(\min_{v^m\in V^m} \psi_m\left(v^m,v^{\hat{m}}\right)\right) - \psi_m\left(v^m,v^{\hat{m}}\right)\right|^p\right]^{1/p} \tag{19}$$

where $\hat{v}(v) = (v^{\hat{1}\top},\ldots,v^{\hat{\mathcal{M}}\top})^\top \in \hat{V} \subseteq \mathbb{R}^{\mathcal{N}(\mathcal{N}-1)}$ ([29, 28]). Here $\varphi_m\left(v^m,v^{\hat{m}}\right)$ is the cost-function of the follower m which plays the strategy $v^m \in V^m$ and the rest of the leaders play the strategy $v^{\hat{m}} \in V^{\hat{m}}$.

If we consider the utopia point

$$\bar{v}^m := \arg\min_{v^m\in V^m} \psi_m\left(v^m,v^{\hat{m}}\right) \tag{20}$$

then, we can rewrite Eq. (13) as follows

$$F_{L_p}(v,\hat{v}(v)) := \sum_{m=1}^{\mathcal{M}} \left[\left|\psi_m\left(\bar{v}^m,v^{\hat{m}}\right) - \psi_m\left(v^m,v^{\hat{m}}\right)\right|^p\right]^{1/p}. \tag{21}$$

The functions $\psi_m\left(v^m,v^{\hat{m}}\right)$ $\left(m=\overline{1,\mathcal{M}}\right)$ are assumed to be convex in all their arguments.

Condition 3.6. The function $F_{L_p}(v,\hat{v}(v))$ satisfies *the* **Nash condition**

$$\max_{\hat{v}(v)\in\hat{V}} f(v,\hat{v}(v)) = \sum_{m=1}^{\mathcal{M}} \psi_m\left(\bar{v}^m,v^{\hat{m}}\right) - \psi_m\left(v^m,v^{\hat{m}}\right) \leq 0 \tag{22}$$

for any $v^m \in V^m$ and all $m=\overline{1,\mathcal{M}}$.

3.2. The Stackelberg game

Leaders and followers together are in a Stackelberg game: the model involves two cooperatively Nash games restricted by a Stackelberg game defined as follows.

Definition 3.7. A *game with $\mathcal{N}$ leaders and $\mathcal{M}$* followers said to be **a cooperatively Stackelberg–Nash** game if

$$G_{L_p}\left(u(\lambda), \hat{u}(u,\lambda)|v\right) := \sum_{l=1}^{\mathcal{N}} \left[\left|\varphi_l\left(\bar{u}^l, u^{\hat{l}}|v\right) - \varphi_l\left(u^l, u^{\hat{l}}|v\right)\right|^p\right]^{1/p}$$

given $\lambda \in \mathcal{S}^{\mathcal{N}}$ such that

$$\max_{\hat{u}(u)\in\hat{U}} g\left(u, \hat{u}(u)|v\right) = \sum_{l=1}^{\mathcal{N}} \varphi_l\left(\bar{u}^l, u^{\hat{l}}|v\right) - \varphi_l\left(u^l, u^{\hat{l}}|v\right) \leq 0$$

where $u^{\hat{l}}$ is a strategy of the rest of the leaders adjoint to u^l, namely,

$$u^{\hat{l}} := \left(u^1, \ldots, u^{l-1}, u^{l+1}, \ldots, u^{\mathcal{N}}\right) \in U^{\hat{l}} := \bigotimes_{h=1,\ h\neq l}^{\mathcal{N}} U^h$$

and

$$\bar{u}^l := \arg\min_{u^l \in U^l} \varphi_l\left(u^l, u^{\hat{l}}|v\right)$$

such that

$$f_{L_p}\left(v(\theta), \hat{v}(v,\theta)|u\right) := \sum_{m=1}^{\mathcal{M}} \left[\left|\psi_m\left(\bar{v}^m, v^{\hat{m}}|u\right) - \psi_m\left(v^m, v^{\hat{m}}|u\right)\right|^p\right]^{1/p} \tag{23}$$

where $\theta \in \mathcal{S}^{\mathcal{N}}$ and given that $v^{\hat{m}}$ is a strategy of the rest of the followers adjoint to v^m, namely,

$$v^{\hat{m}} := \left(v^1, \ldots, v^{m-1}, v^{m+1}, \ldots, v^{\mathcal{M}}\right) \in V^{\hat{m}} := \bigotimes_{q=1,\ q\neq l}^{\mathcal{M}} V^q$$

and

$$\bar{v}^m := \arg\min_{v^m \in V^m} \psi_m\left(v^m, v^{\hat{m}}|u\right).$$

Remark 3.8. In the case of the bi-level approach introduced in Definition (3.7) we employ the restriction $f_{L_p}\left(v(\theta), \hat{v}(v,\theta)|u\right)$ in Eq. (23) for ensuring the followers to play cooperatively.

Definition 3.9. Let $G_{L_p}\left(u(\lambda), \hat{u}(u,\lambda)|v\right)$ be the cost functions of the leaders $\left(l = \overline{1,\mathcal{N}}\right)$. A strategy $u^* \in U$ of the leaders together with the collection $v^* \in V$ of the followers is said to be **a cooperatively Stackelberg–Nash equilibrium** if

$$\begin{gathered}(u^*, v^*) \in Arg \min_{u\in U, \hat{u}(u)\in\hat{U}, \lambda\in\mathcal{S}^{\mathcal{N}}} \max_{v\in V, \hat{v}(v)\in\hat{V}, \theta\in\mathcal{S}^{\mathcal{N}}} \\ \left\{G_{L_p}\left(u(\lambda), \hat{u}(u,\lambda)|v\right) | g\left(u, \hat{u}(u)|v\right) \leq 0, f_{L_p}\left(v(\lambda), \hat{v}(v,\lambda)|u\right) \leq 0\right\}.\end{gathered} \tag{24}$$

Remark 3.10. If $G_{L_p}(u(\lambda), \hat{u}(u,\lambda)|v)$ is strictly convex then

$$(u^*, v^*) = \arg \min_{u\in U, \hat{u}(u)\in\hat{U}, \lambda\in\mathcal{S}^{\mathcal{N}}} \max_{v\in V, \hat{v}(v)\in\hat{V}, \theta\in\mathcal{S}^{\mathcal{N}}} \left\{ G_{L_p}(u(\lambda), \hat{u}(u,\lambda)|v) \,|\, g(u, \hat{u}(u)|v) \le 0, \right. \\ \left. f_{L_p}(v(\theta), \hat{v}(v,\theta)|u) \le 0 \right\}.$$

4. THE EXTRAPROXIMAL METHOD FOR OPTIMIZATION PROBLEMS

4.1. The regularized Lagrange principle application

Applying the Lagrange principle (see, for example, [26]) for Definition 3.9, we may conclude that (24) can be rewritten as

$$(u^*, v^*) \in Arg \min_{u\in U, \hat{u}(u)\in\hat{U}, \lambda\in\mathcal{S}^{\mathcal{N}}} \max_{v\in V, \hat{v}(v)\in\hat{V}, \theta\in\mathcal{S}^{\mathcal{N}}, \omega\ge 0, \xi\ge 0} \mathcal{L}(u, \hat{u}(u), v, \hat{v}(v), \lambda, \theta, \omega, \xi)$$

$$\mathcal{L}(u, \hat{u}(u), v, \hat{v}(v), \lambda, \theta, \omega, \xi) := G_{L_p}(u(\lambda), \hat{u}(u,\lambda)|v) + \omega g(u, \hat{u}(u)|v) + \xi f_{L_p}(v(\theta), \hat{v}(v,\theta)|u).$$

The approximative solution obtained by the Tikhonov's regularization (see [26]) is given by

$$(u^*, v^*) \in \arg \min_{u\in U, \hat{u}(u)\in\hat{U}, \lambda\in\mathcal{S}^{\mathcal{N}}} \max_{v\in V, \hat{v}(v)\in\hat{V}, \theta\in\mathcal{S}^{\mathcal{N}}, \omega\ge 0, \xi\ge 0} \mathcal{L}(u, \hat{u}(u), v, \hat{v}(v), \lambda, \theta, \omega, \xi)$$

$$\mathcal{L}_\delta(u, \hat{u}(u), v, \hat{v}(v), \lambda, \omega, \xi) \\ := G_{L_p,\delta}(u(\lambda), \hat{u}(u,\lambda)|v) + \omega g_\delta(u, \hat{u}(u)|v) + \xi f_{L_p,\delta}(v(\theta), \hat{v}(v,\theta)|u) - \frac{\delta}{2}(\omega^2 + \xi^2) \tag{25}$$

where

$$G_{L_p,\delta}(u(\lambda), \hat{u}(u,\lambda)|v) = \sum_{l=1}^{\mathcal{N}} \left[\left| \varphi_l\left(\bar{u}^l, u^{\hat{l}}|v\right) - \varphi_l\left(u^l, u^{\hat{l}}|v\right) \right|^p \right]^{1/p} + \frac{\delta}{2}(\|u\|^2 + \|\hat{u}(u)\|^2 + \|\lambda\|^2)$$

$$g_\delta(u, \hat{u}(u)|v) = \sum_{l=1}^{\mathcal{N}} \left[\varphi_l\left(\bar{u}^l, u^{\hat{l}}|v\right) - \varphi_l\left(u^l, u^{\hat{l}}|v\right) \right] + \frac{\delta}{2}(\|u\|^2 + \|\hat{u}(u)\|^2)$$

$$f_{L_p,\delta}(v(\theta), \hat{v}(v,\theta)|u)$$

$$= \sum_{m=1}^{\mathcal{M}} \left[\left| \psi_m\left(\bar{v}^m, v^{\hat{m}}|u\right) - \psi_m\left(v^m, v^{\hat{m}}|u\right) \right|^p \right]^{1/p} + \frac{\delta}{2}(\|v\|^2 + \|\hat{v}(v)\|^2 + \|\theta\|^2).$$

Now, the function $G_\delta(u, \hat{u}(u)|v)$ is strictly convex if the Hessian matrix is positive semi-definite, then $G_\delta\left(u, \hat{u}(u)|v\right)$ attains a minimum at $(u, \hat{u}(u)|v)$ if

$$\nabla^2 G_\delta\left(u, \hat{u}(u)|v\right)$$

$$= \begin{bmatrix} \frac{\partial^2}{(\partial u_1)^2} G_\delta\left(u, \hat{u}(u)|v\right) & \cdots & \frac{\partial^2}{\partial u_1 \partial u_{\mathcal{N}}} G_\delta\left(u, \hat{u}(u)|v\right) \\ \frac{\partial^2}{\partial u_2 \partial u_1} G_\delta\left(u, \hat{u}(u)|v\right) & \cdots & \frac{\partial^2}{\partial u_2 \partial u_{\mathcal{N}}} G_\delta\left(u, \hat{u}(u)|v\right) \\ \cdots & \cdots & \cdots \\ \frac{\partial^2}{\partial u_{\mathcal{N}} \partial u_1} G_\delta\left(u, \hat{u}(u)|v\right) & \cdots & \frac{\partial^2}{(\partial u_{\mathcal{N}})^2} G_\delta\left(u, \hat{u}(u)|v\right) \end{bmatrix}$$

$$= \begin{bmatrix} \delta I_{n_1 \times n_1} & \mathcal{DG}_{1,2}(\hat{u}_{1,2}) & \cdots & \mathcal{DG}_{1,\mathcal{N}}(\hat{u}_{1,\mathcal{N}}) \\ \mathcal{DG}_{2,1}(\hat{u}_{2,1}) & \delta I_{n_2 \times n_2} & \cdots & \mathcal{DG}_{3,2}(\hat{u}_{3,2}) \\ \cdots & \cdots & \cdots & \cdots \\ \mathcal{DG}_{3,1}(\hat{u}_{3,1}) & \mathcal{DG}_{3,2}(\hat{u}_{3,2}) & \cdots & \delta I_{n_{\mathcal{N}} \times n_{\mathcal{N}}} \end{bmatrix} > 0$$

or, equivalently, δ should provide the inequality

$$\min_{u \in U, \hat{u} \in \hat{U}} \left[\lambda_{\min}\left(\nabla^2 G_\delta\left(u, \hat{u}(u)|v\right)\right)\right] > 0. \tag{26}$$

Here, $\hat{u}_{ik}$ is independent of $u^{(i)}$ and $u^{(k)}$, that is, $\frac{\partial}{\partial u^{(i)}} \hat{u}_{ik} = 0$ and $\frac{\partial}{\partial u^{(k)}} \hat{u}_{ik} = 0$.

As well as, the function $f_\delta\left(v, \hat{v}(v)|u\right)$ is strictly concave if the Hessian matrix is negative semi-definite, then $f_\delta\left(v, \hat{v}(v)|u\right)$ attains a maximum at $(v, \hat{v}(v)|u)$ if

$$\max_{v \in V, \hat{v} \in \hat{V}} \left[\lambda_{\max}\left(\nabla^2 f_\delta\left(v, \hat{v}(v)|u\right)\right)\right] < 0. \tag{27}$$

With sufficiently large δ, the considered functions provide the uniqueness of the conditional optimization problem (25).

Notice also that the Lagrange function in (25) satisfies the saddle-point ([25]) condition, namely, for all $u \in U$, $\hat{u} \in \hat{U}$, $v \in V$, $\hat{v}(v) \in \hat{V}$, $\lambda \in \mathcal{S}^{\mathcal{N}}$, $\theta \in \mathcal{S}^{\mathcal{N}}$, $\omega \geq 0$ and $\xi \geq 0$ we have

$$\begin{aligned} \mathcal{L}_\delta(u_\delta^*, \hat{u}_\delta^*(u), v_\delta, \hat{v}_\delta(v), \lambda_\delta^*, \theta_\delta, \omega_\delta, \xi_\delta) &\leq \mathcal{L}_\delta(u_\delta^*, \hat{u}_\delta^*(u), v_\delta^*, \hat{v}_\delta^*(v), \lambda_\delta^*, \theta_\delta^*, \omega_\delta^*, \xi_\delta^*) \\ &\leq \mathcal{L}_\delta(u_\delta, \hat{u}_\delta(u), v_\delta^*, \hat{v}_\delta^*(v), \lambda_\delta, \theta_\delta^*, \omega_\delta^*, \xi_\delta^*). \end{aligned} \tag{28}$$

4.2. The proximal format

In the *proximal format* (see, [2]) the relation (25) can be expressed as

$$
\begin{aligned}
\omega_\delta^* &= \arg\max_{\omega\geq 0}\left\{-\tfrac{1}{2}\|\omega-\omega_\delta^*\|^2+\gamma\mathcal{L}_\delta(u_\delta^*,\hat{u}_\delta^*(u),v_\delta^*,\hat{v}_\delta^*(v),\lambda_\delta^*,\theta_\delta^*,\omega_\delta,\xi_\delta^*)\right\}\\
\xi_\delta^* &= \arg\max_{\xi\geq 0}\left\{-\tfrac{1}{2}\|\xi-\xi_\delta^*\|^2+\gamma\mathcal{L}_\delta(u_\delta^*,\hat{u}_\delta^*(u),v_\delta^*,\hat{v}_\delta^*(v),\lambda_\delta^*,\theta_\delta^*,\omega_\delta^*,\xi_\delta)\right\}\\
u_\delta^* &= \arg\min_{u\in U}\left\{\tfrac{1}{2}\|u-u_\delta^*\|^2+\gamma\mathcal{L}_\delta(u_\delta,\hat{u}_\delta^*(u),v_\delta^*,\hat{v}_\delta^*(v),\lambda_\delta^*,\theta_\delta^*,\omega_\delta^*,\xi_\delta^*)\right\}\\
\hat{u}_\delta^* &= \arg\min_{\hat{u}\in\hat{U}}\left\{\tfrac{1}{2}\|\hat{u}-\hat{u}_\delta^*\|^2+\gamma\mathcal{L}_\delta(u_\delta^*,\hat{u}_\delta(u),v_\delta^*,\hat{v}_\delta^*(v),\lambda_\delta^*,\theta_\delta^*,\omega_\delta^*,\xi_\delta^*)\right\}\\
v_\delta^* &= \arg\max_{v\in V}\left\{-\tfrac{1}{2}\|v-v_\delta^*\|^2+\gamma\mathcal{L}_\delta(u_\delta^*,\hat{u}_\delta^*(u),v_\delta,\hat{v}_\delta^*(v),\lambda_\delta^*,\theta_\delta^*,\omega_\delta^*,\xi_\delta^*)\right\}\\
\hat{v}_\delta^* &= \arg\max_{\hat{v}\in\hat{V}}\left\{-\tfrac{1}{2}\|\hat{v}-\hat{v}_\delta^*\|^2+\gamma\mathcal{L}_\delta(u_\delta^*,\hat{u}_\delta^*(u),v_\delta^*,\hat{v}_\delta(v),\lambda_\delta^*,\theta_\delta^*,\omega_\delta^*,\xi_\delta^*)\right\}\\
\lambda_\delta^* &= \arg\min_{\lambda\in\mathcal{S}^N}\left\{\tfrac{1}{2}\|\lambda-\lambda_\delta^*\|^2+\gamma\mathcal{L}_\delta(u_\delta^*,\hat{u}_\delta^*(u),v_\delta^*,\hat{v}_\delta^*(v),\lambda_\delta,\theta_\delta^*,\omega_\delta^*,\xi_\delta^*\right\}\\
\theta_\delta^* &= \arg\max_{\theta\in\mathcal{S}^N}\left\{-\tfrac{1}{2}\|\theta-\theta_\delta^*\|^2+\gamma\mathcal{L}_\delta(u_\delta^*,\hat{u}_\delta^*(u),v_\delta^*,\hat{v}_\delta^*(v),\lambda_\delta^*,\theta_\delta,\omega_\delta^*,\xi_\delta^*\right\}
\end{aligned}
\tag{29}
$$

where the solutions $u_\delta^*,\hat{u}_\delta^*(u),v_\delta^*,\hat{v}_\delta^*(v),\lambda_\delta^*,\omega_\delta^*$ and ξ_δ^* depend on the parameters $\delta,\gamma>0$.

4.3. The Extraproximal method

The *Extraproximal Method* for the conditional optimization problems (25) was suggested in ([2, 30]). We design the method for the static Stackelberg–Nash game in a general format. The general format iterative version ($n=0,1,\ldots$) of the extraproximal method with some fixed admissible initial values ($u_0\in U$, $\hat{u}_0\in U$, $v_0\in V$, $\hat{v}_0\in V$, $\omega_0\geq 0$, $\xi_0\geq 0$, $\lambda_0\in\mathcal{S}^{\mathcal{N}}$ and $\theta_0\in\mathcal{S}^{\mathcal{N}}$) is as follows:

1. The *first half-step* (prediction):

$$
\begin{aligned}
\bar{\omega}_n &= \arg\min_{\omega\geq 0}\left\{\tfrac{1}{2}\|\omega-\omega_n\|^2-\gamma\mathcal{L}_\delta(u_n,\hat{u}_n(u),v_n,\hat{v}_n(v),\lambda_n,\theta_n,\omega,\bar{\xi}_n)\right\}\\
\bar{\xi}_n &= \arg\min_{\xi\geq 0}\left\{\tfrac{1}{2}\|\xi-\xi_n\|^2-\gamma\mathcal{L}_\delta(u_n,\hat{u}_n(u),v_n,\hat{v}_n(v),\lambda_n,\theta_n,\bar{\omega}_n,\xi)\right\}
\end{aligned}
\tag{30}
$$

$$\bar{u}_n = \arg\min_{u \in U} \left\{ \tfrac{1}{2}\|u - u_n\|^2 + \gamma \mathcal{L}_\delta(u, \hat{u}_n(u), v_n, \hat{v}_n(v), \lambda_n, \theta_n, \bar{\omega}_n, \bar{\xi}_n) \right\}$$

$$\bar{\hat{u}}_n = \arg\min_{\hat{u} \in \hat{U}} \left\{ \tfrac{1}{2}\|\hat{u} - \hat{u}_n\|^2 + \gamma \mathcal{L}_\delta(u_n, \hat{u}(u), v_n, \hat{v}_n(v), \lambda_n, \theta_n, \bar{\omega}_n, \bar{\xi}_n) \right\}$$

$$\bar{v}_n = \arg\min_{v \in V} \left\{ \tfrac{1}{2}\|v - v_n\|^2 - \gamma \mathcal{L}_\delta(u_n, \hat{u}_n(u), v, \hat{v}_n(v), \lambda_n, \theta_n, \bar{\omega}_n, \bar{\xi}_n) \right\}$$

$$\bar{\hat{v}}_n = \arg\min_{\hat{v} \in \hat{V}} \left\{ \tfrac{1}{2}\|\hat{v} - \hat{v}_n\|^2 - \gamma \mathcal{L}_\delta(u_n, \hat{u}_n(u), v_n, \hat{v}(v), \lambda_n, \theta_n, \bar{\omega}_n, \bar{\xi}_n) \right\}$$

$$\bar{\lambda}_n = \arg\min_{\lambda \in \mathcal{S}^N} \left\{ \tfrac{1}{2}\|\lambda - \lambda_n\|^2 + \gamma \mathcal{L}_\delta(u_n, \hat{u}_n(u), v_n, \hat{v}_n(v), \lambda, \theta_n, \bar{\omega}_n, \bar{\xi}_n) \right\}$$

$$\bar{\theta}_n = \arg\min_{\theta \in \mathcal{S}^N} \left\{ \tfrac{1}{2}\|\theta - \theta_n\|^2 - \gamma \mathcal{L}_\delta(u_n, \hat{u}_n(u), v_n, \hat{v}_n(v), \lambda_n, \theta, \bar{\omega}_n, \bar{\xi}_n) \right\}.$$

2. The *second* (basic) half-step

$$\begin{aligned}
\omega_{n+1} &= \arg\min_{\omega \geq 0} \left\{ \tfrac{1}{2}\|\omega - \omega_n\|^2 - \gamma \mathcal{L}_\delta(\bar{u}_n, \bar{\hat{u}}_n(u), \bar{v}_n, \bar{\hat{v}}_n(v), \bar{\lambda}_n, \bar{\theta}_n, \omega, \bar{\xi}_n) \right\} \\
\xi_{n+1} &= \arg\min_{\xi \geq 0} \left\{ \tfrac{1}{2}\|\xi - \xi_n\|^2 - \gamma \mathcal{L}_\delta(\bar{u}_n, \bar{\hat{u}}_n(u), \bar{v}_n, \bar{\hat{v}}_n(v), \bar{\lambda}_n, \bar{\theta}_n, \bar{\omega}_n, \xi) \right\} \\
u_{n+1} &= \arg\min_{u \in U} \left\{ \tfrac{1}{2}\|u - u_n\|^2 + \gamma \mathcal{L}_\delta(u, \bar{\hat{u}}_n(u), \bar{v}_n, \bar{\hat{v}}_n(v), \bar{\lambda}_n, \bar{\theta}_n, \bar{\omega}_n, \bar{\xi}_n) \right\} \\
\hat{u}_{n+1} &= \arg\min_{\hat{u} \in \hat{U}} \left\{ \tfrac{1}{2}\|\hat{u} - \hat{u}_n\|^2 + \gamma \mathcal{L}_\delta(\bar{u}_n, \hat{u}(u), \bar{v}_n, \bar{\hat{v}}_n(v), \bar{\lambda}_n, \bar{\theta}_n, \bar{\omega}_n, \bar{\xi}_n) \right\} \\
v_{n+1} &= \arg\min_{v \in V} \left\{ \tfrac{1}{2}\|v - v_n\|^2 - \gamma \mathcal{L}_\delta(\bar{u}_n, \bar{\hat{u}}_n(u), v, \bar{\hat{v}}_n(v), \bar{\lambda}_n, \bar{\theta}_n, \bar{\omega}_n, \bar{\xi}_n) \right\} \\
\hat{v}_{n+1} &= \arg\min_{\hat{v} \in \hat{V}} \left\{ \tfrac{1}{2}\|\hat{v} - \hat{v}_n\|^2 - \gamma \mathcal{L}_\delta(\bar{u}_n, \bar{\hat{u}}_n(u), \bar{v}_n, \hat{v}(v), \bar{\lambda}_n, \bar{\theta}_n, \bar{\omega}_n, \bar{\xi}_n) \right\} \\
\lambda_{n+1} &= \arg\min_{\lambda \in \mathcal{S}^N} \left\{ \tfrac{1}{2}\|\lambda - \lambda_n\|^2 + \gamma \mathcal{L}_\delta(\bar{u}_n, \bar{\hat{u}}_n(u), \bar{v}_n, \bar{\hat{v}}_n(v), \lambda, \bar{\theta}_n, \bar{\omega}_n, \bar{\xi}_n) \right\} \\
\theta_{n+1} &= \arg\min_{\theta \in \mathcal{S}^N} \left\{ \tfrac{1}{2}\|\theta - \theta_n\|^2 - \gamma \mathcal{L}_\delta(\bar{u}_n, \bar{\hat{u}}_n(u), \bar{v}_n, \bar{\hat{v}}_n(v), \bar{\lambda}_n, \theta, \bar{\omega}_n, \bar{\xi}_n) \right\}.
\end{aligned} \tag{31}$$

5. CONVERGENCE ANALYSIS AND UNIQUENESS

The following theorem presents the convergence conditions of (4.3) – (31) and gives the estimate of its rate of convergence for the strong L_p- Stackelberg/Nash equilibrium. As well, we prove that the extraproximal method converges to a unique equilibrium point.

Let us define the following extended vectors

$$\tilde{u} := \begin{pmatrix} u \\ \hat{u} \\ \lambda \end{pmatrix} \in \tilde{U} := U \times \hat{U} \times \mathbb{R}^+, \quad \tilde{z} := \begin{pmatrix} v \\ \hat{v} \\ \theta \\ \xi \\ \omega \end{pmatrix} \in \tilde{Z} := V \times \hat{V} \times \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+.$$

Then, the regularized Lagrange function can be expressed as

$$\tilde{\mathcal{L}}_\delta(\tilde{u},\tilde{z}) := \mathcal{L}_\delta(u_\delta,\hat{u}_\delta,v_\delta,\hat{v}_\delta,\lambda_\delta,\theta_\delta,\xi_\delta,\omega_\delta).$$

The equilibrium point that satisfies (29) can be expressed as

$$\tilde{u}^*_\delta = \arg\min_{\tilde{u}\in\tilde{U}}\left\{\tfrac{1}{2}\|\tilde{u}-\tilde{u}^*_\delta\|^2+\gamma\tilde{\mathcal{L}}_\delta(\tilde{u},\tilde{z}^*_\delta)\right\}$$

$$\tilde{z}^*_\delta = \arg\max_{\tilde{z}\in\tilde{Z}}\left\{\tfrac{1}{2}\|\tilde{z}-\tilde{z}^*_\delta\|^2+\gamma\tilde{\mathcal{L}}_\delta(\tilde{u}^*_\delta,\tilde{z})\right\}.$$

Now let us introduce the following variables

$$\tilde{w}=\begin{pmatrix}\tilde{w}_1\\ \tilde{w}_2\end{pmatrix}\in\tilde{U}\times\tilde{Z},\ \tilde{v}=\begin{pmatrix}\tilde{v}_1\\ \tilde{v}_2\end{pmatrix}\in\tilde{U}\times\tilde{Z}$$

and let define the Lagrangian in term of the previous variables

$$L_\delta(\tilde{w},\tilde{v}) := \tilde{\mathcal{L}}_\delta(\tilde{w}_1,\tilde{v}_2)-\tilde{\mathcal{L}}_\delta(\tilde{v}_1,\tilde{w}_2).$$

For $\tilde{w}_1=\tilde{u}$, $\tilde{w}_2=\tilde{z}$, $\tilde{v}_1=\tilde{v}^*_1=\tilde{u}^*_\delta$ and $\tilde{v}_2=\tilde{v}^*_2=\tilde{z}^*_\delta$ we have

$$L_\delta(\tilde{w},\tilde{v}^*) := \tilde{\mathcal{L}}_\delta(\tilde{u},\tilde{z}^*_\delta)-\tilde{\mathcal{L}}_\delta(\tilde{u}^*_\delta,\tilde{z}).$$

In these variables the relation (29) can be represented by

$$\tilde{v}^*=\arg\min_{\tilde{w}\in\tilde{U}\times\tilde{Z}}\left\{\tfrac{1}{2}\|\tilde{w}-\tilde{v}^*\|^2+\gamma L_\delta(\tilde{w},\tilde{v}^*)\right\}. \tag{32}$$

Finally, we have that the extraproximal method can be expressed by

1. First step
$$\hat{v}_n=\arg\min_{\tilde{w}\in\tilde{U}\times\tilde{Z}}\left\{\tfrac{1}{2}\|\tilde{w}-\tilde{v}_n\|^2+\gamma L_\delta(\tilde{w},\tilde{v}_n)\right\}. \tag{33}$$

2. Second step
$$\tilde{v}_{n+1}=\arg\min_{\tilde{w}\in\tilde{U}\times\tilde{Z}}\left\{\tfrac{1}{2}\|\tilde{w}-\tilde{v}_n\|^2+\gamma L_\delta(\tilde{w},\hat{v}_n)\right\}. \tag{34}$$

Theorem 5.1. Let $\tilde{\mathcal{L}}_\delta(\tilde{u},\tilde{z})$ be differentiable in $\tilde{u}$ and $\tilde{z}$, whose partial derivative with respect to $\tilde{z}$ satisfies the Lipschitz condition with positive constant C_0. Then,

$$\|\tilde{v}_{n+1}-\hat{v}_n\| \le \gamma C_0\|\tilde{v}_n-\hat{v}_n\|. \tag{35}$$

Proof. See [30] □

Theorem 5.2. (Convergence and Uniqueness) Let $\tilde{\mathcal{L}}_\delta(\tilde{u},\tilde{z})$ be differentiable in $\tilde{u}$ and $\tilde{z}$, whose partial derivative with respect to $\tilde{z}$ satisfies the Lipschitz condition with positive constant C. Then, for some δ and

$$C_0^l=\sum\nolimits_{l=1}^{\mathcal{N}} C_{0,l} \ \le\ \mathcal{N}\max_{l=\overline{1,\mathcal{N}}} C_{0,l} \ =\ \mathcal{N}C_0^{l+}$$

and

$$C_0^m = \sum_{m=1}^{\mathcal{M}} C_{0,m} \;\leq\; \mathcal{M} \max_{m=\overline{1,\mathcal{M}}} C_{0,m} \;=\; \mathcal{M} C_0^{m+}$$

there exists a small-enough

$$\gamma_0 = \gamma_0(\delta) < C$$

$$:= \max\left[\min\left\{\tfrac{1}{\sqrt{2}C_0^{l+}\mathcal{N}}, \tfrac{1+\sqrt{1+2\left(C_0^{l+}\right)^2}}{2\left(C_0^{l+}\right)^2\mathcal{N}}\right\}, \min\left\{\tfrac{1}{\sqrt{2}C_0^{m+}\mathcal{M}}, \tfrac{1+\sqrt{1+2\left(C_0^{m+}\right)^2}}{2\left(C_0^{m+}\right)^2\mathcal{M}}\right\}\right]$$

where such that, for any $0 < \gamma \leq \gamma_0$, sequence $\{\tilde{v}_n\}$, which generated by the equivalent extraproximal procedure (4.3) - (31), monotonically converges with exponential rate $q \in (0,1)$ to a unique equilibrium point $\tilde{v}^*$, i. e.,

$$\|\tilde{v}_n - \tilde{v}^*\|^2 \leq e^{n \ln q} \|\tilde{v}_0 - \tilde{v}^*\|^2 \tag{36}$$

where

$$q = 1 + \tfrac{4(\delta\gamma)^2}{1+2\delta\gamma-2\gamma^2C^2} - 2\delta\gamma < 1$$

and $q_{\min}$ is given by

$$q_{\min} = 1 - \tfrac{2\delta\gamma}{1+2\delta\gamma} = \tfrac{1}{1+2\delta\gamma}.$$

Proof. Following Theorem 1 in [30] we obtain that

$$q = 1 - 2\gamma\delta + \tfrac{(2\gamma\delta)^2}{1+2\gamma\delta-2\gamma^2C^2} < 1.$$

Iterating over the previous inequality we have

$$\|\tilde{v}_\delta^* - \tilde{v}_{n+1}\|^2 \leq q \|\tilde{v}_\delta^* - \tilde{v}_n\|^2 \leq \cdots \leq e^{n+1 \ln q} \|\tilde{v}_\delta^* - \tilde{v}_0\|^2. \tag{37}$$

That implies that the series converge and also that the trajectories are bounded. Then, by Eq. (37) we have that

$$\|\tilde{v}_\delta^* - \tilde{v}_{n+1}\|^2 \underset{n\to\infty}{\to} 0.$$

Given that $\tilde{v}$ is a bounded sequence, by the Weierstrass Theorem there exist a point $\tilde{v}'$ such that any subsequence $\tilde{v}_{n_i}$ satisfies that $\tilde{v}_{n_i} \underset{n_i\to\infty}{\to} \tilde{v}'$. In addition, we have that $\|\tilde{v}_{n_i} - \tilde{v}_{n_i+1}\|^2 \to 0$. Fixing, $n = n_i$ in Eq. (32) and computing the limit when $n_i \to \infty$ we have

$$\tilde{v}' = \arg \min_{\tilde{w} \in \tilde{U} \times \tilde{T}} \left\{ \tfrac{1}{2} \|\tilde{w} - \tilde{v}'\|^2 + \gamma L_\delta(\tilde{w}, \tilde{v}') \right\}.$$

Then, we have that $\tilde{v}' = \tilde{v}_\delta^*$, i. e., any limit point of the sequence $\tilde{v}_n$ is a solution of the problem. Given that $\|\tilde{v}_n - \tilde{v}_\delta^*\|^2$ is monotonically decreasing then, there exists a unique limit point (equilibrium point). As a consequence, we have that the sequence $\tilde{v}_n$ satisfies that $\tilde{v}_n \underset{n\to\infty}{\to} \tilde{v}_\delta^*$ with a convergence velocity of $e^{n \ln q}$. □

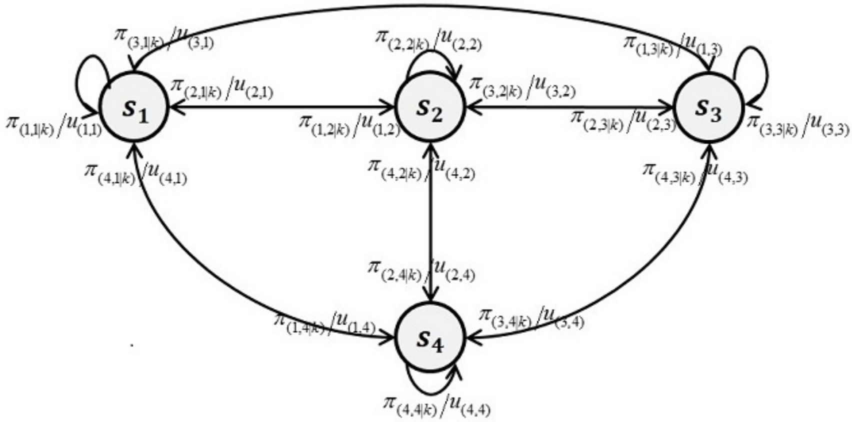

Fig. 1. Supermarket Markov Chain.

6. APPLICATION EXAMPLE

This example analyzes the effectiveness of relationship marketing strategies within the department store sector of the retail industry considering two supermarket leaders with $l = 1, 2$ and two supermarkets followers with $m = 3, 4$. The three supermarkets are branching out into non-food items and they are also department stores in their own right, selling items as clothes, entertainment products for example toys, books, cosmetics, non-prescription drugs and many other household goods. All the supermarkets offer loyalty cards having their own system with the purpose to attract customers, encourage customer loyalty and build strong customer relationships. As well, loyalty cards create an advantage for supermarkets developing profiles of individuals' personal shopping habits. When linked with the personal details that customers disclosed when signing up for the scheme, the store is in a position to target promotions that are tailored around specific customers shopping habits. Based on the available data, supermarkets discretize the client space in four sub-segments according to the regularly of purchasing, using frequency of the loyalty card and the revenue. Figure 1 describes the segments and promotions corresponding to the Markov chain of the marketing problem. Here a customer is said to be in state s_1 if he/she become a Potential customer. A Low-Frequent customer corresponds with the state s_2 and a Regular customer is a frequent customer of the loyalty card that is said to be in state s_3. A Loyal Customer corresponds with the state s_4 and he/she is a high-frequency user of the loyal card. The promotions (actions) offered by the supermarkets include two different benefits: 1) points and 2) discounts. We are interested in contrasting the strategies applied by the supermarkets defined over all possible combinations of states (i, j) and actions (k) given a fixed utility $J_{(i,j,k)}$.

Our goal is to analyze a four-player Stackelberg game for the norm $p = 1$ in a class of ergodic controllable finite Markov chains. Let $N_1 = N_2 = N_3 = N_4 = 4$, $M_1 = M_2 = M_3 = M_4 = 2$. Following Eq. (3) the individual utility for each player are defined by

$$J^{(1)}_{ij|1}=\begin{bmatrix}567 & 822 & 733 & 830\\ 261 & 896 & 85 & 568\\ 30 & 996 & 634 & 261\\ 288 & 90 & 806 & 785\end{bmatrix}\quad J^{(1)}_{ij|2}=\begin{bmatrix}170 & 27 & 57 & 699\\ 275 & 855 & 224 & 919\\ 50 & 205 & 46 & 909\\ 398 & 861 & 751 & 806\end{bmatrix}$$

$$J^{(2)}_{ij|1}=\begin{bmatrix}810 & 36 & 27 & 9\\ 63 & 90 & 567 & 72\\ 81 & 0 & 9 & 45\\ 855 & 594 & 441 & 9\end{bmatrix}\quad J^{(2)}_{ij|2}=\begin{bmatrix}8 & 592 & 48 & 0\\ 64 & 64 & 312 & 16\\ 264 & 32 & 120 & 72\\ 400 & 56 & 40 & 200\end{bmatrix}$$

$$J^{(3)}_{ij|1}=\begin{bmatrix}22 & 7 & 11 & 6\\ 10 & 0 & 19 & 8\\ 23 & 28 & 23 & 9\\ 90 & 5 & 12 & 1\end{bmatrix}\quad J^{(3)}_{ij|2}=\begin{bmatrix}66 & 0 & 126 & 42\\ 18 & 78 & 240 & 6\\ 96 & 18 & 60 & 156\\ 66 & 102 & 180 & 48\end{bmatrix}$$

$$J^{(4)}_{ij|1}=\begin{bmatrix}0 & 60 & 2 & 26\\ 10 & 26 & 36 & 48\\ 14 & 56 & 28 & 24\\ 8 & 12 & 16 & 38\end{bmatrix}\quad J^{(4)}_{ij|2}=\begin{bmatrix}420 & 168 & 378 & 84\\ 0 & 280 & 14 & 112\\ 42 & 56 & 350 & 140\\ 84 & 210 & 336 & 98\end{bmatrix}.$$

The transition matrices for each player are defined as follows

$$\pi^{(1)}_{ij|1}=\begin{bmatrix}0.2759 & 0.4886 & 0.0366 & 0.1989\\ 0.1752 & 0.0953 & 0.3825 & 0.3470\\ 0.1695 & 0.2629 & 0.4103 & 0.1574\\ 0.2612 & 0.1665 & 0.4124 & 0.1600\end{bmatrix}\quad \pi^{(1)}_{ij|2}=\begin{bmatrix}0.0863 & 0.3672 & 0.3201 & 0.2264\\ 0.4339 & 0.1684 & 0.1919 & 0.2058\\ 0.3856 & 0.2349 & 0.1324 & 0.2471\\ 0.1475 & 0.3500 & 0.1903 & 0.3122\end{bmatrix}$$

$$\pi^{(2)}_{ij|1}=\begin{bmatrix}0.1761 & 0.1204 & 0.3883 & 0.3151\\ 0.2207 & 0.1632 & 0.2354 & 0.3807\\ 0.0708 & 0.3708 & 0.1364 & 0.4219\\ 0.0132 & 0.5169 & 0.4127 & 0.0572\end{bmatrix}\quad \pi^{(2)}_{ij|2}=\begin{bmatrix}0.2033 & 0.2456 & 0.2667 & 0.2844\\ 0.2732 & 0.1032 & 0.3046 & 0.3190\\ 0.1207 & 0.0930 & 0.3997 & 0.3866\\ 0.1032 & 0.6976 & 0.1609 & 0.0383\end{bmatrix}$$

$$\pi^{(3)}_{ij|1}=\begin{bmatrix}0.4109 & 0.1654 & 0.0918 & 0.3319\\ 0.3015 & 0.2201 & 0.1029 & 0.3756\\ 0.1709 & 0.5673 & 0.0292 & 0.2326\\ 0.1885 & 0.1491 & 0.3317 & 0.3307\end{bmatrix}\quad \pi^{(3)}_{ij|2}=\begin{bmatrix}0.3046 & 0.2883 & 0.2573 & 0.1498\\ 0.2470 & 0.0978 & 0.3060 & 0.3492\\ 0.3006 & 0.0439 & 0.4387 & 0.2169\\ 0.1141 & 0.3397 & 0.1855 & 0.3607\end{bmatrix}$$

$$\pi^{(4)}_{ij|1}=\begin{bmatrix}0.2610 & 0.3145 & 0.2088 & 0.2158\\ 0.3777 & 0.1968 & 0.1574 & 0.2681\\ 0.2593 & 0.0308 & 0.5113 & 0.1986\\ 0.3401 & 0.4638 & 0.1200 & 0.0761\end{bmatrix}\quad \pi^{(4)}_{ij|2}=\begin{bmatrix}0.0316 & 0.4652 & 0.2221 & 0.2811\\ 0.1624 & 0.3245 & 0.3691 & 0.1440\\ 0.1448 & 0.5777 & 0.2087 & 0.0688\\ 0.2536 & 0.1996 & 0.3231 & 0.2237\end{bmatrix}.$$

Given δ and γ and applying the extraproximal method we obtain the convergence of the strategies in terms of the variable $c^{l}_{i|k}$ for the leaders (see Figure 2) and the convergence of the strategies in terms of the variable $c^{m}_{i|k}$ for the followers (see Figure 3). In addition, the Figure 4 shows the convergence of the parameters Ji and Omega.

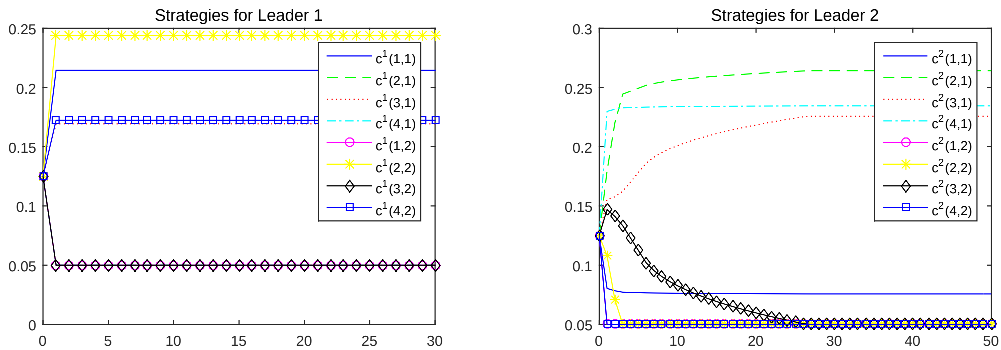

Fig. 2. Convergence of the strategies of the leader 1 (left) leader 2 (right).

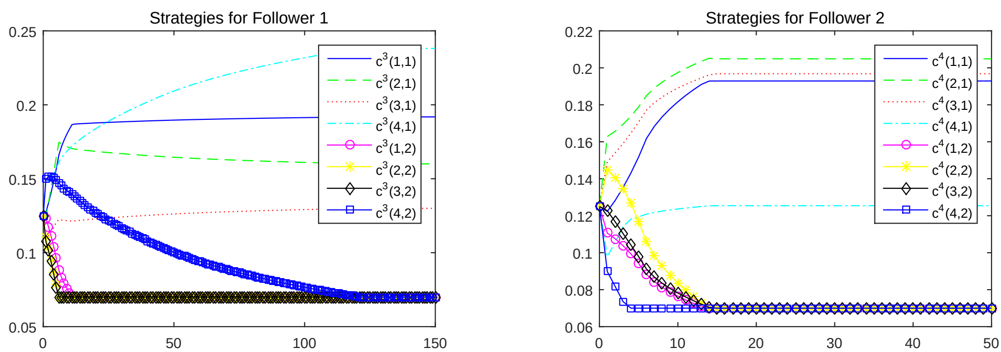

Fig. 3. Convergence of the strategies of the follower 1 (left) and follower 2 (right).

With final values $\lambda^{(1)*} = 0.5063$ and $\lambda^{(2)*} = 0.4937$ for the leaders, and $\theta^{(1)*} = 0.5258$ and $\theta^{(2)*} = 0.4792$ for the followers (see Figure 5), the mixed strategies obtained for determining the strong Stackelberg/Nash equilibrium for all the players applying (9) are as follows

$$d^{(1)*} = \begin{bmatrix} 0.8110 & 0.1890 \\ 0.1701 & 0.8299 \\ 0.7720 & 0.2280 \\ 0.2249 & 0.7751 \end{bmatrix} \qquad d^{(2)*} = \begin{bmatrix} 0.6023 & 0.3977 \\ 0.8408 & 0.1592 \\ 0.8187 & 0.1813 \\ 0.8242 & 0.1758 \end{bmatrix} \tag{38}$$

$$d^{(3)*} = \begin{bmatrix} 0.6478 & 0.3522 \\ 0.7078 & 0.2922 \\ 0.6455 & 0.3545 \\ 0.6442 & 0.3558 \end{bmatrix} \qquad d^{(4)*} = \begin{bmatrix} 0.7337 & 0.2663 \\ 0.7454 & 0.2546 \\ 0.7376 & 0.2624 \\ 0.6418 & 0.3582 \end{bmatrix}.$$

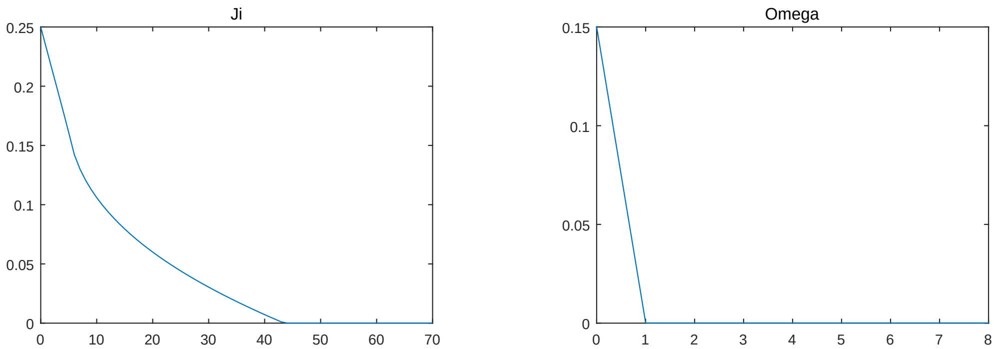

Fig. 4. Convergence of the parameters Ji and Omega.

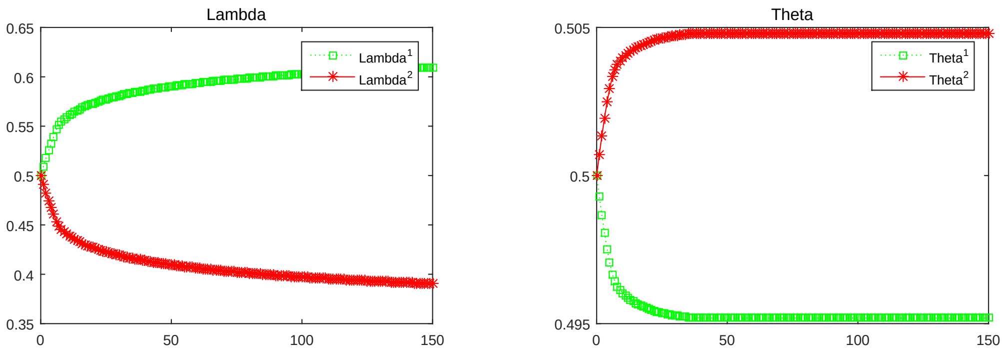

Fig. 5. Convergence of Lambda and Theta.

The resulting utilities by segment are as follows:

$$J^{(1)}(s_i) = \begin{bmatrix} 129,130 \\ 92,790 \\ 84,590 \\ 121,520 \end{bmatrix} \qquad J^{(2)}(s_i) = \begin{bmatrix} 13,102 \\ 22,635 \\ 1,113 \\ 64,809 \end{bmatrix} \tag{39}$$

$$J^{(3)}(s_i) = \begin{bmatrix} 551 \\ 1,295 \\ 746 \\ 1,494 \end{bmatrix} \qquad J^{(4)}(s_i) = \begin{bmatrix} 3,914 \\ 2,113 \\ 2,158 \\ 3,467 \end{bmatrix}. \tag{40}$$

The resulting utilities by promotion are as follows:

$$J^{(1)}(k_i) = \begin{bmatrix} 226,830 & 201,190 \end{bmatrix} \qquad J^{(2)}(k_i) = \begin{bmatrix} 93,930 & 7,729 \end{bmatrix} \tag{41}$$

$$J^{(3)}(k_i) = \begin{bmatrix} 437 & 3,650 \end{bmatrix} \qquad J^{(4)}(k_i) = \begin{bmatrix} 609 & 11,045 \end{bmatrix}. \tag{42}$$

Relationship marketing recognizes that the focus of marketing is to build a relationship with existing customers. The main purpose of the game is to discover the extent to which customers use and are influenced by relationship marketing strategies. In addition, it is to analyze the impact that these strategies have on customer loyalty and the development of customer-department store relationship. The supermarket leaders (players 1 and 2) fix their strategies (38) to ensure high degrees of customer loyalty and retention as well utility by segment (39) and promotion. For segment 1, the leader 1 made a strong emphasis on offering points (0.8110) for attracting Potential customers. Instead, the leader 2 made emphasis on offering points (0.6023) and discounts (0.3977) for the same segment. Looking at the utilities of the leaders (39), the follower1 decided for offering points (0.6478) and discounts (0.3522). Instead, the follower 2 resolved for competing highlighting points (0.7337). For segment 2 corresponding to Low-Frequent customers the leader 1 promoted points (0.1701) and discounts (0.8299) and, the leader 2 chose offering points (0.8408) and discounts (0.1592). However, for competing with the leaders, follower 1 and follower 2 made emphasis on points (0.7078 and 0.7454 respectively). For Regular customers the leader 1 focused on points (0.7720) and discounts (0.2280) and, the leader 2 made emphasis on points (0.8187). The follower 1 preferred offering points (0.6455) and discounts (0.3545). Instead, follower 2 made emphasis on points (0.7376) and discounts (0.2624). For Loyal customers the leader 1 made emphasis on points (0.2249) and discounts (0.7751), leader 2 focus on points (0.8242) and discounts (0.1758) as well, follower 1 chose the same strategies – points (0.6442) and discounts (0.3558) –. The follower 2 made emphasis on points (0.6418) and discounts (0.3582). For the leaders the most profitable segments are the Potential customers and the Loyal customers (see 39 vs. 40). An insight into the mind of the consumer is obvious from the findings the importance that is placed on a given policy: the utilities obtained by action for the leaders and followers are shown in Eqs. 41 and 42 respectively.

7. CONCLUSION

In this paper we presented a novel approach in Markov games for computing the strong L_p-Stackelberg/Nash equilibrium in case of a metric state space (determined by the positive orthant in the Euclidian space). The existence of the L_p-Stackelberg/Nash equilibrium was characterized as a strong Pareto policy, which is the closest in the Euclidean norm to the utopian point. The optimization problem was reduced to find a Pareto solution. We proposed to employ the extraproximal approach for solving the problem and converge to a strong L_p- Stackelberg/Nash equilibrium. We developed the method for the Stackelberg game in terms of nonlinear programming problems implementing the Lagrange principle and employed the Tikhonov's regularization method to ensure the convergence to a unique strong Stackelberg/Nash equilibrium. Finally, we applied our approach in a numerical example related to the effectiveness of relationship marketing strategies for supermarkets showing the usefulness of the solution.

(Received February 19, 2015)

REFERENCES

[1] E. Aiyoshi and K. Shimizu: Hierarchical decentralized systems and its new solution by abarrier method. IEEE Trans. Systems, Man, and Cybernet. *11* (1981), 444–449.

DOI:10.1109/tsmc.1981.4308712

[2] A. S. Antipin: An extraproximal method for solving equilibrium programming problems and games. Comput. Math. and Math. Phys. *45* (2005), 11, 1893–1914.

[3] J. Bard: Coordination of a multidivisional organization through two levels of management. Omega *11* (1983), 5, 457–468. DOI:10.1016/0305-0483(83)90038-5

[4] J. Bard: Practical Bilevel Optimization: Algorithms and Applications, vol 30. Kluwer Academic, Dordrecht 1998. DOI:10.1007/978-1-4757-2836-1

[5] J. Bard and J. Falk: An explicit solution to the multi-level programming problem. Comput. Oper. Res. *9* (1982), 77–100. DOI:10.1016/0305-0548(82)90007-7

[6] L. Bianco, M. Caramia, and S. Giordani: A bilevel flow model for hazmat transportation network design. Transport. Res. Part C: Emerging Technol. *17* (2009), 2, 175–196. DOI:10.1016/j.trc.2008.10.001

[7] L. Brotcorne, M. Labb, P. Marcotte, and G. Savard: A bilevel model for toll optimization on a multicommodity transportation network. Transport. Sci. *35* (2001), 345–358. DOI:10.1287/trsc.35.4.345.10433

[8] J. B. Clempner and A. S. Poznyak: Simple computing of the customer lifetime value: A fixed local-optimal policy approach. J. Systems Sci. and Systems Engrg. *23* (2014), 4, 439–459. DOI:10.1007/s11518-014-5260-y

[9] J. B. Clempner and A. S. Poznyak: Stackelberg security games: Computing the shortest-path equilibrium. Expert Systems Appl. *42* (2015), 8, 3967–3979. DOI:10.1016/j.eswa.2014.12.034

[10] J. Cote, P. Marcotte, and G. Savard: A bilevel modelling approach to pricing and fare optimisation in the airline industry. J. Revenue and Pricing Management *2* (2003), 1, 23–36. DOI:10.1057/palgrave.rpm.5170046

[11] K. Deb and A. Sinha: An efficient and accurate solution methodology for bilevel multi-objective programming problems using a hybrid evolutionary-local-search algorithm. Evolutionary Comput. J. *18* (2010), 3, 403–449. DOI:10.1162/evco_a_00015

[12] S. Dempe: Discrete Bilevel Optimization Problems. Technical ReportInstitut fur Wirtschaftsinformatik, Universitat Leipzig 2001.

[13] S. Dempe, V. Kalashnikov, and R Rios-Mercado: Discrete bilevel programming: Application to a natural gas cash-out problem. Europ. J. Oper. Res. *166* (2005), 2, 469–488. DOI:10.1016/j.ejor.2004.01.047

[14] S. DeNegre and T. Ralphs: A branch-and-cut algorithm for integer bilevel linear programs. Oper. Res. Cyber-Infrastruct. *47* (2009), 65–78. DOI:10.1007/978-0-387-88843-9_4

[15] M. Fampa, L. Barroso, D. Candal, and L. Simonetti: Bilevel optimization applied to strategic pricing in competitive electricity markets. Comput. Optim. Appl. *39* (2008), 2, 121–142. DOI:10.1007/s10589-007-9066-4

[16] J. Fortuny-Amat and B. McCarl: A representation and economic interpretation of a two-level programming problem. J. Oper. Res. Soc. *xx* (1981), 783–792. DOI:10.1057/jors.1981.156

[17] Y. B. Germeyer: Introduction to the Theory of Operations Research. Nauka, Moscow 1971.

[18] Y. B. Germeyer: Games with Nonantagonistic Interests. Nauka, Moscow 1976.

[19] J. Herskovits, A. Leontiev, G. Dias, and G. Santos: Contact shape optimization: a bilevel programming approach. Struct. Multidiscipl. Optim. *20* (2000), 214–221. DOI:10.1007/s001580050149

[20] M. Koppe, M. Queyranne, and C. T. Ryan: A parametric integer programming algorithm for bilevel mixed integer programs. J. Optim. Theory Appl. *146* (2009), 1, 137–150. DOI:10.1007/s10957-010-9668-3

[21] M. Labbe, P. Marcotte, and G. Savard: A bilevel model of taxation and its application to optimal highway pricing. Management Sci. *44* (1998), 1608–1622.

[22] C. Lim and J. Smith: Algorithms for discrete and continuous multicommodity flow network interdiction problems. IIE Trans. *39* (2007), 1, 15–26. DOI:10.1287/mnsc.44.12.1608

[23] D. Morton, F. Pan, and K. Saeger: Models for nuclear smuggling interdiction. IIE Trans. *39* (2007), 1, 3–14. DOI:10.1080/07408170500488956

[24] J. Naoum-Sawaya and S. Elhedhli: Controlled predatory pricing in a multiperiod stackelberg game: an mpec approach. J. Global Optim. *50* (2011), 345–362. DOI:10.1007/s10898-010-9585-x

[25] A. S. Poznyak: Advance Mathematical Tools for Automatic Control Engineers. Vol 2 Deterministic Techniques. Elsevier, Amsterdam 2009.

[26] A. S. Poznyak, K. Najim, and E. Gomez-Ramirez: Self-Learning Control of Finite Markov Chains. Marcel Dekker, New York 2000.

[27] J. Salmeron, K. Wood, and R. Baldick: Analysis of electric grid security under terrorist threat. IEEE Trans. Power Syst. *19* (2004), 2, 905–912. DOI:10.1109/tpwrs.2004.825888

[28] K. Tanaka: The closest solution to the shadow minimum of a cooperative dynamic game. Comp. Math. Appl. *18* (1989), 1–3, 181–188. DOI:10.1016/0898-1221(89)90135-1

[29] K. Tanaka and K. Yokoyama: On ϵ-equilibrium point in a noncooperative n-person game. J. Math. Anal. Appl. *160* (1991), 413–423. DOI:10.1016/0022-247x(91)90314-p

[30] K. K. Trejo, J. B. Clempner, and A. S. Poznyak: Computing the Stackelberg/Nash equilibria using the extraproximal method: Convergence analysis and implementation details for Markov chains games. Int. J. Appl. Math. Computer Sci. *25* (2015), 2, 337–351. DOI:10.1515/amcs-2015-0026

[31] K. K. Trejo, J. B. Clempner, and A. S. Poznyak: A stackelberg security game with random strategies based on the extraproximal theoretic approach. Engrg. Appl. Artif. Intell. *37* (2015), 145–153. DOI:10.1016/j.engappai.2014.09.002

Kristal K. Trejo, Department of Control Automatics, Center for Research and Advanced Studies Av. IPN 2508, Col. San Pedro Zacatenco, 07360 Mexico City. Mexico.
e-mail: ktrejo@ctrl.cinvestav.mx

Julio B. Clempner, Centro de Investigaciones Economicas, Administrativas y Sociales, Instituto Politecnico Nacional, Lauro Aguirre 120, col. Agricultura, Del. Miguel Hidalgo, 11360, Mexico City. Mexico.
e-mail: julio@clempner.name

Alexander S. Poznyak, Department of Control Automatics, Center for Research and Advanced Studies, Av. IPN 2508, Col. San Pedro Zacatenco, 07360 Mexico City. Mexico.
e-mail: apoznyak@ctrl.cinvestav.mx