Preprints of the 20th World Congress
The International Federation of Automatic Control
Toulouse, France, July 9-14, 2017

Nash Bargaining Equilibria for Controllable Markov Chains Games

Kristal K. Trejo * **Julio B. Clempner** ** **Alexander S. Poznyak** *

* *Department of Automatic Control, CINVESTAV-IPN, Av. IPN 2508, Col. San Pedro Zacatenco, 07360 Mexico City, Mexico, (e-mail: ktrejo(apoznyak)@ctrl.cinvestav.mx).*
** *Centro de Investigaciones Económicas, Administrativas y Sociales, Instituto Politécnico Nacional, Lauro Aguirre 120, col. Agricultura, Del. Miguel Hidalgo, 11360, Ciudad de México, México, (e-mail: julio@clempner.name)*

Abstract: A classical bargaining situation involves two individuals who have the opportunity to collaborate for mutual benefit. In this paper we present a novel approach for computing the Nash bargaining equilibrium for controllable Markov chains games. We follow the solution introduced by Nash considering the disagreement point as the Nash equilibrium of the problem. For solving the bargaining process we consider the game formulation in terms of nonlinear programming equations implementing the regularized Lagrange method. For computing the equilibrium point we employ the extraproximal optimization approach. We present the convergence and rate of convergence of the method. Finally a numerical example for a two-person bargaining situation illustrates the effectiveness of the method.

Keywords: Controlled Markov chains, bargaining, optimization, Nash, game theory.

1. INTRODUCTION

The starting point of bargaining theory is the Nash [1950] formulation, who presented the bargaining situation as a new treatment of a classical economic problem. A two-person bargaining situation involves two individuals who have the opportunity to collaborate for mutual benefit. This is defined as a pair (L, f^*) where L is a compact, convex subset of $\mathbb{R}^2$ containing both f^* and a point that strictly dominates f^*. Points $f = (f_1, f_2) \in L$ represents levels of utility for players 1 and 2 that can be reached by an outcome of the game which is feasible for the two players when they do cooperate, and $f^* = (f_1^*, f_2^*)$ is the level of utility that players receive if the two players do not cooperate with each other. The goal is to find an outcome in B which will be agreeable to both players.

There are many models that present an extension to the Nash model for delayed agreements. Perles and Maschler [1981] proposed a solution concept with the property that its followers will always prefer to reach an immediate agreement, and Rubinstein [1982] showed a model where every player bears fixed bargaining cost for each period.

Nash proved that a solution for all convex bargaining problems always maximizes the product of individuals' utilities under four axioms that describe the behavior of players and provide a unique solution: Symmetry, Pareto optimality, Invariance with respect to Positive Affine Transformations, and Independence of Irrelevant Alternatives; however, the Kalai and Smorodinsky [1975] solution presented the axiom of Monotonicity, which leads to another unique solution. In the same way the Dagan et al. [2002] characterization replaced the axiom of Independence of Irrelevant Alternatives with three independent axioms.

There are several applications of the bargaining theory with Markov chains in the economy area. Kalandrakis [2004] analyzed an infinitely repeated divide-the-dollar bargaining game, in each period a new dollar is divided among three legislators according to the proposal of a randomly recognized member or according to previous period's allocation otherwise. Cripps [1998] considered an alternating offer bargaining game which is played by a risk neutral buyer and seller. Kennan [2001] analysed repeated contract negotiations involving the same buyer and seller where the contracts are linked because the buyer has persistent private information. Coles and Muthoo [2003] studied an alternating offers Nash bargaining model in which the set of possible utility pairs evolves through time in a non-stationary, but smooth manner.

The objective of this paper is to present a novel method for computing the Nash bargaining solution in a class of controlled Markov chains games. We solve the disagreement point considering the Nash equilibrium point; for the bargaining solution we present the Nash model in terms of coupled nonlinear programming problems implementing the regularized Lagrange principle. For solving the problem we employ the extraproximal method. The usefulness of the method is demonstrated by a numerical example.

The remainder of the paper is organized as follows. The next section presents the Nash bargaining model and establishes the formulation and solution of the bargaining problem. A numerical example that validates the proposed method, including the solution for the disagreement point as a Nash equilibrium is presented in Section 3. Final comments are outlined in Section 4.

2. THE BARGAINING MODEL

Nash bargaining solution is based on a model in which the players are assumed to negotiate on which point of the set of feasible payoffs $L \subset \mathbb{R}^{\mathcal{N}}$ will be agreed upon and realized by concerted actions of the members of the coalition $l = 1, ..., \mathcal{N}$. A pivotal element of the model is a fixed disagreement vector $f^* \in \mathbb{R}^{\mathcal{N}}$ which plays the role of a deterrent. The players are committed to the disagreement point in the case of failing to reach a consensus on which feasible payoff to realize. Thus the whole bargaining problem B will be concisely given by the pair $B = (L, f^*)$. We will call this form the condensed form of the bargaining problem (see Nash [1950], Forgó et al. [1999]).

A bargaining problem can be derived from the normal form of an $\mathcal{N}$-person game $G = C_1, ..., C_{\mathcal{N}}; f_1, ..., f_{\mathcal{N}}$. The set of all feasible payoffs is defined as $F = f : f = (f_1(c), ..., f_{\mathcal{N}}(c)), c \in C$ where $C = C_1 \times ... \times C_{\mathcal{N}}$.

Given a disagreement vector $f^* \in \mathbb{R}^{\mathcal{N}}$, $B = (F, f^*)$ is a bargaining problem in condensed form. We can derive another bargaining problem $B = (L, f^*)$ from G by extending the set of feasible outcomes F to its convex hull L. Notice that any element $\varphi \in L$ can be represented as

$$\varphi = \sum_{k=1}^{m} \lambda_k f_k, \qquad (m \leq \mathcal{N} + 1),$$

where $f_k = (f_1(c), ..., f_{\mathcal{N}}(c)), (c \in C)$, $\lambda_k \geq 0$ for all k, and $\sum_{k=1}^{m} \lambda_k = 1$.

The payoff vector φ can be realized by playing the strategies c_k with probability λ_k, and so φ is the expected payoff to the players. Thus, when the players face the bargaining problem B the question is, which point of L should be selected taking into account the different position and strength of the players that is reflected in the set L of extended payoffs and the disagreement point f^*.

Nash approached this problem by assigning a one-point solution to B in an axiomatic manner. Let B denote the set of all pairs (L, f^*) such that

(1) $L \subset \mathbb{R}^{\mathcal{N}}$ is compact, convex;
(2) there exists at least one $f \in L$ such that $f > f^*$.

A Nash solution to the bargaining problem is a function $\psi : B \to \mathbb{R}^{\mathcal{N}}$ such that $\psi(L, f^*) \in L$. We shall confine ourselves to functions satisfying the following axioms and we still call there functions solution (see Nash [1950], Forgó et al. [1999], Muthoo [2002]).

(1) Feasibility: $\psi(L, f^*) \in L$.
(2) Rationality: $\psi(L, f^*) \geq f^*$.
(3) Pareto Optimality: For every $(L, f^*) \in B$ there is $f \in L$ such that $f \geq \psi(L, f^*)$ and imply $f = \psi(L, f^*)$.
(4) Symmetry: If for a bargaining problem $(L, f^*) \in B$, there exist indices i, j such that $\varphi = (\varphi_1, ..., \varphi_{\mathcal{N}}) \in L$ if and only if $\bar{\varphi} = (\bar{\varphi}_1, ..., \bar{\varphi}_{\mathcal{N}}) \in L$, $(\bar{\varphi}_k = \varphi_k, k \neq i, k \neq j, \bar{\varphi}_i = \varphi_j, \bar{\varphi}_j = \varphi_i)$ and $f_i^* = f_j^*$ for $f^* = (f_1^*, ..., f_{\mathcal{N}}^*)$, then $\psi_i = \psi_j$ for the solution vector $\psi(L, f^*) = (\psi_1, ..., \psi_{\mathcal{N}})$.
(5) Invariance with respect to affine transformations of utility: Let $\alpha_k > 0, \beta_k, (k = 1, ..., n)$ be arbitrary constants and let
$$f^{*\prime} = (\alpha_1 f_1^* + \beta_1, ..., \alpha_{\mathcal{N}} f_{\mathcal{N}}^* + \beta_{\mathcal{N}})$$
with $f^* = (f_1, ..., f_{\mathcal{N}})$ and
$L' = (\alpha_1 \varphi_1 + \beta_1, ..., \alpha_{\mathcal{N}} \varphi_{\mathcal{N}} + \beta_{\mathcal{N}}) : (\varphi_1, ..., \varphi_{\mathcal{N}}) \in L$.
Then $\psi(L', f^{*\prime}) = (\alpha_1 \psi_1 + \beta_1, ..., \alpha_{\mathcal{N}} \psi_{\mathcal{N}} + \beta_{\mathcal{N}})$, where $\psi(L, f^*) = (\psi_1, ..., \psi_{\mathcal{N}})$.
(6) Independence of irrelevant alternatives: If (L, f^*) and (T, f^*) are bargaining pairs such that $L \subset T$ and $\psi(T, f^*) \in L$, then $\psi(T, f^*) = \psi(L, f^*)$.

Theorem 1. There is a unique function ψ satisfying axioms 1-6, furthermore for all $(L, f^*) \in B$, the vector $\psi(L, f^*) = (\psi_1, ..., \psi_{\mathcal{N}})$ is the unique solution of the optimization problem

$$\begin{aligned} &\text{maximize } \ g(\psi) = \prod_{k=1}^{\mathcal{N}} (\psi_k - f_k^*) \\ &\text{subject to } \quad \psi \in L, \psi \geq f^* \end{aligned} \tag{1}$$

The objective function of problem (1) is usually called the Nash product.

Proof. See Forgó et al. [1999]

2.1 Formulation of the problem

Let S be a finite set, called the *state space*, consisting of finite set of states $\{s_{(1)}, ..., s_{(N)}\}$, $N \in \mathbb{N}$. A *Stationary Markov chain* (see Clempner and Poznyak [2014]) is a sequence of S-valued random variables $s(n)$, $n \in \mathbb{N}$, satisfying the *Markov condition*:

$$P\left(s(n+1) = s_{(j)} | s(n) = s_{(i)}\right) =: \pi_{(i,j)} \tag{2}$$

Following Poznyak et al. [2000], a *controllable Markov chain* is a 4-tuple

$$MC = \{S, A, \mathbb{K}, \Pi\} \tag{3}$$

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 $s \in S$;
- $\mathbb{K} = \{(s, a) | s \in S, a \in A(s)\}$ is the set of admissible state-action pairs;
- $\Pi_{(k)} = \left[\pi_{(i,j|k)}\right]$ is a stationary controlled transition matrix, where
$\pi_{(i,j|k)} \equiv P\left(s(n+1) = s_{(j)} | s(n) = s_{(i)}, a(n) = a_{(k)}\right)$
represents the probability associated with the transition from state $s_{(i)}$ to state $s_{(j)}$ under an action $a_{(k)} \in A\left(s_{(i)}\right)$, $k = 1, ..., M$.

A *Markov Decision Process* is a pair

$$MDP = \{MC, U\} \tag{4}$$

where:

- MC is a controllable Markov chain (3)
- $U : S \times \mathbb{K} \to \mathbb{R}$ is a utility function, associating to each state a real value.

The strategy (policy)

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

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

The game for Markov chains consists of $\mathcal{N}$ players ($l = \overline{1,\mathcal{N}}$) and begins at the initial state $s^l(0)$ which is assumed to be completely measurable. Each of the players l is allowed to randomize, with distribution $d^l_{(k|i)}(n)$ over the action choices $a^l_{(k)}$, $i = \overline{1,N_l}$ and $k = \overline{1,M_l}$. From now on, we will consider only stationary strategies $d^l_{(k|i)}(n) = d^l_{(k|i)}$. When all Markov chains are ergodic for any stationary strategy $d^l_{(k|i)}$ the distributions $P^l(s^l(n+1) = s_{(j_l)})$ exponentially quickly converge to their limits $P^l\left(s = s_{(i)}\right)$ satisfying

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

A $\mathcal{N}$-person bargaining game is a situation in which $\mathcal{N}$ players have a common interest to cooperate, but have conflicting interests over exactly how to cooperate. This process involves the players making offers and counter-offers to each other. Let us denote the disagreement utility that depends on the strategies $c^l_{(i_l,k_l)}$ as $\psi^*_l(c^1,..,c^{\mathcal{N}})$ for each player ($l = 1,..,\mathcal{N}$), and the solution for the Nash bargaining problem as the point $(\psi_1,...,\psi_{\mathcal{N}})$. The utilities $\psi_l = \psi_l\left(c^1,..,c^{\mathcal{N}}\right)$, as well as the disagreement utilities, for Markov chains are defined as follows

$$\psi_l\left(c^1,..,c^{\mathcal{N}}\right) := \sum_{i_1,k_1}^{N_1,M_1} .. \sum_{i_{\mathcal{N}},k_{\mathcal{N}}}^{N_{\mathcal{N}},M_{\mathcal{N}}} W^l_{(i_1,k_1,..,i_{\mathcal{N}},k_{\mathcal{N}})}\prod_{l=1}^{\mathcal{N}}c^l_{(i_l,k_l)} \quad (5)$$

where

$$c^l_{(i_l,k_l)} = d^l_{(i_l,k_l)}P\left(s^l_{(i)}\right)$$

and

$$W^l_{(i_1,k_1,..,i_{\mathcal{N}},k_{\mathcal{N}})} = \sum_{j_1}^{N_1} .. \sum_{j_{\mathcal{N}}}^{N_{\mathcal{N}}} J^l_{(i_1,j_1,k_1,..,i_{\mathcal{N}},j_{\mathcal{N}},k_{\mathcal{N}})}\prod_{l=1}^{\mathcal{N}}\pi^l_{(i_l,j_l|k_l)}$$

where J^l represent the utilities matrices of each player.

The bargaining solution is better than the disagreement point, therefore must satisfy that $\psi_l > \psi^*_l$. The process to solve the bargaining problem consists of two main steps, firstly to find the disagreement point we define it as the Nash equilibrium point of the problem (see Nash [1951]); while for the solution of the bargaining process we follow the model presented by Nash [1950].

The function for finding the solution to the Nash Bargaining problem is

$$g(c^1,...,c^{\mathcal{N}}) = \prod_{l=1}^{\mathcal{N}}(\psi_l - \psi^*_l)^{\alpha^l\chi(\psi_l>\psi^*_l)} \quad (6)$$

where $\alpha^l \geq 0$ ($l = 1,..,\mathcal{N}$), which are weighting parameters for each player. We can rewrite (6) for purposes of implementation as follows

$$\tilde{g}(c^1,...,c^{\mathcal{N}}) = \sum_{l=1}^{\mathcal{N}}\alpha^l\chi(\psi_l > \psi^*_l)\ln(\psi_l - \psi^*_l)$$

The strategy $x^* = (c^1,...,c^{\mathcal{N}}) \in X_{adm} := \bigotimes_{l=1}^{\mathcal{N}} C^l_{adm}$, is the solution for the Nash bargaining problem

$$x^* \in \ arg \max_{x\in X_{adm}} \ \left\{\tilde{g}(c^1,..,c^{\mathcal{N}})\right\}$$

where the strategies c^l satisfy the restrictions C^l_{adm}

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

Applying the Lagrange principle, (Poznyak et al. [2000])

$$\mathcal{L}_\delta(x,\mu,\eta) = \tilde{g}(c^1,...,c^{\mathcal{N}}) - \sum_{l=1}^{\mathcal{N}}\sum_{j_l=1}^{N_l}\mu^l_{(j_l)}h^l_{(j_l)}(c^l) - \sum_{l=1}^{\mathcal{N}}\sum_{i_l,k_l}^{N_l,M_l}\eta^l\left(c^l_{(i_l,k_l)} - 1\right)$$

The approximative solution obtained by the Tikhonov's regularization (see Poznyak et al. [2000]) is given by

$$x^*,\mu^*,\eta^* = \arg\max_{x\in X}\min_{\mu,\eta\geq 0}\mathcal{L}_\delta(x,\mu,\eta)$$

where

$$\mathcal{L}_\delta(x,\mu,\eta) = \tilde{g}(c^1,...,c^{\mathcal{N}}) - \sum_{l=1}^{\mathcal{N}}\sum_{j_l=1}^{N_l}\mu^l_{(j_l)}h^l_{(j_l)}(c^l) - \sum_{l=1}^{\mathcal{N}}\sum_{i_l,k_l}^{N_l,M_l}\eta^l\left(c^l_{(i_l,k_l)} - 1\right) - \frac{\delta}{2}\|x\|^2 + \frac{\delta}{2}\sum_{l=1}^{\mathcal{N}}\sum_{j_l=1}^{N_l}\left(\mu^l_{(j_l)}\right)^2 + \frac{\delta}{2}\sum_{l=1}^{\mathcal{N}}\left(\eta^l\right)^2 \quad (7)$$

Notice that the Lagrange function (7) satisfies the saddle-point (Poznyak [2009]) condition, namely, for all $x \in X$, and $\mu,\eta \geq 0$ we have

$$\mathcal{L}_\delta(x_\delta,\mu^*_\delta,\eta^*_\delta) \leq \mathcal{L}_\delta(x^*_\delta,\mu^*_\delta,\eta^*_\delta) \leq \mathcal{L}_\delta(x^*_\delta,\mu_\delta,\eta_\delta)$$

2.2 The proximal format

In the proximal format (see, Antipin [2005]) the relation (7) can be expressed as

$$\begin{aligned} \mu^*_\delta &= \arg\min_{\mu\geq 0}\left\{\frac{1}{2}\|\mu-\mu^*_\delta\|^2 + \gamma\mathcal{L}_\delta(x^*_\delta,\mu,\eta^*_\delta)\right\} \\ \eta^*_\delta &= \arg\min_{\eta\geq 0}\left\{\frac{1}{2}\|\eta-\eta^*_\delta\|^2 + \gamma\mathcal{L}_\delta(x^*_\delta,\mu^*_\delta,\eta)\right\} \\ x^*_\delta &= \arg\max_{x\in X}\left\{-\frac{1}{2}\|x-x^*_\delta\|^2 + \gamma\mathcal{L}_\delta(x,\mu^*_\delta,\eta^*_\delta)\right\} \end{aligned} \quad (8)$$

where the solutions x^*_δ, μ^*_δ and η^*_δ depend on the parameters $\delta > 0$ and $\gamma > 0$.

2.3 The Extraproximal method

The Extraproximal Method for (7) was suggested in (Antipin [2005], Trejo et al. [2015]). We design the method for the static Nash bargaining game in a general format with some fixed admissible initial values ($x_0 \in X$ and $\mu_0,\eta_0 \geq 0$ as follows:

1. The *first half-step*:

$$\begin{aligned} \bar{\mu}_n &= \arg\max_{\mu\geq 0}\left\{-\frac{1}{2}\|\mu-\mu_n\|^2 - \gamma\mathcal{L}_\delta(x_n,\mu,\eta_n)\right\} \\ \bar{\eta}_n &= \arg\max_{\eta\geq 0}\left\{-\frac{1}{2}\|\eta-\eta_n\|^2 - \gamma\mathcal{L}_\delta(x_n,\bar{\mu}_n,\eta)\right\} \\ \bar{x}_n &= \arg\max_{x\in X}\left\{-\frac{1}{2}\|x-x_n\|^2 + \gamma\mathcal{L}_\delta(x,\bar{\mu}_n,\bar{\eta}_n)\right\} \end{aligned} \quad (9)$$

2. The *second half-step*:

$$\mu_{n+1} = \arg\max_{\mu\geq 0}\left\{-\frac{1}{2}\|\mu-\mu_n\|^2 - \gamma\mathcal{L}_\delta(\bar{x}_n,\mu,\bar{\eta}_n)\right\}$$

$$\eta_{n+1} = \arg\max_{\eta\geq 0}\left\{-\frac{1}{2}\|\eta-\eta_n\|^2 - \gamma\mathcal{L}_\delta(\bar{x}_n,\bar{\mu}_n,\eta)\right\} \quad (10)$$

$$x_{n+1} = \arg\max_{x\in X}\left\{-\frac{1}{2}\|x-x_n\|^2 + \gamma\mathcal{L}_\delta(x,\bar{\mu}_n,\bar{\eta}_n)\right\}$$

2.4 Convergence and Uniqueness Analysis

The following theorem presents the convergence conditions of (9 - 10) and the estimate of its rate of convergence for the Nash bargaining equilibrium. We prove that the extraproximal method converges to a unique equilibrium point. Let us define the following extended vectors

$$\tilde{x} = x\in\tilde{X}, \quad \tilde{\mu} = \begin{pmatrix}\mu\\ \eta\end{pmatrix}\in\mathbb{R}^+\times\mathbb{R}^+$$

The regularized Lagrange function can be expressed as $\tilde{\mathcal{L}}_\delta(\tilde{x},\tilde{\mu}) := \mathcal{L}_\delta(x,\mu,\eta)$, and the equilibrium point that satisfies (8) can be expressed as

$$\tilde{\mu}^*_\delta = \arg\min_{\tilde{\mu}\geq 0}\left\{\frac{1}{2}\|\tilde{\mu}-\tilde{\mu}^*_\delta\|^2 + \gamma\tilde{\mathcal{L}}_\delta(\tilde{x}^*_\delta,\tilde{\mu})\right\}$$

$$\tilde{x}^*_\delta = \arg\max_{\tilde{x}\in\tilde{X}}\left\{-\frac{1}{2}\|\tilde{x}-\tilde{x}^*_\delta\|^2 + \gamma\tilde{\mathcal{L}}_\delta(\tilde{x},\tilde{\mu}^*_\delta)\right\}$$

Let us introduce the following variables

$$\tilde{w} = \begin{pmatrix}\tilde{w}_1\\ \tilde{w}_2\end{pmatrix}\in X\times\mathbb{R}^+,\ \tilde{v} = \begin{pmatrix}\tilde{v}_1\\ \tilde{v}_2\end{pmatrix}\in X\times\mathbb{R}^+$$

For $\tilde{w}_1=\tilde{x}$, $\tilde{w}_2=\tilde{\mu}$, $\tilde{v}_1=\tilde{v}^*_1=\tilde{x}^*_\delta$ and $\tilde{v}_2=\tilde{v}^*_2=\tilde{\mu}^*_\delta$, let define the Lagrange function in term of $\tilde{w}$ and $\tilde{v}$

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

In these variables the relation (8) can be represented by

$$\tilde{v}^* = \arg\max_{\tilde{w}\in\tilde{X}\times\mathbb{R}^+}\left\{-\tfrac{1}{2}\|\tilde{w}-\tilde{v}^*\|^2 + \gamma L_\delta(\tilde{w},\tilde{v}^*)\right\} \quad (11)$$

The extraproximal method can be expressed by

1. First step

$$\hat{v}_n = \arg\max_{\tilde{w}\in\tilde{X}\times\mathbb{R}^+}\left\{-\tfrac{1}{2}\|\tilde{w}-\tilde{v}_n\|^2 + \gamma L_\delta(\tilde{w},\tilde{v}_n)\right\} \quad (12)$$

2. Second step

$$\tilde{v}_{n+1} = \arg\max_{\tilde{w}\in\tilde{X}\times\mathbb{R}^+}\left\{-\tfrac{1}{2}\|\tilde{w}-\tilde{v}_n\|^2 + \gamma L_\delta(\tilde{w},\hat{v}_n)\right\} \quad (13)$$

Lemma 2. Let $\tilde{\mathcal{L}}_\delta(\tilde{x},\tilde{\mu})$ be differentiable in $\tilde{x}$ and $\tilde{\mu}$, whose partial derivative with respect to $\tilde{\mu}$ satisfies the Lipschitz condition with positive constant C_0. Then,

$$\|\tilde{v}_{n+1}-\hat{v}_n\| \leq \gamma C_0\|\tilde{v}_n-\hat{v}_n\|$$

Proof. See Trejo et al. [2017].

Lemma 3. Let us consider the set of regularized solutions of a non-empty game. The behavior of the regularized function is described by the following inequality:

$$L_\delta(\tilde{w},\tilde{w}) - L_\delta(\tilde{v}^*_\delta,\tilde{w}) \geq \delta\|\tilde{w}-\tilde{v}^*_\delta\|$$

for all $\tilde{w}\in\{\tilde{w}\mid\tilde{w}\in X\times\mathbb{R}^+\}$ and $\delta>0$.

Proof. See Trejo et al. [2017].

Theorem 4. Let $\tilde{\mathcal{L}}_\delta(\tilde{x},\tilde{\mu})$ be differentiable in $\tilde{x}$ and $\tilde{\mu}$, whose partial derivative with respect to $\tilde{\mu}$ satisfies the Lipschitz condition with positive constant C. Then, for any $\delta>0$ there exists a small-enough

$$\gamma_0 = \gamma_0(\delta) < C := \min\left\{\tfrac{1}{\sqrt{2}C_0}, \tfrac{1+\sqrt{1+2(C_0)^2}}{2(C_0)^2}\right\}$$

where such that, for any $0<\gamma\leq\gamma_0$, sequence $\{\tilde{v}_n\}$, which generated by the equivalent extraproximal procedure (12 - 13), monotonically converges with exponential rate $q\in(0,1)$ to a unique equilibrium point $\tilde{v}^*$, i.e.,

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

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. See the complete proof in Trejo et al. [2017].

3. NUMERICAL EXAMPLE

Our goal is to analyze a 2-player Nash Bargaining situation in a class of ergodic controllable finite Markov chains. Let us denote the disagreement cost that depends on the strategies $c^l_{(i_l,k_l)}$ $(l=1,2)$ for players 1 and 2 as $\psi^*_1(c^1,c^2)$ and $\psi^*_2(c^1,c^2)$ respectively, and the solution for the Nash bargaining problem as the point (ψ_1,ψ_2). Let the states $N_1=N_2=3$, and the number of actions $M_1=M_2=2$. The individual utility for each player are defined by

$$J^1_{(i,j|1)} = \begin{bmatrix}7 & 17 & 13\\ 0 & 1 & 18\\ 13 & 7 & 10\end{bmatrix}\ J^1_{(i,j|2)} = \begin{bmatrix}18 & 3 & 10\\ 9 & 0 & 7\\ 15 & 6 & 16\end{bmatrix}$$

$$J^2_{(i,j|1)} = \begin{bmatrix}9 & 11 & 6\\ 9 & 17 & 3\\ 11 & 1 & 4\end{bmatrix}\ J^2_{(i,j|2)} = \begin{bmatrix}10 & 18 & 0\\ 12 & 7 & 18\\ 17 & 6 & 10\end{bmatrix}$$

The transition matrices for each player are as follows

$$\pi^1_{(i,j|1)} = \begin{bmatrix}0.5144 & 0.2877 & 0.1978\\ 0.3775 & 0.0893 & 0.5332\\ 0.3305 & 0.2703 & 0.3992\end{bmatrix}\ \pi^1_{(i,j|2)} = \begin{bmatrix}0.3438 & 0.3846 & 0.2717\\ 0.2484 & 0.0756 & 0.6759\\ 0.1378 & 0.4655 & 0.3968\end{bmatrix}$$

$$\pi^2_{(i,j|1)} = \begin{bmatrix}0.3541 & 0.1945 & 0.4514\\ 0.5929 & 0.2559 & 0.1512\\ 0.4288 & 0.2434 & 0.3278\end{bmatrix}\ \pi^2_{(i,j|2)} = \begin{bmatrix}0.6435 & 0.0216 & 0.3349\\ 0.2990 & 0.3905 & 0.3105\\ 0.5575 & 0.2203 & 0.2221\end{bmatrix}$$

3.1 The disagreement point - Nash equilibrium

Let us introduce the variables, see Trejo et al. [2015]

$$x := col\ c^l,\ \hat{x} := col\ c^{\hat{l}},\ \left(l=\overline{1,\mathcal{N}}\right)$$

The strategies of the players are denoted by the vector x, and $\hat{x}$ is a strategy of the rest of the players adjoint to x. Players try to find a join strategy $x^*=(c^1,...,c^{\mathcal{N}})$ satisfying

$$f(x,\hat{x}) := \sum_{l=1}^{\mathcal{N}}\left[\psi^l\left(c^l,c^{\hat{l}}\right) - \left(\max_{c^l\in C^l}\psi^l\left(c^l,c^{\hat{l}}\right)\right)\right] \quad (14)$$

Here $\psi^l\left(c^l,c^{\hat{l}}\right)$ is the utility-function of the player l which plays the strategy c^l and the other player plays the strategy $c^{\hat{l}}$. If we consider the utopia point

$$\bar{c}^l := \arg\max_{c^l\in C^l}\psi^l\left(c^l,c^{\hat{l}}\right)$$

then, we can rewrite (14) as follows

$$f(x,\hat{x}) := \sum_{l=1}^{\mathcal{N}} \left[\psi^l \left(c^l, c^{\hat{l}} \right) - \psi^l \left(\bar{c}^l, c^{\hat{l}} \right) \right] \quad (15)$$

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

The function $f(x,\hat{x})$ satisfies the Nash condition

$$\psi^l \left(c^l, c^{\hat{l}} \right) - \psi^l \left(\bar{c}^l, c^{\hat{l}} \right) \leq 0 \quad (16)$$

for any $c^l \in C^l$ and all $l = \overline{1,\mathcal{N}}$.

A strategy x^* is said to be a Nash equilibrium if

$$x^* \in \ arg \max_{x \in X_{adm}} \ \{f(x,\hat{x})\}$$

Applying the regularized Lagrange principle we have the solution for the Nash equilibrium

$$x^*, \hat{x}^*, \mu^*, \eta^* = \arg \max_{x \in X, \hat{x} \in \hat{X}} \ \min_{\mu, \eta \geq 0} \ \mathcal{L}_{\theta,\delta}(x, \hat{x}, \mu, \eta)$$

$$\mathcal{L}_{\theta,\delta}(x,\hat{x},\mu,\eta) := (1-\theta) f(x,\hat{x}) - \sum_{l=1}^{\mathcal{N}} \sum_{j_l=1}^{N_l} \mu^l_{(j_l)} h^l_{(j_l)}(c^l) - \sum_{l=1}^{\mathcal{N}} \sum_{i_l,k_l}^{N_l,M_l} \eta^l \left(c^l_{(i_l,k_l)} - 1 \right) - \frac{\delta}{2}(\|x\|^2 + \|\hat{x}\|^2) + \frac{\delta}{2} \sum_{l=1}^{\mathcal{N}} \sum_{j_l=1}^{N_l} \left(\mu^l_{(j_l)} \right)^2 + \frac{\delta}{2} \sum_{l=1}^{\mathcal{N}} (\eta^l)^2 \quad (17)$$

Notice also that the Lagrange function (17) satisfies the saddle-point condition

$$\mathcal{L}_{\theta,\delta}(x_\delta, \hat{x}_\delta, \mu^*_\delta, \eta^*_\delta) \leq \mathcal{L}_{\theta,\delta}(x^*_\delta, \hat{x}^*_\delta, \mu^*_\delta, \eta^*_\delta) \leq \mathcal{L}_{\theta,\delta}(x^*_\delta, \hat{x}^*_\delta, \mu_\delta, \eta_\delta)$$

The proximal format. The relation (17) can be expressed in the proximal format as

$$\mu^*_\delta = \arg\min_{\mu \geq 0} \left\{ \frac{1}{2} \|\mu - \mu^*_\delta\|^2 + \gamma \mathcal{L}_{\theta,\delta}(x^*_\delta, \hat{x}^*_\delta, \mu, \eta^*_\delta) \right\}$$

$$\eta^*_\delta = \arg\min_{\eta \geq 0} \left\{ \frac{1}{2} \|\eta - \eta^*_\delta\|^2 + \gamma \mathcal{L}_{\theta,\delta}(x^*_\delta, \hat{x}^*_\delta, \mu^*_\delta, \eta) \right\}$$

$$x^*_\delta = \arg\max_{x \in X} \left\{ -\frac{1}{2} \|x - x^*_\delta\|^2 + \gamma \mathcal{L}_{\theta,\delta}(x, \hat{x}^*_\delta, \mu^*_\delta, \eta^*_\delta) \right\}$$

$$\hat{x}^*_\delta = \arg\max_{\hat{x} \in \hat{X}} \left\{ -\frac{1}{2} \|\hat{x} - \hat{x}^*_\delta\|^2 + \gamma \mathcal{L}_{\theta,\delta}(x^*_\delta, \hat{x}, \mu^*_\delta, \eta^*_\delta) \right\}$$

where the solutions x^*_δ, $\hat{x}^*_\delta(u)$, μ^*_δ and η^*_δ depend on the parameters $\delta, \gamma > 0$.

The Extraproximal method. We design the method for the static Nash game in a general format with some fixed admissible initial values ($x_0 \in X$, $\hat{x}_0 \in \hat{X}$, and $\mu_0, \eta_0 \geq 0$, as follows:

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

$$\bar{\mu}_n = \arg\max_{\mu \geq 0} \left\{ -\frac{1}{2} \|\mu - \mu_n\|^2 - \gamma \mathcal{L}_{\theta,\delta}(x_n, \hat{x}_n, \mu, \eta_n) \right\}$$

$$\bar{\eta}_n = \arg\max_{\eta \geq 0} \left\{ -\frac{1}{2} \|\eta - \eta_n\|^2 - \gamma \mathcal{L}_{\theta,\delta}(x_n, \hat{x}_n, \mu_n, \eta) \right\}$$

$$\bar{x}_n = \arg\max_{x \in X} \left\{ -\frac{1}{2} \|x - x_n\|^2 + \gamma \mathcal{L}_{\theta,\delta}(x, \hat{x}_n, \bar{\mu}_n, \bar{\eta}_n) \right\}$$

$$\bar{\hat{x}}_n = \arg\max_{\hat{x} \in \hat{X}} \left\{ -\frac{1}{2} \|\hat{x} - \hat{x}_n\|^2 + \gamma \mathcal{L}_{\theta,\delta}(x_n, \hat{x}, \bar{\mu}_n, \bar{\eta}_n) \right\}$$

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

$$\mu_{n+1} = \arg\max_{\mu \geq 0} \left\{ -\frac{1}{2} \|\mu - \mu_n\|^2 - \gamma \mathcal{L}_{\theta,\delta}(\bar{x}_n, \bar{\hat{x}}_n, \mu, \bar{\eta}_n) \right\}$$

$$\eta_{n+1} = \arg\max_{\eta \geq 0} \left\{ -\frac{1}{2} \|\eta - \eta_n\|^2 - \gamma \mathcal{L}_{\theta,\delta}(\bar{x}_n, \bar{\hat{x}}_n, \bar{\mu}_n, \eta) \right\}$$

$$x_{n+1} = \arg\max_{x \in X} \left\{ -\frac{1}{2} \|x - x_n\|^2 + \gamma \mathcal{L}_{\theta,\delta}(x, \bar{\hat{x}}_n, \bar{\mu}_n, \bar{\eta}_n) \right\}$$

$$\hat{x}_{n+1} = \arg\max_{\hat{x} \in \hat{X}} \left\{ -\frac{1}{2} \|\hat{x} - \hat{x}_n\|^2 + \gamma \mathcal{L}_{\theta,\delta}(\bar{x}_n, \hat{x}, \bar{\mu}_n, \bar{\eta}_n) \right\}$$

Computing the disagreement point. Given δ, γ and applying the extraproximal method we obtain the convergence of the strategies for each player in the disagreement point in terms of the variable $c^l_{(i_l,k_l)}$ (Fig. 1 and Fig. 2).

$$c^1 = \begin{bmatrix} 0.1683 & 0.1551 \\ 0.1829 & 0.0973 \\ 0.1853 & 0.2111 \end{bmatrix} \quad c^2 = \begin{bmatrix} 0.2618 & 0.2122 \\ 0.0673 & 0.1320 \\ 0.1305 & 0.1962 \end{bmatrix}$$

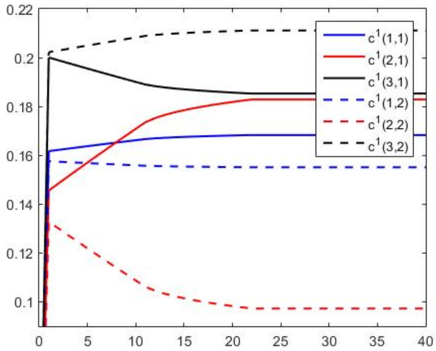

Fig. 1. Convergence of the strategies for player 1

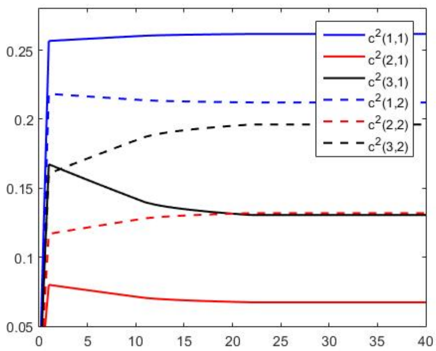

Fig. 2. Convergence of the strategies for player 2.

With the strategies calculated, the resulting utilities in the disagreement point for each player are as follows:

$$\psi^*_1(c^1, c^2) = 120.3001 \quad \psi^*_2(c^1, c^2) = 97.0832 \quad (18)$$

3.2 Bargaining solution

The Nash's solution has a simple geometric interpretation in a two-person game: given a bargaining pair, for every point (ψ_1, ψ_2), consider the product (area of a rectangle) $(\psi_1 - \psi^*_1)(\psi_2 - \psi^*_2)$. Then (ψ_1, ψ_2) is the unique point in the Pareto front that maximizes this product (Muthoo [2002]).

Computing the Nash Bargaining solution Following the method in Section 2 and applying the extraproximal method for the Nash bargaining problem (9 - 10), we obtain the convergence of the strategies for the bargaining solution in terms of the variable $c^l_{(i_l,k_l)}$ for each player (see Fig. 3 and Fig. 4).

$$c^1 = \begin{bmatrix} 0.1890 & 0.1178 \\ 0.3057 & 0.0010 \\ 0.0010 & 0.3854 \end{bmatrix} \quad c^2 = \begin{bmatrix} 0.3463 & 0.0881 \\ 0.0010 & 0.2325 \\ 0.0010 & 0.3310 \end{bmatrix}$$

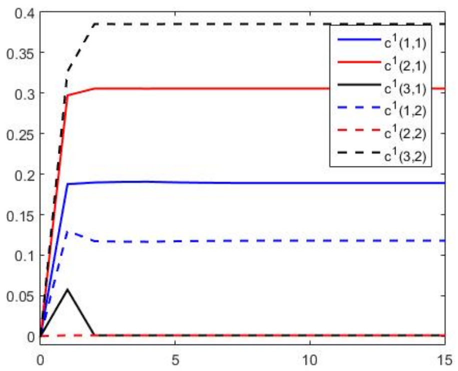

Fig. 3. Convergence of the strategies of player 1

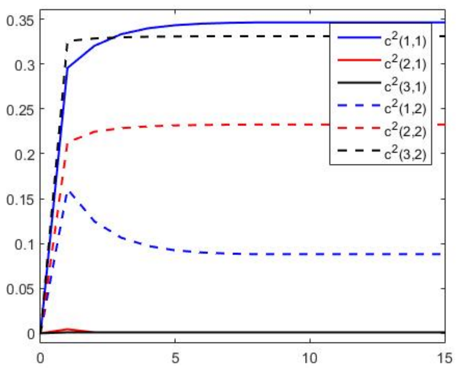

Fig. 4. Convergence of the strategies of player 2.

With the strategies calculated, the resulting utilities in the bargaining solution are as follows:

$$\psi_1(c^1, c^2) = 139.6854 \quad \psi_2(c^1, c^2) = 119.4296 \qquad (19)$$

We can see that the profits obtained at the point of Nash bargaining solution are greater than those obtained at the disagreement point.

4. CONCLUSION

This paper developed a method to find the solution of the Nash bargaining process. The solution of the problem is restricted to a class of controlled, ergodic and finite Markov chains games. We considered the characterization of the disagreement point as the Nash equilibrium. Both, the disagreement point and the Nash bargaining solution were represented in terms of nonlinear programming equations implementing the regularized Lagrange method based-on the Tikhonov's regularization approach for ensuring convergence to a unique equilibrium point. For solving the problem we employed the extraproximal method, a two-step iterated procedure for computing the equilibrium point. We presented the convergence and rate of convergence of the method. A numerical example validates the proposed method.

REFERENCES

Antipin, A.S. (2005). An extraproximal method for solving equilibrium programming problems and games. *Computational Mathematics and Mathematical Physics*, 45(11), 1893–1914.

Clempner, J.B. and Poznyak, A.S. (2014). Simple computing of the customer lifetime value: A fixed local-optimal policy approach. *Journal of Systems Science and Systems Engineering*, 23(4), 439–459.

Coles, M.G. and Muthoo, A. (2003). Bargaining in a non-stationary environment. *Journal of Economic Theory*, 109(1), 70–89.

Cripps, M.W. (1998). Markov bargaining games. *Journal of Economic Dynamics and Control*, 22(3), 341–355.

Dagan, N., Volij, O., and Winter, E. (2002). A characterization of the nash bargaining solution. *Social Choice and Welfare*, 19(4), 811–823.

Forgó, F., Szép, J., and F., S. (1999). *Introduction to the Theory of Games: Concepts, Methods, Applications.* Kluwer Academic Publishers.

Kalai, E. and Smorodinsky, M. (1975). Other solutions to nashs bargaining problem. *Econometrica*, 43(3), 513–518.

Kalandrakis, A. (2004). A three-player dynamic majoritarian bargaining game. *Journal of Economic Theory*, 116(2), 294–322.

Kennan, J. (2001). Repeated bargaining with persistent private information. *Review of Economic Studies*, 68, 719–755.

Muthoo, A. (2002). *Bargaining Theory with Applications.* Cambridge University Press.

Nash, J.F. (1950). The bargaining problem. *Econometrica*, 18(2), 155–162.

Nash, J.F. (1951). Non-cooperative games. *Annals of Mathematics*, 54, 286–295.

Perles, M.A. and Maschler, M. (1981). The super-additive solution for the nash bargaining game. *International Journal of Game Theory*, 10(3), 163–193.

Poznyak, A.S. (2009). *Advance Mathematical Tools for Automatic Control Engineers. Vol 2 Deterministic Techniques.* Elsevier, Amsterdam.

Poznyak, A.S., Najim, K., and Gomez-Ramirez., E. (2000). *Self-learning control of finite Markov chains.* Marcel Dekker, Inc., New York,.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, 50(1), 97–109.

Trejo, K.K., Clempner, J.B., and Poznyak, A.S. (2015). Computing the stackelberg/nash equilibria using the extraproximal method: Convergence analysis and implementation details for markov chains games. *International Journal of Applied Mathematics and Computer Science*, 25(2), 337–351.

Trejo, K.K., Clempner, J.B., and Poznyak, A.S. (2017). Computing the strong l_p-nash equilibrium for markov chains games: Convergence and uniqueness. *Applied Mathematical Modelling*, 41, 399–418.