U N I V E R S I T Ä T   H A M B U R G

Stochastic games
for distributed players on graphs

R. K. Chornei, H. Daduna, P. S. Knopov

Preprint–No. 2002-06   Dezember 2002

FACHBEREICH MATHEMATIK

SCHWERPUNKT MATHEMATISCHE STATISTIK
UND STOCHASTISCHE PROZESSE

BUNDESSTR. 55, D – 20146 HAMBURG

Stochastic games for distributed players on graphs

Ruslan K. Chornei Hans Daduna Pavel S. Knopov *

December 2, 2002

Summary: We introduce the concept of local acting distributed players into sequential stochastic games with general compact state and action spaces. The state transition function for the system is of local structure as well and this results in Markov properties in space and time for the describing processes. We prove that we can reduce optimality problems for local strategies to only considering Markov strategies. We further prove the existence of optimal strategies and of a value for the game with respect to the asymptotic average reward criterion.

MSC (2000) subject classification: 91A15, 60K35, 90C40

Key Words: Sequential games, multiperson games, coalitions, random fields, local strategies.

1 Introduction

Stochastic games will be understood in the following as multiperson stochastic sequential games with an infinite horizon. The description of the behavior of such systems is therefore by construction of suitable stochastic processes with discrete time scale.

There is a wide area of applications for the games we have in mind. Indeed many phenomena of management science and economics, biology and psychology, and especially military affairs are modeled by using stochastic sequential games with very different framework. A short introduction with many specific examples can be found in [15][Chapter 9].

Only recently emerged anew interest in many–player stochastic games as an essential ingredience of the control procedures for large networks of interconnected stations, especially network nodes of computer and communications networks which share common resources. The most prominent example being by now the Internet. Classical control of queueing systems is modeled and performed either by using stochastic dynamic optimization or by using 2–person games, where one of the players represents the nature which perturbes the functioning of the system; an introduction into the field is [10] summarizing those techniques as *Competetitive Markov Decision Processes*. For applications of game theoretical methods to queueing systems and small networks see [1], [2], [3], and the references therein.

However, in case of the large networks we have in mind, typically a large number of players (nodes of the networks) interact and compete for the networks resources. These players can act cooperatively or noncooperatively. In the first case this leads to building coalitions, which usually have partly antagonistic criteria for the performance measures of the network. The

*Part of this research was done when the third author held a grant of the DFG at Hamburg University.

above mentioned introduction [15][Chapter 9] is centered around noncooperative multiperson games. General schemes for cooperation under incomplete information with bargaining are described from an abstract point of view in [23]. A monography on classical topics of stochastic games is [22]. A more recent survey including problems of multiperson games is [24]. A classical paper on the general subject with strong connections to stochastic dynamic optimization is [9].

Control of stochastic networks (networks of queues) is considered explicitly as a multiperson stochastic game in [31].

The standard models of multiperson systems which compete and/or cooperate for common resources allow the players to make their decision on the basis of either the whole history of the system, or as can be proved to be almost similarly effective on the basis of the knowledge of the complete actual state of the system. However, for large networks this seems to be an optimistic, resp. unrealistic assumption because any player in real large networks will be able to take into consideration for his decision making only what is known about the state of his (suitably defined) neighborhood. It follows that in a natural way there emerges a neighborhood structure in the set of decision makers (players), resp. in the set of nodes, where the players reside. Such neighborhoods are usually determined by a neighborhood graph, which in our case will be undirected, but this restriction can be removed easily.
Models from economics where neighborhood systems emerge and are developed as structural property to determine equilibrium behavior of the system and to construct efficient control mechanisms for optimizing the system are described in [7] and[17].
Neighborhood systems are well known structures in stochastic process theory and their applications: They determine which coordinates interact in the interacting processes. The most prominent source to the field is [19], where the systems under consideration are interacting particle systems, which traditionally are thought to come from the area of statistical physics [26], [28]. A more recent study in this field is the follower [20] of [19]. Most of the models from statistical physics have a regular neighborhood structure, usually some sublattice of $\mathbb{Z}^n$. Contrasting to that, queueing networks for computer and communications system modelling are quite irregular, which enforces us to consider general graphs for describing neighborhoods. But nevertheless this immediately leads to the definition of random fields in parallel to the classical setting, and then to Markov random fields, see e.g. [16][Chapter 9] or [25].

Random fields describe the spatial distribution of a complex system at a fixed time instant, or as it is assumed in statistical physics, they describe the equilibrium state of a system. If the spatial dependencies are restricted to nearest neighborhoods, we arrive at the notion of a Markov random field. To consider evolution over time of Markovian random fields and to construct stochastic processes (usually Markov processes) dates back to the work of Spitzer, see [19] for references. We are faced here with similar problems, when trying to control the space–time behavior of interacting node systems.

With respect to control of large networks it is pointed out in [1] that there is a strong need and “ growing interest in situations of several controllers and several objectives, which gives rise to non–cooperative models“. Introducing coalitions into the game structure here can therefore be seen as continuing the study of that paper where the interplay of, say, an admission controller to a queue and a service controller of that queue is considered. Here we would interprete one coalition as the set of admission controllers to a data transmission and

communication network and the other to be the set of service controllers in that network. While usually the first coalition is interested in small transmission delay, the second coalition traditionally has as main objective high throughputs.
Note that for such control problems it is again natural to assume that the controllers use only information available in their neighborhood. In highspeed networks it is almost necessary to restrict oneself to use only policies depending on neighborhood states, because too much information incorporated would diminish the performance of the control regime. This is applied e.g. in the definition of the *Dynamic Alternative Routing* algorithm, see [12], and further developed in the suggested *Balanced Dynamic Alternative Routing* algorithm, see [4].

Starting from these observations and similarly from other network structures with locality properties (see the introduction in [5] and [6] for further examples) where interacting players can be thought to occur, our aim in this paper is to develop a general model for describing the interaction of players and coalitions which live on graph structured networks and have to optimize their behavior in sequential stochastic games. This leads immediately to the main questions of the existence of optimal strategies and a value of the games, and to the question whether these optimal strategies can be chosen from easy to handle subclasses of strategies. The main point is that we shall consider only local strategies as being suitable for the situations described above. Within this class we shall prove the required existence theorems and structural properties, allowing for the technical framework rather general assumptions.

The paper is organized as follows. We start in Section 2 with descriptions of neighborhood graphs and the general state spaces of the systems under consideration. The stochastic development of uncontrolled Markov processes with transition probabilities which are synchronous and local is used as a starting point for construction of local policies (strategies) for distributed players: We define admissable local strategies, with sometimes additional properties, being e.g., stationary, Markovian, and deterministic. The transition mechanism of our systems are homogeneous in time, locally determined, and with general Markovian properties in time and space.
We assume that the set of players is split into two disjoint coalitions and that the rewards are paid between the coalitions by the players.
In Section 3 we show that (given some technical requirements) with respect to the asymptotic average reward criterion we can find in the class of local strategies even Markov strategies which yield optimal rewards. The required technical conditions are then shown to hold under rather mild conditions (see Assumption I and II, Section 3.2). It then follows that the game has a value under this condition and both coalitions have optimal local stationary Markovian strategies.

Summarizing this, the contribution of this paper is to introduce the concept of locality into general stochastic games. This can be quoted also as another concept of incomplete information, different from that usually dealt with, see e.g. [23], where each player has complete knowledge about his own *type* (which summarizes his properties of relevance to the development of the system), and has only a (probabilistic) guess about the *type* of the other players.

2 Distributed players which act locally

2.1 Preliminaries

For systems with locally interacting coordinates the interaction structure is defined via an undirected finite neighborhood graph $\Gamma = (V, B)$ with set of vertices (nodes) V and set of edges B. Denote by $\{k, j\}$ the edge of the graph connecting vertices k and j. The neighborhood of vertex k is the node set $N(k) = \{j : \{k, j\}) \in B\}$. The complete neighborhood of vertex k is $\widetilde{N}(k) = N(k) \cup \{k\}$, i. e. the neighborhood of the vertex k, and including k. For any $K \subset V$ we define the neighborhood $N(K) = \bigcup_{k \in K} N(k) - K$, and the complete neighborhood $\widetilde{N}(K) = N(K) \cup K$. For every node $i \in V$ let X_i be some compact metric space with countable basis, the local state space at that node. Then $X := \underset{i \in V}{\times} X_i$ is the global state space of the system. For every subset of nodes $K \subset V$ we denote the marginal vector of state x by $x_K = (x_k,\ k \in K) \in X_K = \underset{i \in K}{\times} X_i$.

Let $\mathfrak{X}_i$ be the sigma-field of Borel subsets of X_i and $\mathfrak{X}_K = \sigma\{\underset{i \in K}{\times} \mathfrak{X}_i\}$ the sigma-field generated by $\underset{i \in K}{\times} \mathfrak{X}_i$, and $K \subseteq V$, $\mathfrak{X}_V = \mathfrak{X}$.

Throughout the paper we fix the underlying probability space $(\Omega, \mathcal{F}, P)$ where all random variables are defined on.

Definition 2.1. (1) A random variable ξ defined on $(\Omega, \mathcal{F}, P)$ with values in X is called a random field over (V, B) (or simply a random field over V). The marginal random variables with values in the space X_K are denoted by ξ_K. For $K = \{k\}$ we write ξ_k.

(2) A random field ξ on (V, B) is a Markov field if for all $k \in V$ the following holds:

$$P(\xi_k \in C_k / \xi_{V-\{k\}} = x_{V-\{k\}}) = P(\xi_k \in C_k / \xi_{N(k)} = x_{N(k)}), \qquad \forall\, x \in X, \quad C_k \in \mathfrak{X}_k, \quad (1)$$

where $P(\xi_k \in C_k / \xi_{N(k)} = x_{N(k)})$, resp. $P(\xi_k \in C_k / \xi_{V-\{k\}} = x_{V-\{k\}})$ are the regular conditional probability on $(X_k, \mathfrak{X}_k)$ given $\xi_{N(k)} = x_{N(k)}$, resp. $\xi_{V-\{k\}} = x_{V-\{k\}}$. ⊙

Remark 2.1. Equation (1) is to be read as follows:

$$P(\xi_k \in C_k / \xi_{V-\{k\}} = x_{V-\{k\}}) = P(\xi_k \in C_k / \xi_{V-\{k\}} = y_{V-\{k\}}) \quad \text{if } x_{N(k)} = y_{N(k)} \text{ holds.} \quad (2)$$

Similar readings will apply throughout the paper. ⊙

In many problems related to economical or engineering applications random fields describe the state of a system for some fixed time instant. The evolution over time of that system is then described by a stochastic process with random fields as one-dimensional marginals in time. In this case we write $\xi = (\xi^t)$ to emphasize the time-dependence. We assume that t takes discrete values $t = 0, 1, \ldots$. The subscript k in ξ_k^t refers to the vertex k, ξ_k^t therefore denotes the marginal distribution for time t and node k of some vector valued process $\xi = (\xi^t : t = 0, 1, \ldots)$.

If a Markov process ξ from its very construction is related to the neighborhood system $\{N(k) : k \in V\}$ of (V, B), it is natural to assume that with respect to the evolution over time the probability for $\xi_k^t = x_k$ at the k-th vertex depends on the previous states of the whole system only through the values of the vertices in $\widetilde{N}(k)$ (including k) at $t - 1$. To describe this we introduce the notion of Markov property in space and in time along the lines of Definition 2.1.

Definition 2.2. Let $\xi = \{\xi^t, t = 0, 1, \ldots\}$ be a discrete time Markov process with state space X. The transition probabilities of ξ are said to be *local* [30, p. 100] if for all $k \in V$, $x^0, \ldots, x^{t+1} \in X$, $C_k \in \mathfrak{X}_k$

$$P\{\xi_k^{t+1} \in C_k / \xi^t = x^t, \ldots, \xi^0 = x^0\} = P\{\xi_k^{t+1} \in C_k / \xi^t_{\widetilde{N}(k)} = x^t_{\widetilde{N}(k)}\} \tag{3}$$

holds, i. e. the transition probability at vertex k depends on the state of its complete neighborhood at the previous time instant only.

The transition probabilities of ξ are said to be *synchronous* [30, p. 100] if

$$P\{\xi_K^{t+1} \in C_K / \xi^t = x^t\} = \prod_{k \in K} P\{\xi_k^{t+1} \in C_k / \xi^t = x^t\}, \qquad \forall\, K \subset V, \quad x^t \in X, \tag{4}$$

where $C_K = \underset{k \in K}{\times} C_k \in \mathfrak{X}_K$.

If ξ fulfills (3) and (4) it is called a *Markov process with locally interacting synchronous components over* (Γ, X), shortly, a *(time dependent) Markov random field.* $\odot$

Before we shall present the details of our stochastic games we recall some Definitions and Theorems from [18].

Definition 2.3. For topological spaces X and A with Borel σ–algebras $\mathfrak{X}$ and $\mathfrak{A}$ denote by $(2)^A_{set}$ the set of all subsets of A, and let $F\colon X \to (2)^A_{set}$ be a set valued function which associates with $x \in X$ a nonempty closed set $F(x) =: A(x) \subseteq A$ such that the set $\Delta = \{(x,a)\colon x \in X,\ a \in A(x)\} \subset X \times A$ is Borel measurable. (The function F will later determine the admissible decision (control) set $A(x)$ for all states x.)

The map $F\colon X \to (2)^A_{set}$ is called open, closed or Borel measurable if $\{x\colon F(x) \cap E \neq \varnothing\} \in \mathfrak{X}$, for all open, closed or Borel subsets E of A, respectively.

The function $f\colon X \to A$ is called a selector for the map F if $f(x) \in F(x)$ for all $x \in X$. $\odot$

Theorem of Choice. [18, p. 74]. In the setting of Definition 2.3 any Borel measurable map has a Borel measurable selector. $\odot$

Theorem of Choice for Semicontinuous Maps. [18, p. 71]. If A is a compact metric space then any semicontinuous map $F\colon X \to (2)^A_{set}$ has a selector belonging to Baire class 1. $\odot$

2.2 Distributed players and local policies for coalitions

We now describe the $|V|$ - person stochastic games which we shall investigate in this paper. The players are located at the vertices of the interaction graph Γ, the edges represent the connections between the players. Our aim is to model the behavior of two fixed coalitions of the players in a way that the stochastic process model for the interaction of players resembles Markovian neighborhood systems, and that the development of the game over time is a Markov process with a transition mechanism similar to that described in Definition 2.2.

This will result in the notion of a controlled Markov process with locally interacting synchronous components, shortly: a controlled time dependent Markov random field, as it was developed in [5] for the case of finite state spaces and in [6] for the case of compct state spaces. For easier understanding of the modeling procedure we can think that some time dependent Markov random field $\xi = \{\xi^t \colon t \in \mathbb{N}\}$ according to Definition 2.2 is given as our starting point which describes the behavior of the $|V|$ persons which interact locally and synchronous but do not have the possibility of making decisions. ξ will be equipped in the sequel with a control structure which governs the transition kernel of ξ. As usual in multi person games the control and decision mechanism is synchronized and will be localized according to the definitions below, which resemble (3) and therefore will result in a Markov random field behavior in space for a fixed time instant. Over time we construct as usual a Markovian interaction process.

Definition 2.4. (1) The set of actions (control values) usable at control instants is $A = \underset{i\in V}{\times} A_i$ over Γ, where A_i is a set of possible actions (decisions) for vertex i. We assume that A_i is a compact metric space with countable basis and with Borel σ - algebra $\mathcal{A}_i$. $\mathcal{A}$ is the Borel σ -algebra over A.

(2) If for the decision maker at node i at time t with state $\xi^t = x$ the set of control actions is restricted to $A_i^t(x) \subset A_i$, we call $A_i^t(x)$ the set of admissible actions (decisions) at time t in state x.

(3) We assume that the so defined set valued maps

$$F_i^t \colon X \to (2)^{A_i}_{set}, \quad x \to F_i^t(x) = A_i^t(x), \qquad i \in V,$$

depend on $x \in X$ only through $x_{\widetilde{N}(i)} \in X_{\widetilde{N}(i)}$ and are Borel measurable in the sense of Definition 2.3. Thus the F_i^t determine admissibility as a local property. We therefore write $A_i^t(x) =: A_i^t(x_{\widetilde{N}(i)})$. We also assume the sets

$$\Delta_i^t = \{(x_{\widetilde{N}(i)}, a_i) \colon x_{\widetilde{N}(i)} \in X_{\widetilde{N}(i)}, a_i \in A_i^t(x_{\widetilde{N}(i)})\}$$

to be Borel measurable sets of the product space $X_{\widetilde{N}(i)} \times A_i$ and $\Delta^t = \underset{i\in V}{\times} \Delta_i^t$ to be Borel measurable in the product space $\widetilde{X} \times A$, $\widetilde{X} = \underset{i\in V}{\times} X_{\widetilde{N}(i)}$ and $\widetilde{\mathfrak{X}} = \sigma\{\underset{i\in V}{\times} X_{\widetilde{N}(i)}\}$. $\odot$

Our players are located at the edges of the finite graph $\Gamma = (V, B)$, so the coalitions which the players build can be described by a decomposition of the node set $V = V_1 \cup V_2$ $V_1, V_2 \neq \varnothing$, $V_1 \cap V_2 = \varnothing$. V_1 is the set of vertices where players of coalition K_1, V_2 is a set of vertices where players of coalition K_2.
For a shorthand notation we shall often identify an vertex and the player who resides there.

Definition 2.5. (1) Let α_i^t denote the action chosen by the player (decision maker) at node i at time t, $\alpha^t := (\alpha_i^t : i \in V)$ the joint decision vector of all players at time $t \in N$.

(2) A strategy (or policy) π for coalition K_1 with interacting components is defined as vector of coordinate policies $\pi = (\pi_i,\ i \in V_1)$, where for node i $\pi_i = \{\pi_i^0, \dots, \pi_i^t, \dots\}$ is a sequence of transition probabilities $\pi_i^t = \pi_i^t(\cdot/x^0, a^0, \dots, x^{t-1}, a^{t-1}, x^t)$.
Similarly a strategy (or policy) γ for coalition K_2 is $\gamma = (\gamma_j,\ j \in V_2)$ where for node j $\gamma_j = \{\gamma_j^0, \dots, \gamma_j^t, \dots\}$ is a sequence of transition probabilities $\gamma_j^t = \gamma_j^t(\cdot/x^0, a^0, \dots, x^{t-1}, a^{t-1}, x^t)$.

So for any history $h^t = (x^0, a^0, \ldots, x^{t-1}, a^{t-1}, x^t)$ of the system up to time t π_i^t is a probability measure on $(A_i, \mathfrak{A}_i)$ and γ_j^t is a probability measure on $(A_j, \mathfrak{A}_j)$ which are measurably dependent on the history $h^t = (x^0, a^0, \ldots, x^{t-1}, a^{t-1}, x^t)$ up to time t. We therefore have

$$\begin{aligned}
&P(\alpha_i^t \in B_i / \xi^0 = x^0, \alpha^0 = a^0, \ldots, \xi^{t-1} = x^{t-1}, \alpha^{t-1} = a^{t-1}, \xi^t = x^t) = \\
&\qquad = \pi_i^t(B_i / x^0, a^0, \ldots, x^{t-1}, a^{t-1}, x^t) \ \forall \ B_i \in \mathfrak{A}_i, \ i \in V_1, \\
&P(\alpha_j^t \in B_j / \xi^0 = x^0, \alpha^0 = a^0, \ldots, \xi^{t-1} = x^{t-1}, \alpha^{t-1} = a^{t-1}, \xi^t = x^t) = \\
&\qquad = \gamma_j^t(B_j / x^0, a^0, \ldots, x^{t-1}, a^{t-1}, x^t)) \ \forall \ B_j \in \mathfrak{A}_j, \ j \in V_2.
\end{aligned}$$

$\odot$

We assume that the decisions of the players and the coalitions are synchronized according to the following scheme, which is, following the saying in (4), a synchronous control kernel. Note however, that this form of sychronization is a standard assumption for decision making in discrete time sequential games. Even games with perfect information may be considered as allowing at each decision instant for one of the customers only one "decision", see [9][Section 5], which formally results in the general synchronization.

$$\begin{aligned}
&P(\alpha^t \in \mathop{\times}_{i \in V} B_i / \xi^0 = x^0, \alpha^0 = a^0, \ldots, \xi^{t-1} = x^{t-1}, \alpha^{t-1} = a^{t-1}, \xi^t = x^t) = \\
&\qquad = \prod_{i \in V} P(\alpha_i^t \in B_i / \xi^0 = x^0, \alpha^0 = a^0, \ldots, \xi^{t-1} = x^{t-1}, \alpha^{t-1} = a^{t-1}, \xi^t = x^t) = \\
&\qquad = \prod_{i \in V_1} \pi_i^t(B_i / x^0, a^0, \ldots, x^{t-1}, a^{t-1}, x^t) \prod_{j \in V_2} \gamma_j^t(B_j / x^0, a^0, \ldots, x^{t-1}, a^{t-1}, x^t), \\
&\qquad\qquad B_k \in \mathfrak{A}_k, \quad a^s \in A, \quad x^s \in X.
\end{aligned}$$

We now introduce various forms of structured strategies. We shall give the definition formally only for coalition K_1. Strategies for coalition K_2 are defined similarly.

Definition 2.6. (1) If at any time $t = 0, 1, \ldots$ the decision α_i^t for node i is made according to the probability π_i^t on basis of the history of the complete neighborhood $\widetilde{N}(i)$ of i only, i. e. on the basis of $h_i^t = (x_{\widetilde{N}(i)}^0, a_i^0, \ldots, x_{\widetilde{N}(i)}^{t-1}, a_i^{t-1}, x_{\widetilde{N}(i)}^t)$, and if $\pi_i^t(F_i^t(x)/h_i^t) = \pi_i^t(A_i^t(x_{\widetilde{N}(i)}^t)/h_i^t) = 1$, $x_{\widetilde{N}(i)}^s \in X_{\widetilde{N}(i)}$, $a_i^s \in A_i^s(x_{\widetilde{N}(i)}^s)$, then the sequence of transition probabilities (decisions) $\pi_i = \{\pi_i^t, \ t \in \mathbb{N}\}$ is called admissible local strategy for vertex i, and $\pi = (\pi_i : i \in V_1)$ is said to be locally admissible,.

(2) An admissible local strategy $\pi = (\pi_i, \ i \in V_1)$ is called admissible local Markov strategy for coalition K_1 if

$$\pi_i^t(\cdot / x_{\widetilde{N}(i)}^0, a_i^0, \ldots, x_{\widetilde{N}(i)}^{t-1}, a_i^{t-1}, x_{\widetilde{N}(i)}^t) = \pi_i^t(\cdot / x_{\widetilde{N}(i)}^t), \qquad i \in V_1.$$

(3) An admissible local Markov strategy $\pi = (\pi_i, \ i \in V_1)$ is called admissible stationary local strategy if $\pi_i^{t'}(\cdot / x_{\widetilde{N}(i)}) = \pi_i^{t''}(\cdot / x_{\widetilde{N}(i)})$, $i \in V_1$, for all t' and t'' and all x.

(4) An admissible deterministic (nonrandomized) stationary local strategy $\pi = (\pi_i, \ i \in V_1)$ is called admissible stationary local strategy if $\pi_i^t(\cdot / x_{\widetilde{N}(i)})$, $i \in V$, are one–point measures on $A_i^t(x_{\widetilde{N}(i)})$, $i \in V$, for each $x \in X$.

The class of all admissible local strategies is denoted by LS; the subclass of admissible local Markov strategies by LS_M. By $LS_S \subseteq LS_M$ we denote the class of admissible stationary local strategies with time invariant restriction sets: $\pi^t = \pi^{t'}$ for all $t, t' \in \mathbb{N}$ and $A^t(x) = A(x) = \times_{i \in V} A_i(x_{\widetilde{N}(i)})$ for all t, and by $LS_D \subseteq LS_S$ the set of admissible deterministic stationary local strategies.
If it is clear from the context we shall henceforth omit the *"admissible"* in the description of the strategies under consideration. $\odot$

Similar to (3) and (4) we shall construct our process in a way that the Markov property holds: The law of motion of the system is characterized by a set of time invariant transition probabilities. Whenever the state of the system is $\xi^t = x^t$ and decision $\alpha^t = a^t$ is taken, then the transition probability is $P\{\xi^{t+1} \in C/\xi^t = x^t, \alpha^t = a^t\} =: Q(C/x^t, a^t)$, and therefore independent of the past given the present (generalized) state. We further assume, that this transition probability is independent of t, the motion is homogeneous in time.
Applying a control policy to a time dependent Markov random field from Definition 2.2 we shall call the triple (ξ, π, γ) a controlled version of ξ using strategies π and γ. The controlled process $\xi = (\xi^t)$ in general is not Markovian because at first the sequence (α^t) of decisions depends not only on the actual local states $x^t_{\widetilde{N}(i)}$, $i \in V$, but on the previous local states $x^0_{\widetilde{N}(i)}, \ldots, x^{t-1}_{\widetilde{N}(i)}$ as well. An immediate consequence of the Definition 2.6 will be, that if Markov strategies π and γ are applied to control the time dependent Markov random field then we obtain a Markov process ξ from (ξ, π, γ).

Definition 2.7. A triple (ξ, π, γ) is called a controlled stochastic process with locally interacting synchronous components with respect to the finite interaction graph $\Gamma = (V, B)$ with coalitions K_1 (located at V_1) and K_2 (located at V_2), $V = V_1 \cup V_2$, $V_1 \cap V_2 = \varnothing$, if $\xi = (\xi^t : t \in \mathbb{N})$ is a stochastic process with state space $X = \times_{i \in V} X_i$, $\pi = (\pi_i : i \in V_1)$ and $\gamma = (\gamma_j : j \in V_2)$ are admissible local strategies, and the transitions of ξ are determined as given below in (5). If the strategies (π, γ) of both coalitions are stationary local Markov then the triple (ξ, π, γ) is called a controlled Markov process with locally interacting synchronous components and has a time invariant transition law. The rationale behind the construction of this transition law is along the following lines.

$$
\begin{aligned}
&P\{\xi_K^{t+1} \in C_K/\xi^0 = x^0, \alpha^0 = a^0, \ldots, \xi^{t-1} = x^{t-1}, \alpha^{t-1} = a^{t-1}, \xi^t = y, \alpha^t = a\} \\
&\quad \stackrel{(1)}{=} P\{\xi_K^{t+1} \in C_K/\xi^0 = x^0, \ldots, \xi^{t-1} = x^{t-1}, \xi^t = y, \alpha^t = a\} \\
&\quad \stackrel{(2)}{=} P\{\xi_K^{t+1} \in C_K/\xi^t = y, \alpha^t = a\} \\
&\quad \stackrel{(3)}{=} \prod_{j \in K} P\{\xi_j^{t+1} \in C_j/\xi^t = y, \alpha^t = a\} \\
&\quad \stackrel{(4)}{=} \prod_{j \in K} P\{\xi_j^{t+1} \in C_j/\xi^t_{\widetilde{N}(j)} = y_{\widetilde{N}(j)}, \alpha_j^t = a_j\} \\
&\quad \stackrel{(5)}{=} \prod_{j \in K} Q_j\{C_j/y_{\widetilde{N}(j)}, a_j\} \\
&\quad \stackrel{(6)}{=} Q_K(C_K/y, a), \qquad K \subseteq V, \quad y \in X, \quad a_j \in A_j(y_{\widetilde{N}(j)}).
\end{aligned} \tag{5}
$$

If $K = V$ we shall write $Q_V(C_V/y, a) =: Q(C/y, a)$.
The Markov kernel $Q = \prod_{j \in V} Q_j$ is said to be local and synchronuos. ξ will then shortly be called a controlled time dependent random field. $\odot$

Some remarks on the modeling principle behind this definition are in order:
It is assumed that the first expression is well defined, then $\overset{(1)}{=}$ introduces a memoryless (Markov) properties with respect to past decision rules into the definition, given the state history of the process. $\overset{(2)}{=}$ is the similar assumption of a Markovian transition law for the complete state changes. $\overset{(3)}{=}$ expresses that the coordinates act synchronously, and $\overset{(4)}{=}$ is due to the locality of the transition law with respect to the state information. Independence of time for the one–step transition probabilities then lead to the form $\overset{(5)}{=}$ of the Markov kernel, and $\overset{(6)}{=}$ again results from the coordinates acting synchronuosly.

At every time instant $t \in \mathbb{N}$ the players from both coalitions make decisions on the basis of the complete information about the state history of their full neighbourhoods and of all their own preceding decisions. It is possible that only players of one's own coalition or players of both coalitions belong to a player's neighbourhood. After the joint decision of all players the i-th player (at the i-th vertex) receives from the j-th player a (positive or negative) payment (reward) $r_{ij}(x_i, x_j, a_i, a_j)$, $x_i \in X_i$, $x_j \in X_j$, $a_i \in A_i$, $a_j \in A_j$, $i \in V$, $j \in N(i)$, where we assume $r_{ij} = -r_{ji}$. Thereafter the state of the system changes according to the transition law Q in (5).

The overall profits of coalitions K_1 and K_2 at each time instant are $r_1(x, a) = \sum\limits_{i \in V_1} \sum\limits_{j \in N(i)} r_{ij}(x_i, x_j, a_i, a_j)$ and $r_2(x, a) = \sum\limits_{j \in V_2} \sum\limits_{i \in N(j)} r_{ji}(x_j, x_i, a_j, a_i)$. We have a zero-sum game: $r_1(x, a) = -r_2(x, a)$, and the sum of payments inside each coalition equal zero. Therefore we write $r(x, a) = r_1(x, a) = -r_2(x, a)$.
We assume that the reward function r and the transition probabilities P are Borel measurable, and that furthermore

$$|r(x, a)| \leq L < \infty, \qquad x \in X, \quad a \in A.$$

The tupel (X, A, P, V_1, V_2, r) defines the stochastic game on neighborhood graph Γ.

The asymptotic average mean reward for initial state $\xi_V^0 = x$ of ξ, and when policies π, γ are applied is

$$\phi(x, \pi, \gamma) = \liminf_{n \to \infty} \frac{1}{n+1} E_{\pi,\gamma}^x \sum_{k=0}^{n} r(\xi^k, \alpha^k),$$

where $E_{\pi,\gamma}^x$ is the expectation associated with the controlled process (ξ, π, γ) under $\xi^0 = x$.

Definition 2.8. The strategy $\pi^\star \in LS$ is called optimal in the class LS of admissible local strategies for coalition K_1 if for all admissible local strategies γ' of coalition K_2 and all initial states $x \in X$ of the system

$$\phi(x, \pi^\star, \gamma') \geq \inf_{\gamma \in LS} \sup_{\pi \in LS} \phi(x, \pi, \gamma).$$

The strategy $\gamma^\star \in LS$ is called optimal in the class LS of admissible local strategies for coalition K_2 if for all admissible local strategies π' of coalition K_1 and all initial states

$x \in X$ of the system

$$\phi(x, \pi', \gamma^\star) \leq \sup_{\pi \in LS} \inf_{\gamma \in LS} \phi(x, \pi, \gamma).$$

The stochastic game has a value if

$$\inf_{\gamma \in LS} \sup_{\pi \in LS} \phi(x, \pi, \gamma) = \sup_{\pi \in LS} \inf_{\gamma \in LS} \phi(x, \pi, \gamma), \qquad x \in X.$$

$\odot$

2.3 Relation to stochastic dynamic optimization: Reduction to Markov strategies

To study a stochastic game on some graph we will introduce a reduced model (R-model) similarly to the case of stochastic 2–person games [27, 21, 13]. The idea behind this procedure is in the 2–person case, to fix the strategy of one of the players, say, player 2, and incorporate the then known stochastic effects of his decision into the transition kernel for the state transition mechanism. Then theorems, methods, and algorithms of standard stochastic dynamic optimization can be applied to optimize the decisions of player 1 to find at least a conditioned optimal behavior.
In the situation of our n–person games with two coalitions we mimic this by fixing the strategies for one coalition and then will end in the framework of control and optimization of Markov interacting processes, or time dependent random fields.

If A is a compact metric space, we denote by P_A the space of all probability measures on the Borel sets of A. It is well known that P_A, endowed with the weak topology, is a compact metric space. Throughout the paper, the space of probability measures on the Borel subsets of a compact metric space is endowed with the weak topology.

Recall $a = (a_{V_1}, a_{V_2})$ and put

$$r(x, \zeta, \lambda) = \iint r(x, a)\, d\zeta(a_{V_1})\, d\lambda(a_{V_2}),$$
$$Q(\cdot/x, \zeta, \lambda) = \iint Q(\cdot/x, a)\, d\zeta(a_{V_1})\, d\lambda(a_{V_2}), \qquad x \in X, \quad \zeta \in P_{A_{V_1}}, \quad \lambda \in P_{A_{V_2}}.$$

Let g be the fixed local stationary Markov strategy of coalition K_2, i. e. $g\colon \widetilde{X}_{V_2} \to P_{A_{V_2}}$, where $\widetilde{X}_S = \times_{i \in S} X_{\widetilde{N}(S)}$.

The R-model can be looked upon as a controlled Markov random field with discrete time with phase space X and action space A_{V_1} as defined in [6]). The reward function r' and the transition probability Q' have to be defined as follows:

$$r'(x, \zeta) = r(x, \zeta, g(\widetilde{x}_{V_2})),$$
$$Q'(C/x, \zeta) = Q(C/x, \zeta, g(\widetilde{x}_{V_2})), \qquad C \in \mathfrak{X}, \quad x \in X, \quad \widetilde{x}_{V_2} \in \widetilde{X}_{V_2}, \quad \zeta \in P_{A_{V_1}},$$

and the actions on nodes of K_2 are formally *"Do nothing !"*. If with this assumption π' is a strategy (similar to definition 2.6 with $V_1 \equiv V$, or according to the definitions from [6]) in the R-model, the asymptotic average mean reward for initial state $\xi_V^0 = x$ of ξ, and when policy π' is applied, is

$$\phi'(x, \pi') = \liminf_{n\to\infty} \frac{1}{n+1} E^x_{\pi'} \sum_{k=0}^{n} r'(\xi^k, \alpha^k_{V_1}),$$

where $E^x_{\pi'}$ is the expectation associated with the controlled process (ξ, π'), see [6].

The following lemma follows immediately from the definition of the R-model.

Lemma 2.1. If π is stationary local Markov strategy of coalition K_1 in the stochastic game then π is deterministic stationary local strategy in the R-model and $\phi'(x, \pi) = \phi(x, \pi, g)$, $x \in X$.

Conversely, if π is deterministic stationary local strategy in the R-model then π is a stationary local Markov strategy of coalition K_1 in the stochastic game and $\phi(x, \pi, g) = \phi'(x, \pi)$, $x \in X$. $\odot$

3 Optimal strategies for locally acting players

3.1 Stationary local Markov strategies

We are now ready to show that at least for any coaltion and for any local strategy of this coalition there exists a (conditioned on the other coalition playing a fixed Markov strategy) Markovian local policy which reproduces the long time average reward for the game.

Lemma 3.1. Assume that coalition K_2 applies the fixed local stationary Markov strategy g. If π is any admissible local strategy of coalition K_1 in the stochastic game then there exists a local Markov strategy $\pi^\star$ of coalition K_1 such that $\phi(x, \pi^\star, g) = \phi(x, \pi, g)$, $x \in X$. $\odot$

Proof. The proof will use arguments borrowed from the proofs in the framework of standard controlled processes (see e.g. Theorem 4.1 in [29]) and standard stochastic games (see e.g. Lemma 3.2 in [21]).

Let $e_\pi(\cdot/x^0)$ be the conditional distribution of the entire history $h = (x^0, a^0, x^1, a^1, \ldots)$ of the system, given the initial state $x^0 \in X$, where coalition K_1 uses π and coalition K_2 uses g. For each $n \geq 0$ the conditional distribution under $e_\pi(\cdot/x^0)$ of $a^n_i \in A_i$, given x^n, is denoted by $\overset{\star}{\pi}{}^n_i(a^n_i/x^n_{\widetilde{N}(i)})$, $i \in V_1$, and

$$\overset{\star}{\pi}{}^n(a^n/\widetilde{x}^n_{V_1}) = \prod_{i\in V_1} \overset{\star}{\pi}{}^n_i(a^n_i/x^n_{\widetilde{N}(i)}).$$

Let $\pi^\star = (\overset{\star}{\pi}{}^0, \overset{\star}{\pi}{}^1, \ldots)$. Then $\pi^\star$ is the required Markov strategy for coalition K_1. To see this, denote the conditional distribution of the entire history $h = (x^0, a^0, x^1, a^1, \ldots)$ of the system by $e_{\pi^\star}(\cdot/x^0)$, given the initial state is x^0, when coalition K_1 uses $\pi^\star$ and coalition K_2 uses g. First, we will prove that for every bounded measurable function p on $X \times A$, and every $n \geq 0$

$$\int p(x^n, a^n)\, de_\pi(a^n/x^0) = \int p(x^n, a^n)\, de_{\pi^\star}(a^n/x^0) \qquad (6)$$

for every $x^0 \in X$.

The proof proceeds by induction. For $n = 0$ note that $\pi^0 = \overset{\star}{\pi}{}^0$. Consequently,

$$\begin{aligned}\int p(x^0, a^0)\, de_\pi(a^0/x^0) &= \iint p(x^0, a^0)\, d\pi^0(a^0_{V_1}/\widetilde{x}^0_{V_1})\, dg(\widetilde{x}^0_{V_2})(a^0_{V_2}) \\ &= \iint p(x^0, a^0)\, d\overset{\star}{\pi}{}^0(a^0_{V_1}/\widetilde{x}^0_{V_1})\, dg(\widetilde{x}^0_{V_2})(a^0_{V_2}) \\ &= \int p(x^0, a^0)\, de_{\pi^\star}(a^0/x^0).\end{aligned}$$

This proves that (6) holds for $n = 0$.
Suppose now that (6) holds for all $n < N$ and for all bounded Borel functions on $X \times A$. Below, all conditional expectations are relative to $e_\pi(\cdot/x^0)$. We have

$$\begin{aligned}\int p(x^N, a^N)\, de_\pi(a^n/x^0) &= E_{x_0} p(\xi^N, \alpha^N) \\ &= E_{x_0}(E_{x_0}[p(\xi^N, \alpha^N)/\xi^N]) \\ &= E_{x_0} u(\xi^N),\end{aligned}$$

where

$$\begin{aligned}u(\xi^N) &= E_{x_0}[p(\xi^N, \alpha^N)/\xi^N] \\ &= \iint p(x^N, a^N)\, d\overset{\star}{\pi}{}^N(a^N_{V_1}/\widetilde{x}^N_{V_1})\, dg(\widetilde{x}^N_{V_2})(a^N_{V_2}).\end{aligned}$$

Thus,

$$\begin{aligned}\int p(x^N, a^N)\, de_\pi(\cdot/x^0) &= E_{x_0}(E_{x_0}[u(\xi^N)/\xi^{N-1}, \alpha^{N-1}]) \\ &= E_{x_0} w(\xi^{N-1}, \alpha^{N-1}),\end{aligned}$$

where

$$\begin{aligned}w(\xi^{N-1}, \alpha^{N-1}) &= E_{x_0}[u(\xi^N)/\xi^{N-1}, \alpha^{N-1}] \\ &= \iiint p(x^N, a^N)\, d\overset{\star}{\pi}{}^N(a^N_{V_1}/\widetilde{x}^N_{V_1})\, dg(\widetilde{x}^N_{V_2})(a^N_{V_2})\, dQ(x^{N-1}, a^{N-1}).\end{aligned}$$

Now, the induction hypothesis impliesis that

$$\begin{aligned}E_{x_0} w(\xi^{N-1}, \alpha^{N-1}) &= \int w(x^{N-1}, a^{N-1})\, de_\pi(\alpha^{N-1}/x^0) \\ &= \int w(x^{N-1}, a^{N-1})\, de_{\pi^\star}(\alpha^{N-1}/x^0).\end{aligned}$$

Consequently,

$$\begin{aligned}\int p(x^N, a^N)\, de_\pi(a^N/x^0) &= \int w(x^{N-1}, a^{N-1})\, de_{\pi^\star}(a^N/x^0) \\ &= \iiiint p(x^N, a^N)\, d\overset{\star}{\pi}{}^N(a^N_{V_1}/\widetilde{x}^N_{V_1}) \times \\ &\quad \times dg(\widetilde{x}^N_{V_2})(a^N_{V_2})\, dQ(x^N/x^{N-1}, a^{N-1})\, de_{\pi^\star}(a^{N-1}/x^0) \\ &= \int p(x^N, a^N)\, de_{\pi^\star}(a^N/x^0).\end{aligned}$$

Thus, (6) has been verified for $n = N$. By (6) we have

$$\begin{aligned}\phi(x,\pi,g) &= \liminf_{n\to\infty} \frac{1}{n+1} E_{\pi,g} \sum_{k=0}^{n} r(\xi^k,\alpha^k) \\ &= \liminf_{n\to\infty} \frac{1}{n+1} E_{\pi^\star,g} \sum_{k=0}^{n} r(\xi^k,\alpha^k) \\ &= \phi(x,\pi^\star,g), \qquad x\in X.\end{aligned}$$

The proof is completed. □

Remark 3.1. A similar statement as for coalition K_1 in Lemma 3.1 holds for policies of coalition K_2. ⊙

The next two theorems will under some technical conditions provide us with the existence of values for games and optimal stationary local Markov strategies. These conditions will enable us later on to prove general existence theorems under mild conditions on the data of the systems.

Theorem 3.1. Assume that for the stochastic game there exist admissible local stationary (Markov) strategies f and g for coalitions K_1 and K_2 respectively, a constant c and a bounded measurable function $v(x)$ on X such that

$$c + v(x) = \sup_{\zeta\in P_{A_{V_1}}} \Big\{ r(x,\zeta,g(\widetilde{x}_{V_2})) + \int v(y) Q(dy/x,\zeta,g(\widetilde{x}_{V_2})) \Big\}, \tag{7}$$

$$c + v(x) = \inf_{\delta\in P_{A_{V_2}}} \Big\{ r(x,f(\widetilde{x}_{V_1}),\delta) + \int v(y) Q(dy/x,f(\widetilde{x}_{V_1}),\delta) \Big\}, \tag{8}$$

for all $x \in X$. Then the stochastic game has a value which is identically equal to c and strategies f and g are optimal in LS for coalitions K_1 and K_2, respectively. ⊙

Proof. Because we restrict our computations to the set of local admissible strategies LS, so any $\sup_\gamma$, $\inf_\pi$, etc., is to be read with this restriction. Let coalition K_2 be allowed to use only strategy g. Then the optimality equation (13) in Theorem 3.2 of [6] implies that for any strategy π' in the R-model $\phi'(x,\pi') \le c$ holds, for all $x \in X$. From this and using Lemma 2.1 below we then obtain for any strategy π of coalition K_1 $\phi(x,\pi,g) \le c$, $x \in X$, in the stochastic game under consideration, i. e.

$$\sup_\pi \phi(x,\pi,g) \le c, \qquad x\in X, \tag{9}$$

and, consequently,

$$\inf_\gamma \sup_\pi \phi(x,\pi,\gamma) \le c, \qquad x\in X. \tag{10}$$

Similarly, fixing the strategy f for the coalition K_1 and considering an appropriate R-model we obtain

$$\inf_\gamma \phi(x,f,\gamma) \ge c, \qquad x\in X. \tag{11}$$

and, consequently,

$$\sup_{\pi} \inf_{\gamma} \phi(x,\pi,\gamma) \geq c, \qquad x \in X. \tag{12}$$

¿From (10), (12) and inequality

$$\inf_{\gamma} \sup_{\pi} \phi(x,\pi,\gamma) \geq \sup_{\pi} \inf_{\gamma} \phi(x,\pi,\gamma), \qquad x \in X,$$

we have

$$\inf_{\gamma} \sup_{\pi} \phi(x,\pi,\gamma) \equiv \sup_{\pi} \inf_{\gamma} \phi(x,\pi,\gamma) \equiv c.$$

We shall prove optimality of strategy f for coalition K_1. (7) and (8) implies that

$$c + v(x) = r(x, f(\widetilde{x}_{V_1}), g(\widetilde{x}_{V_2})) + \int v(y) Q(dy/x, f(\widetilde{x}_{V_1}), g(\widetilde{x}_{V_2})), \qquad x \in X. \tag{13}$$

We argue now as in the proof of Theorem 3.1 of [6] and obtain from (13) $\phi(x,f,g) \equiv c$. From this we see that in (9) and (11) equality holds. Therefore, for any strategy $\gamma' \in LS$ of coalition K_2

$$\begin{aligned} \phi(x,f,\gamma') &\geq \inf_{\gamma} \phi(x,f,\gamma) \\ &= c \\ &= \inf_{\gamma} \sup_{\pi} \phi(x,\pi,\gamma), \qquad x \in X. \end{aligned}$$

Optimality of strategy g for coalition K_2 is proved similarly. $\square$

For every function $u(x)$ in Banach space $M(X)$ of bounded measurable functions on X we denote

$$K_u(x,\zeta,\lambda) = r(x,\zeta,\lambda) + \int u(y) Q(dy/x,\zeta,\lambda), \qquad x \in X, \quad \zeta \in P_{A_{V_1}}, \quad \lambda \in P_{A_{V_2}},$$

Theorem 3.2. Let there exist a constant c and a function $v \in M(X)$ such that

$$c + v(x) = \sup_{\zeta \in P_{A_{V_1}}} \inf_{\lambda \in P_{A_{V_2}}} K_v(x,\zeta,\lambda), \qquad x \in X, \tag{14}$$

and assume further that the map F_v, defined by

$$F_u(x) = \big\{(\zeta',\lambda') \colon \inf_{\lambda \in P_{A_{V_2}}} K_u(x,\zeta',\lambda) = \sup_{\zeta \in P_{A_{V_1}}} K_u(x,\zeta,\lambda')\big\}, \tag{15}$$

is well defined, i.e., $F_v(x) \neq \emptyset$ and closed in $P_{A_1} \times P_{A_2}$, for all $x \in X$, and Borel measurable in the sense of Definition 2.3.

Then the stochastic game has a value which is constant c and both coalitions K_1 and K_2 have optimal stationary local strategies. $\odot$

Proof. Borel measurability of F_v by the Theorem of Choice guarantees the existence of Borel functions $f\colon \widetilde{X}_{V_1} \to P_{A_{V_1}}$ and $g\colon \widetilde{X}_{V_2} \to P_{A_{V_2}}$ for which

$$\inf_{\lambda \in P_{A_{V_2}}} K_v(x, f(\widetilde{x}_{V_1}), \lambda) = \sup_{\zeta \in P_{A_{V_1}}} K_v(x, \zeta, g(\widetilde{x}_{V_2})), \qquad x \in X.$$

¿From this we obtain

$$\begin{aligned} \sup_{\zeta \in P_{A_{V_1}}} \inf_{\lambda \in P_{A_{V_2}}} K_v(x,\zeta,\lambda) &\geq \inf_{\lambda \in P_{A_{V_2}}} K_v(x, f(\widetilde{x}_{V_1}), \lambda) \\ &= \sup_{\zeta \in P_{A_{V_1}}} K_v(x, \zeta, g(\widetilde{x}_{V_2})) \\ &\geq \inf_{\lambda \in P_{A_{V_2}}} \sup_{\zeta \in P_{A_{V_1}}} K_v(x,\zeta,\lambda), \qquad x \in X. \end{aligned} \tag{16}$$

Because

$$\inf_{\lambda \in P_{A_{V_2}}} \sup_{\zeta \in P_{A_{V_1}}} K_v(x,\zeta,\lambda) \geq \sup_{\zeta \in P_{A_{V_1}}} \inf_{\lambda \in P_{A_{V_2}}} K_v(x,\zeta,\lambda), \qquad x \in X,$$

all inequalities in (16) are replaced by equalities. Taking into account (14) we obtain (7) and (8), and by Theorem 3.1 the proof is completed. □

Remark 3.2. It is easy to see that for countable phase space X the Borel measurability of map F_v can be replaced by the condition that for all $x \in X$ the sets $F_v(x)$ are nonempty. ⊙

The proofs of the following Lemmas are borrowed from the proofs of Lemmas 2.1 and 3.1 in [21] for the 2–person games. They will enable us to prove and apply smoothness properties in our general setting.

Lemma 3.2. Let $p(x,a)$ be a continuous real-valued function on $X \times A$, where X and A are compact spaces. Then, $p(x,\zeta,\lambda) := \iint p(x,a)\, d\zeta(a_{V_1})\, d\lambda(a_{V_2})$, $x \in X$, $\zeta \in P_{A_{V_1}}$, $\lambda \in P_{A_{V_2}}$, is a continuous function on $X \times P_{A_{V_1}} \times P_{A_{V_2}}$. ⊙

Proof. Let $x^n \to x^0$ in X, $\zeta^n \to \zeta^0$ in $P_{A_{V_1}}$ and $\lambda^n \to \lambda^0$ in $P_{A_{V_2}}$; and let $\epsilon > 0$. We have

$$\begin{aligned} &|p(x^n,\zeta^n,\lambda^n) - p(x^0,\zeta^0,\lambda^0)| \leq \\ &\quad \leq \left| \iint p(x^n,a)\, d\zeta^n(a_{V_1})\, d\lambda^n(a_{V_2}) - \iint p(x^0,a)\, d\zeta^n(a_{V_1})\, d\lambda^n(a_{V_2}) \right| + \\ &\quad + \left| \iint p(x^0,a)\, d\zeta^n(a_{V_1})\, d\lambda^n(a_{V_2}) - \iint p(x^0,a)\, d\zeta^0(a_{V_1})\, d\lambda^0(a_{V_2}) \right|. \end{aligned} \tag{17}$$

As $X \times A$ is a compact space $p(x,a)$ is uniformly continuous. Hence, there exists N_1 such that whenever $n \leq N_1$ $|p(x^n,a) - p(x^0,a)| < \epsilon/2$ for all $a \in A$. Consequently, the first term on the right-hand side of (17) is at most $\epsilon/2$ for all $n \leq N_1$. Next, the product measures $\zeta^n \times \lambda^n$ on A converge weakly to $\zeta^0 \times \lambda^0$. To see this, use the following facts: (a) every real-valued continuous function f on A can be approximated uniformly by a sequence of function of the form $\sum_{i=1}^{k} f_i g_i$, where f_i are continuous, real-valued functions on

A_{V_1} and g_i are continuous, real-valued functions on A_{V_2} (Stone-Weierstrass theorem); (b) if f, g are continuous real-valued functions on A_{V_1}, A_{V_2}, respectively, then

$$\iint f(a_{V_1})g(a_{V_2})\, d\zeta^n(a_{V_1})\, d\lambda^n(a_{V_2}) \to \iint f(a_{V_1})g(a_{V_2})\, d\zeta^0(a_{V_1})\, d\lambda^0(a_{V_2}).$$

Since $p(x^0,\cdot)$ is a continuous function on A henceforth there exists N_2 such that wherever $n \leq N_2$ the second term on the right-hand side of (17) is at most $\epsilon/2$. Since ϵ is arbitrary this proves that $p(x^n,\zeta^n,\lambda^n) \to p(x^0,\zeta^0,\lambda^0)$. □

Lemma 3.3. Assume X, $P_{A_{V_1}}$, $P_{A_{V_2}}$ are compact spaces, and let u be a bounded, continuous function on $X \times P_{A_{V_1}} \times P_{A_{V_2}}$. Then the functions $u_\star(x) = \max\limits_{\zeta\in P_{A_{V_1}}} \min\limits_{\lambda\in P_{A_{V_2}}} u(x,\zeta,\lambda)$ and $u^\star(x) = \min\limits_{\lambda\in P_{A_{V_2}}} \max\limits_{\zeta\in P_{A_{V_1}}} u(x,\zeta,\lambda)$ are continuous. ⊙

Proof. At first we shall prove that function $u'(x,\lambda) = \max\limits_{\zeta\in P_{A_{V_1}}} u(x,\zeta,\lambda)$ is continuous. Let $(x^n,\lambda^n) \to (x^0,\lambda^0)$ in $X \times P_{A_{V_2}}$. Since $\{u'(x^n,\lambda^n),\ n\geq 1\}$ is a bounded sequence of real numbers it has a convergent subsequence. We will prove that every convergent subsequence of $\{u'(x^n,\lambda^n),\ n\geq 1\}$ goes to $u'(x^0,\lambda^0)$. Assume then that $\{u'(x^{n'},\lambda^{n'})\}$ is a convergent subsequence. Then there exists $\zeta^{n'} \in P_{A_{V_1}}$ such that $u'(x^{n'},\lambda^{n'}) = u(x^{n'},\zeta^{n'},\lambda^{n'})$. As $P_{A_{V_1}}$ is compact there exists subsequence $\{\zeta^{n''}\}$ of $\{\zeta^{n'}\}$ such that $\zeta^{n''} \to \zeta^0$ (say). It follows from the continuity of u that $u(x^{n''},\zeta^{n''},\lambda^{n''}) \to u(x^0,\zeta^0,\lambda^0)$. Moreover, if $\zeta \in P_{A_{V_1}}$, $u(x^{n''},\zeta,\lambda^{n''}) \leq u(x^{n''},\zeta^{n''},\lambda^{n''})$ and so $u(x^0,\zeta,\lambda^0) \leq u(x^0,\zeta^0,\lambda^0)$, i. e. $u'(x^0,\lambda^0) = u(x^0,\zeta^0,\lambda^0)$. Hence, $u'(x^{n''},\lambda^{n''}) \to u'(x^0,\lambda^0)$ and so $u'(x^{n'},\lambda^{n'}) \to u'(x^0,\lambda^0)$. This proves that $u'(x^n,\lambda^n) \to u'(x^0,\lambda^0)$.

Similarly, the function $u''(x,\zeta) = \min\limits_{\lambda\in P_{A_{V_2}}} u(x,\zeta,\lambda)$ is continuous and, therefore, the functions $u_\star(x) = \max\limits_{\zeta\in P_{A_{V_1}}} \min\limits_{\lambda\in P_{A_{V_2}}} u(x,\zeta,\lambda)$ and $u^\star(x) = \min\limits_{\lambda\in P_{A_{V_2}}} \max\limits_{\zeta\in P_{A_{V_1}}} u(x,\zeta,\lambda)$ are continuous. □

3.2 Existence of optimal local policies and values of the game

We have now prepared the ground for to state and prove our main theorems. We need the following standard regularity assumptions for the transition kernels which will provide conditions to prove optimality properties of local strategies. The proofs of the theorems utilize ideas from the proofs of Theorems 3 and 4 in [13]. Foundations for the development in that paper and technical details can be found in [14].

Assumption I. $Q = \prod_{j\in V} Q_j$ is a local and synchronous Markov kernel according to Definition 2.7 for (ξ,π,γ) and the admissible control sets are independent of time: $A(x) := A^t(x) = \times_{i\in V} A_i^t(x_{\tilde{N}(i)}) =: \times_{i\in V} A_i(x_{\tilde{N}(i)})$.

There exists a non-negative measure μ on $(X,\mathfrak{X})$ with $\mu(X) > 0$ such that:
$\mu(C) \leq Q(C/x,a), \qquad C\in\mathfrak{X}, \quad (x,a)\in\Delta.$ ⊙

Assumption II. $Q = \prod_{j\in V} Q_j$ is a local and synchronous Markov kernel according to Definition 2.7 for (ξ,π,γ) and the admissible control sets are independent of time: $A(x) := A^t(x) = \times_{i\in V} A_i^t(x_{\tilde{N}(i)}) =: \times_{i\in V} A_i(x_{\tilde{N}(i)})$.

There exists a non-negative measure η on $(X,\mathfrak{X})$ with $\eta(X) < 2$ such that: $\eta(C) \geq Q(C/x,a), \qquad C \in \mathfrak{X}, \quad (x,a) \in \Delta.$ $\odot$

Theorem 3.3. Assume that X, A are compact spaces, $r(x,a)$ is continuous on $X \times A$, and the transition probability $Q(\cdot/x,a)$ is weakly continuous as a function on $X \times A$. If Assumption I or II holds, then the stochastic game has a value which is constant and coalitions K_1 and K_2 have optimal stationary local (Markovian) strategies. $\odot$

Proof. Let Assumption I be fulfilled. Define an operator U on the Banach space $C(X)$ of continuous bounded functions on X by

$$Uu(x) = \sup_{\zeta \in P_{A_{V_1}}} \inf_{\lambda \in P_{A_{V_2}}} K_u(x,\zeta,\lambda) - \int u(y)\, d\mu(y). \tag{18}$$

Lemmas 3.2 and 3.3 imply that the functions $K_u(x,\zeta,\lambda)$ and $\sup\limits_{\zeta \in P_{A_{V_1}}} \inf\limits_{\lambda \in P_{A_{V_2}}} K_u(x,\zeta,\lambda)$ are continuous on $X \times P_{A_{V_1}} \times P_{A_{V_2}}$, resp. X, for every function $u \in C(X)$ and, consequently, sup and inf can be replaced by max and min respectively ($P_{A_{V_1}}$ and $P_{A_{V_2}}$ are compact). Using Assumption I just as in Theorem 3.3 of [6] it is easy to check that the operator U is a contraction with contraction factor $1-\mu(X)$ ($0 \leq 1-\mu(X) < 1$). From Banach's Fixed Point Theorem there exists in $C(X)$ a function $u^\star(x)$ such that $Uu^\star = u^\star$. From (18) we obtain relation (14) with $v(x) = u^\star(x)$ and $c = \int u^\star(y)\, d\mu(y)$.

We shall show that the map $F_{u^\star}\colon X \to (2)_{set}^{P_{A_{V_1}} \times P_{A_{V_2}}}$ defined by (14) is upper semicontinuous and, therefore, measurable. For all $x \in X$ the function $K_{u^\star}(x,\zeta,\lambda)$ on $P_{A_{V_1}} \times P_{A_{V_2}}$ satisfies all conditions of minimax Theorem (Theorem 1 of [8]) and, consequently,

$$\max_{\zeta \in P_{A_{V_1}}} \min_{\lambda \in P_{A_{V_2}}} K_{u^\star}(x,\zeta,\lambda) = \min_{\lambda \in P_{A_{V_2}}} \max_{\zeta \in P_{A_{V_1}}} K_{u^\star}(x,\zeta,\lambda). \tag{19}$$

(19) together with the continuity of the functions $\max\limits_{\zeta \in P_{A_{V_1}}} K_{u^\star}(x,\zeta,\lambda)$ and $\min\limits_{\lambda \in P_{A_{V_2}}} K_{u^\star}(x,\zeta,\lambda)$ (see the proof of Lemma 3.3) imply that the sets $F_{u^\star}(x) \subseteq P_{A_{V_1}} \times P_{A_{V_2}}$ are nonempty and closed in the product topology of $P_{A_{V_1}} \times P_{A_{V_2}}$. To prove semicontinuity of the map $F_{u^\star}$ it is enough to show that

$$\lim_{n\to\infty} (\zeta^n, \lambda^n) = (\zeta', \lambda'), \quad \lim_{n\to\infty} x^n = x, \quad (\zeta^n,\lambda^n) \in F_{u^\star}(x^n) \tag{20}$$

implies

$$(\zeta',\lambda') \in F_{u^\star}(x). \tag{21}$$

From (20) we have

$$\max_{\zeta \in P_{A_{V_1}}} K_{u^\star}(x^n,\zeta,\lambda^n) = \min_{\lambda \in P_{A_{V_2}}} K_{u^\star}(x^n,\zeta^n,\lambda). \tag{22}$$

Passing to the limit as $n \to \infty$ from continuity of both sides of (22) we obtain

$$\max_{\zeta \in P_{A_{V_1}}} K_{u^\star}(x,\zeta,\lambda') = \min_{\lambda \in P_{A_{V_2}}} K_{u^\star}(x,\zeta',\lambda),$$

which is equivalent to (21).

Theorem 3.2 implies that the stochastic game has a constant value and coalitions K_1 and K_2 have optimal stationary local strategies.

When Assumption II is fulfilled instead of Assumption I the proof will be carried out similarly using an operator U' operating in $C(X)$ by

$$U'u(x) = - \max_{\zeta \in P_{A_{V_1}}} \min_{\lambda \in P_{A_{V_2}}} K_u(x, \zeta, \lambda) - \int u(y)\, d\eta(y).$$

The proof is completed. $\square$

Note that the resulting (in general random) strategies of the Theorem 3.3 can be chosen as smooth versions: By the Theorem of Choice for semicontinuous maps the functions f and g in Theorem 3.2 here can be chosen from Baire class 1 of $P_{A_{V_i}}-$ valued functions on $\widetilde{X}_{V_i}$.

If the state space X is countable, we write $q(y/x,a) = Q(\{y\}/x,a)$, $x, y \in X$, $a \in A$. We have the following Theorem.

Theorem 3.4. Assume X is a countable space and A is compact. Assume further that the functions $r(x,a)$ and $q(y/x,a)$ are continuous on A for every fixed $x, y \in X$, and that Assumption II holds.

Then the stochastic game has a constant value and both coalitions K_1 and K_2 have optimal stationary local strategies. $\odot$

Proof. The existence of the constant c and the bounded function $v(x)$ satisfying (14) is proved by applying a fixed-point theorem as in Theorem 3.3, replacing space $C(X)$ by the space of bounded sequences.

Assumption II implies that series $\sum_{y \in X} v(y) q(y/x,a)$ converges uniformly in x, a (Weierstrass criterion). Therefore, in every fixed $x \in X$ function

$$K_v(x,a) = r(x,a) + \sum_{y \in X} q(y/x,a)$$

is continuous on A and by Lemma 3.2 the function

$$K_v(x, \zeta, \lambda) = \iint K_v(x,a)\, d\zeta(a_{V_1})\, d\lambda(a_{V_2})$$

is continuous on $P_{A_{V_1}} \times P_{A_{V_2}}$. As in Theorem 3.3 this implies nonemptiness (and closedness) of the sets $F_v(x)$ for every $x \in X$. By Theorem 3.2 and Remark 3.2 the proof is completed. $\square$

Remark 3.3. The results formulated in this section remain valid for stochastic games on graphs where the reward function r for each coalition depends on the states of all vertices . $\odot$

References

[1] Altman E (1996) Nonzero-sum stochastic games in admission, service and routing control in queueing systems. QUESTA 23:259–279

[2] Altman E (1999) A Markov game iapproach for optimal routing into a queueing network. In: Bardi M, Raghavan TES, Parthasarathy T (Eds.) Stochastic and Differential Games. Annals of the Internatioanl Society of Dynamic Games, Vol 4. 359–376. Birkhäuser, Boston

[3] Altman E, Hordijk, A (1995) Zero-sum Markov games and worst–case optimal control of queueing systems. QUESTA 21:415–447

[4] Anagnostopoulos A, Kontoyiannis I, Upfal E (2002) Steady state analysis of balanced-allocation routing. Preprint Computer Science Department, Brown University, Providence

[5] Chornei RK, Daduna H, Knopov PS (2000) Controlled Markov fields with finite state space on graphs. Preprint 00-08, Inst. Math. Stochastics, University of Hamburg, to appear: Stochastic Models

[6] Daduna H, Knopov PS, Chornei RK (2001) Local control of interacting Markov processes on graphs with compact state space. Kibernetika i systemny analiz 3:62–77 [in Russian]

[7] David PA, Foray D (1992) Percolation structures, Markov random fields. The economics of EDI standards diffusion. Center for Economical Policy Research, Stanford University

[8] Fan Ky (1953) Minimax theorems. Proceeding of the National Academy of Science U.S.A. 39:42–47

[9] Federgruen A (1978) On n-person stochastic games with denumerable state space. Adv. Appl. Probab. 10:452–471

[10] Filar J, Vrieze K (1996) Competitive Markov decision Processes. Springer, New York

[11] Frid EB (1973) On stochastic games. Teor. Verojatnost. i ejo Primenen. 18:408–413 [in Russian]

[12] Gibbens RJ, Kelly FP, Key PB (1995) Dynamic alternative routing. In: Steenstrup ME (Ed), Routing in Communications Networks: 13–47, Prentice Hall

[13] Gubenko LG (1972) On multistage stochastic games. Teor. Verojatnost. i Mat. Stat. 8:35–49 [in Russian]

[14] Gubenko LG, Shtatland ES (1972) On discrete time Markov decision processes. Teor. Verojatnost. i Mat. Stat. 7:51–64, English translation in: Theory of Probability and Mathematical Statistics 7:51–64 (1975)

[15] Heyman DP, Sobel MJ (1984) Stochastic Models in Operations Research, Vol II. McGraw-Hill, New York

[16] Kelly FP (1979) Reversibility and Stochastic Networks Wiley, Chichester

[17] Knopov PS (1996) Markov fields and their applications in economics. Obchysljuvalna ta prykladna matematyka 80:33–46 [in Russian]

[18] Kuratowski K (1969) Topology 2. Academic Press, New York

[19] Liggett T (1989) Interacting particle systems. Springer-Verlag, New York

[20] Liggett T (1999) Stochastic interacting systems: Contact, voter, and exclusion processes. Springer-Verlag, Berlin

[21] Maitra A, Parthasarathy T (1970) On Stochastic Games. J. Optimization Theory and Appl. 5:289–300

[22] Maitra AP, Sudderth WD (1996) Discrete Gambling and Stochastic Games. Springer-Verlag, New York

[23] Myerson RB (1984) Cooperative games with incomplete information. Internat. J. Game Theory 13:69–96

[24] Nowak AS, Szajowski K (1999) Non-zero sum stochastic games. In: Bardi M, Raghavan TES, Parthasarathy T (Eds.) Stochastic and Differential Games. Annals of the Internatioanl Society of Dynamic Games, Vol 4. 297 - 342. Birkhäuser, Boston

[25] Preston C (!974) Gibbs States on Countable Sets. Cambridge University Press,London

[26] Ruelle D (1969) Statistical mechanics. WA Benjamin Inc., New York Amsterdam

[27] Shapley LS (1953) Stochastic games. Proc. Nat. Acad. Sci. USA 39:1095–1100

[28] Sinay YaG (1980) The Theory of phase transitions. Nauka, Moscow [in Russian]

[29] Strauch RE (1966) Negative dynamic programming. Ann. Math. Statist. 37:871–890

[30] Vasilyev, NB (1978) Bernoulli and Markov stationary measures in discrete local interactions. in: Dobrushin RL, Kryukov VI, Toom AL (ed.)(1978) Locally Interacting Systems and Their Application in Biology. Lecture Notes in Mathematics 653 Springer Berlin, pp. 99–112

[31] Yao, D (1995) S-modular games, with queueing applications. Queueing Systems and Their Applications 21, 449 - 475