

Article



On a Simplified Method of Defining Characteristic Function in Stochastic Games

Elena Parilina *,[†] and Leon Petrosyan [†]

Department of Mathematical Game Theory and Statistical Decisions, Saint Petersburg State University, 7/9 Universitetskaya nab., Saint Petersburg 199034, Russia; l.petrosyan@spbu.ru

* Correspondence: e.parilina@spbu.ru

+ These authors contributed equally to this work.

Received: 29 May 2020; Accepted: 9 July 2020; Published: 11 July 2020



Abstract: In the paper, we propose a new method of constructing cooperative stochastic game in the form of characteristic function when initially non-cooperative stochastic game is given. The set of states and the set of actions for any player is finite. The construction of the characteristic function is based on a calculation of the maximin values of zero-sum games between a coalition and its anti-coalition for each state of the game. The proposed characteristic function has some advantages in comparison with previously defined characteristic functions for stochastic games. In particular, the advantages include computation simplicity and strong subgame consistency of the core calculated with the values of the new characteristic function.

Keywords: cooperative stochastic game; strong subgame consistency; characteristic function; core

1. Introduction

When a non-cooperative game is initially defined, the problem of construction of a cooperative version of the game is actual if players start acting as a unique coalition to maximize their joint payoff or minimize joint costs. The classical approach is to define cooperative game in a form of characteristic function that assigns the value for any coalition of players. Subsequently, based on this function one can calculate the imputation of the joint payoff allocating it among players. The component of the imputations may vary if we calculate them based on different characteristic functions. Therefore, the way of defining this function is important and it has influence on the players' payoffs in cooperative game. Moreover, some approaches to define characteristic function make it impossible to apply in dynamic or differential games because of computational difficulties. Additionally, the way of constructing characteristic function also influences on the consistency properties of cooperative solutions that are realized in dynamics.

The choice of the approach on how to define characteristic function also depends on the background of the considered problem if it arises from an applied area. The existence and uniqueness issues are also actual when one chooses the way of constructing characteristic function. There exist different approaches that can be applied to stochastic game. The so-called maxmin and minmax approaches define the value of the function for coalition *S* as maxmin and minmax payoff of coalition *S* in zero-sum game against coalition of all left-out players [1,2]. Another approach is proposed in [3,4] when the value of coalition *S* is defined as its payoff in the Nash equilibrium in the non-cooperative game between coalition *S* and left-out players acting individually. The calculation of characteristic function in two-step procedure is proposed in [5], in which the authors first find an *n*-player non-cooperative equilibrium and then allow coalition *S* to optimize its payoff, assuming that left-out players use their Nash equilibrium actions found at the first step. The properties of this function are examined in [6,7]. Another two-stage approach for defining characteristic function

is proposed in [8], in which the strategies maximizing total payoff of the players are first found. Subsequently, these strategies are used by the players from coalition *S*, while the out-coalition players use the strategies minimizing the total payoff of players from *S*. The joint payoff of players from the coalition equals the value of characteristic function for this coalition.

The new simplified method of constructing characteristic function in multistage games is introduced in [9]. They examine the properties of this function and proved that the corresponding core is strongly subgame-consistent in multistage game. This property cannot be proved in general case when the characteristic function is constructed with the classical approaches, like maxmin or minmax.

In the paper, we adopt the method of constructing the characteristic function proposed in [9] to stochastic games. Based on the values of the characteristic function, one can determine the core. Moreover, the core satisfies the strong subgame consistency property, which is a refinement of subgame consistency on the case of set-valued cooperative solutions. The problem of subgame consistency is originally examined for differential games in [10,11]. The construction of a special payment scheme, called imputation distribution procedure (see [11]), allows for coping with the problem of time inconsistency of cooperative solutions. This problem is described for stochastic games in [12–14] in the case of unique-valued cooperative solutions. The node-consistent core is constructed in dynamic games played over event trees in [15]. The strong subgame consistency of the set-valued cooperative solution, like the core, guarantees players to obtain, in total, the solution from initially defined core. It means that, in any intermediate time period, the solution is the sum of obtained payments up to the current period, and the core elements of subgame starting from the next time period. The strong subgame consistency condition is proposed in [16]. The subcore satisfying strong subgame property is constructed for multistage games in [17]. The problem of subgame consistency is actual for different classes of dynamic and differential games and it is examined in [18] for stochastic games with finite duration, in [19] for differential games with finite time horizon, in [20] for multistage games. In the paper, we construct characteristic function for stochastic game in a special way and calculate the core while using the values of this function. The core satisfies strong subgame consistency property. To prove this result, we define the imputation distribution procedure, which determines the payments to the players in any state realized in the game process.

The rest of the paper is organized, as follows. We describe the model of stochastic games in Section 2.1. In Section 2.2, we define the new approximated characteristic function for stage games, and then extend this approach to the case of stochastic game in Section 2.3. We formulate the definition of the imputation distribution procedure for stochastic games and describe the idea of strongly subgame consistency of the core in Section 3. We briefly conclude in Section 4.

2. Cooperative Stochastic Games

2.1. Model

Consider a non-cooperative stochastic game *G* given by

$$G = \left(N, \Omega, \{ \Gamma(\omega) \}_{\omega \in \Omega}, \pi_0, \left\{ p(\omega'' | \omega', a^{\omega'}) \right\}_{a^{\omega'} \in \prod_{i \in N} A_i^{\omega'}}, \delta \right),$$
(1)

where

- $N = \{1, \dots, n\}$ is the set of players.
- $\Omega = \{\omega_1, \dots, \omega_k\}$ is the finite set of states.
- Γ(ω) is the game in normal form associated with state ω. The set of players N is common for any state ω. Let A_i^ω be a finite set of actions of player i ∈ N in state ω, a_i^ω ∈ A_i^ω be an action of player i ∈ N in this state; K_i^ω : Π_{i∈N} A_i^ω → ℝ be a payoff function of player i in state ω.
- $p(\cdot|\omega, a^{\omega}) : \Omega \times A^{\omega} \to \Delta(\Omega)$ is a transition function from state ω when action profile $a^{\omega} \in \prod_{i \in N} A_i^{\omega}$ is realized, where $\Delta(\Omega)$ is a probability distribution over set Ω .

• $\pi_0 = (\pi_0^{\omega_1}, \dots, \pi_0^{\omega_k})$ is an initial state distribution.

• $\delta \in (0, 1)$ is a common discount factor.

Denote by G^{ω} the subgame of *G* starting from state ω defined by (1) with π_0 , such that $\pi_0^{\omega} = 1$ and $\pi_0^{\omega'} = 0$ for any state $\omega' \neq \omega$.

We assume that, in stochastic game *G*, the set of any player's strategies H_i is stationary. The stationary strategy of player *i* is η_i assigning action (maybe mixed) $a_i \in \Delta(A_i^{\omega})$ to any state ω . The vector $(\eta_1, \ldots, \eta_n) \in \prod_{j \in N} H_j$ is a stationary strategy profile in stochastic game *G*. It is obvious that a stationary strategy η_i of player $i \in N$ in game *G* is the stationary strategy of this player in any subgame G^{ω} .

By the payoff of player *i*, we assume the expected payoff in stochastic subgame G^{ω} given by

$$E_i^{\omega}(\eta) = K_i^{\omega}(a^{\omega}) + \delta \sum_{\omega' \in \Omega} p(\omega'|\omega, a^{\omega}) E_i^{\omega'}(\eta).$$
⁽²⁾

where $\eta \in H = \prod_{j \in N} H_j$ is a stationary strategy profile such that $\eta(\omega) = a^{\omega} \in \prod_{j \in N} A_j^{\omega}$. We rewrite Equation (2) in a vector form and obtain

$$E_i(\eta) = K_i(a) + \delta \Pi(\eta) E_i(\eta), \tag{3}$$

where $E_i(\eta) = (E_i^{\omega_1}(\eta), \dots, E_i^{\omega_k}(\eta))'$, $K_i(a) = (K_i^{\omega_1}(a^{\omega_1}), \dots, K_i^{\omega_k}(a^{\omega_k}))'$. A matrix of transition probabilities is formed in the following way

$$\Pi(\eta) = \begin{pmatrix} p(\omega_1|\omega_1, a^{\omega_1}) & \dots & p(\omega_k|\omega_1, a^{\omega_1}) \\ p(\omega_1|\omega_2, a^{\omega_2}) & \dots & p(\omega_k|\omega_2, a^{\omega_2}) \\ \dots & \dots & \dots \\ p(\omega_1|\omega_k, a^{\omega_k}) & \dots & p(\omega_k|\omega_k, a^{\omega_k}) \end{pmatrix}$$
(4)

in which each row contains transition probabilities from a corresponding state.

Equation (3) implies the explicit formula to calculate the expected payoff of player *i* when the stationary strategy profile η is realized:

$$E_i(\eta) = \left(\mathbb{I}_k - \delta \Pi(\eta)\right)^{-1} K_i(a),$$

where \mathbb{I}_k is an identity matrix of size $k \times k$. Inverted matrix $(\mathbb{I}_k - \delta \Pi(\eta))^{-1}$ always exists for discount factor $\delta \in (0, 1)$.

Taking into account the probability distribution π_0 , we calculate the expected payoff in game G, as

$$\bar{E}_{i}(\eta) = \pi_{0} E_{i}(\eta) = \pi_{0} \left(\mathbb{I}_{k} - \delta \Pi(\eta)\right)^{-1} K_{i}(a).$$
(5)

If players cooperate, they find the cooperative strategy profile η^* maximizing the total expected payoff, which is

$$\eta^* = \arg \max_{\eta \in H} \sum_{i \in N} \bar{E}_i(\eta).$$

We should notice that η^* is a pure stationary strategy profile. The profile η^* is such that $\eta_i^*(\omega) = a_i^{\omega^*} \in A_i^{\omega}, \omega \in \Omega$. We also assume that the profile η^* is such that $\max_{\eta \in H} \sum_{i \in N} E_i^{\omega}(\eta) = \sum_{i \in N} E_i^{\omega}(\eta^*)$ for any state $\omega \in \Omega$, which means that the cooperative strategy profile maximizes the total payoff of the players independently of which state is initial. This assumption is usually satisfied for most stochastic games.

To define cooperative game when the non-cooperative stochastic game is given, we use the classical approach and define it in the form of characteristic function $v : 2^N \to \mathbb{R}^1$ whose values

estimate the "power" of any coalition or the subset of players. In [21], the characteristic function value for coalition *S* in subgame starting at any state ω is defined in as maxmin value, which is

$$v^{\omega}(S) = val \, G_S^{\omega},\tag{6}$$

where G_S^{ω} is a zero-sum stochastic subgame starting at state ω , in which coalition *S* is a maximizing player, coalition $N \setminus S$ is a minimizing player. Existence of the value of game G_S^{ω} for stochastic games is proved in [22].

2.2. Approximated Characteristic Function for State Games

Before we define a characteristic function in a new form, we need to make additional calculations. First, we consider state games and propose a scheme of calculation of the approximated characteristic function values for any state. Define characteristic function for a state $\omega \in \Omega$ or one-shot game $\Gamma(\omega)$ given in normal form while using the maxmin approach:

$$v(\omega, S) = \max_{a_{S} \in \prod_{j \in S} A_{j}^{\omega}} \min_{a_{N \setminus S} \in \prod_{j \in N \setminus S} A_{j}^{\omega}} \sum_{i \in S} K_{i}^{\omega}(a_{S}^{\omega}, a_{N \setminus S}^{\omega}),$$
(7)

where maxmin in (7) is found in pure strategies.

Let $C(\omega)$ be a non-empty core in the game defined in state ω using c.f. (7), which is

$$C(\omega) = \left\{ (\alpha_1(\omega), \dots, \alpha_n(\omega)) : \sum_{i \in S} \alpha_i(\omega) \ge v(\omega, S), \forall S \subset N, \sum_{i \in N} \alpha_i(\omega) = v(\omega, N) \right\}$$
(8)

Remark 1. We assume that conditions under which the core $C(\omega)$ exists for any state ω are satisfied. The core $C(\omega)$ is non-empty if and only if for any function $\psi : 2^N \setminus \emptyset \to [0,1]$, where $\sum_{S \in 2^N: S \ni i} \psi(S) = 1$ for any $i \in N$, condition (see [23,24])

$$\sum_{S \in 2^N \setminus \emptyset} \psi(S) v(\omega, S) \le v(\omega, N)$$
(9)

holds. Characteristic function $v(\omega, S)$ is defined by (7). We refer to the book [25] for further discussion of non-emptiness of the core.

Second, for any coalition $S \subseteq N$ define maximal value of characteristic function (7) over set Ω :

$$\hat{w}(S) = \max_{\omega \in \Omega} v(\omega, S), \tag{10}$$

which is the maximal value that coalition *S* can obtain in state games.

S

The next step is to define the approximated value of the characteristic function for any state in the following way. Let for any state $\omega \in \Omega$ the approximated characteristic function $w(\omega, S)$ be given as

$$w(\omega, S) = \begin{cases} \sum_{i \in S} K_i^{\omega}(a^{\omega*}), & \text{if } S = N, \\ \hat{w}(S), & \text{if } S \neq N. \end{cases}$$
(11)

In Equation (11), the summarized payoff of the players adopting cooperative action profile $a^{\omega *}$ is assigned to the grand coalition. The approximated (maximal possible value over all possible states) values of characteristic function $\hat{w}(S)$ given by (10) are assigned to any coalition *S* different from *N*. Denote the core constructed with the values of characteristic function (11) as $D(\omega)$ and assume that it is non-empty for any state ω ,

$$D(\omega) = \{(\alpha_1(\omega), \dots, \alpha_n(\omega)) : \sum_{i \in S} \alpha_i(\omega) \ge w(\omega, S), \forall S \subset N, \sum_{i \in N} \alpha_i(\omega) = w(\omega, N)\}.$$
 (12)

Lemma 1. Let for any coalition $S \subset N$, $S \neq N$, the inequality $\hat{w}(S) < \min_{\omega \in O} v(\omega, N)$ hold. If condition

$$\sum_{i\in N} K_i^{\omega}(a^{\omega*}) = \max_{a^{\omega}\in\prod\limits_{i\in N} A_j^{\omega}} \sum_{i\in N} K_i^{\omega}(a^{\omega}),$$
(13)

is true, and the core $D(\omega)$ *is non-empty for any* ω *, and then* $D(\omega) \subset C(\omega)$ *.*

Proof. If there exists coalition $S \subset N$, $S \neq N$, such that $\hat{w}(S) \ge \min_{\omega \in \Omega} v(\omega, N)$, then the core $D(\omega)$ is empty. Assuming the non-emptiness of the core $D(\omega)$, we consider any imputation $\alpha(\omega) \in D(\omega)$. If condition (13) is true, it means that $\sum_{i=1}^{n} \alpha_i(\omega) = v(\omega, N) = w(\omega, N)$.

If condition (13) is true, it means that $\sum_{i \in N} \alpha_i(\omega) = v(\omega, N) = w(\omega, N)$. Subsequently, for any coalition $S \subset N$, we have $\sum_{i \in S} \alpha_i(\omega) \ge w(\omega, S) = \hat{w}(S) = \max_{\omega \in \Omega} v(\omega, S) \ge v(\omega, S)$, which proves that $\alpha(\omega) \in C(\omega)$. \Box

Remark 2. Condition (13) states that the maximal total payoff of the players in state ω coincides with their payoff if players adopt actions prescribed by the cooperative strategy profile. It may not be satisfied in general case in dynamic games. If condition (13) is not true, the main result of the paper can be proved, but it requires a modification in the method of characteristic function definition. We leave this case for future research.

Remark 3. We assume that the approximated core $D(\omega)$ is non-empty for any ω . The conditions under which it is non-empty are similar to the ones given in Remark 1, but in Equation (9) characteristic function $w(\omega, S)$ given by (11) is used. If the conditions of Lemma 1 are satisfied, then $D(\omega) \subset C(\omega)$, and non-emptiness of approximated core $D(\omega)$ implies non-emptiness of core $C(\omega)$.

Example 1. Consider three-player stochastic game with two states (ω_1 and ω_2). The sets of actions of player 1, 2, and 3 in state ω_1 (ω_2) are $\{a_1, a_2\}$, $\{b_1, b_2\}$ and $\{c_1, c_2\}$ ($\{\alpha_1, \alpha_2\}$, $\{\zeta_1, \zeta_2\}$, $\{\gamma_1, \gamma_2\}$), respectively. The payoff functions are given by the following matrices:

• *in state* ω_1 :

 $c_1: \qquad \begin{array}{c} b_1 & b_2 & & b_1 & b_2 \\ a_1 \begin{pmatrix} (10,10,8) & (0,15,0) \\ (15,0,0) & (5,5,5) \end{pmatrix} & c_2: & \begin{array}{c} a_1 & \begin{pmatrix} (0,0,15) & (2,4,4) \\ (4,4,2) & (0,0,0) \end{pmatrix} \end{array}$

• *in state* ω_2 :

$$\gamma_1: \qquad \begin{array}{ccc} \zeta_1 & \zeta_2 & & \zeta_1 & \zeta_2 \\ \alpha_1 & \begin{pmatrix} (2,1,1) & (4,0,2) \\ (0,4,2) & (7,5,3) \end{pmatrix} & \gamma_2: & \begin{array}{ccc} \alpha_1 & \begin{pmatrix} (2,3,0) & (4,2,4) \\ (3,4,3) & (7,5,7) \end{pmatrix} \end{array}$$

Player 1 chooses a row, player 2 chooses a column and player 3 chooses a matrix.

The transition probabilities are written in the matrices:

• for state ω_1 :

$$c_1: \qquad \begin{array}{c} b_1 & b_2 & & b_1 & b_2 \\ a_1 & \begin{pmatrix} (0.5, 0.5) & (0, 1) \\ (0, 1) & (0, 1) \end{pmatrix} & c_2: & a_1 & \begin{pmatrix} (0, 1) & (0.5, 0.5) \\ (0.5, 0.5) & (1, 0) \end{pmatrix} \end{array}$$

• for state ω_2 :

$$\gamma_1: \qquad \begin{array}{ccc} \zeta_1 & \zeta_2 & & \zeta_1 & \zeta_2 \\ \alpha_1 & \begin{pmatrix} (0,1) & (1,0) \\ (1,0) & (1,0) \end{pmatrix} & \gamma_2: & \alpha_1 & \begin{pmatrix} (0.2,0.8) & (0,1) \\ (0,1) & (1,0) \end{pmatrix} \end{array}$$

The first (second) element in any entry of the matrix is the probability of transition from the particular state and action profile to state ω_1 (state ω_2). One can easily notice that the probabilistic transitions are defined in state ω_1 when players choose action profiles (a_1, b_1, c_1) , (a_2, b_1, c_2) and (a_1, b_2, c_2) , and in state ω_2 when players choose action profiles $(\alpha_1, \zeta_1, \gamma_2)$. All other transitions are deterministic.

The discount factor equals 0.9. Cooperative strategy profile $\eta^* = (\eta_1^*, \eta_2^*, \eta_3^*)$ *is such that*

$$\eta_1^* = (a_1, \alpha_2), \quad \eta_2^* = (b_1, \zeta_2), \quad \eta_2^* = (c_1, \gamma_2),$$
(14)

which prescribes any player to choose the first action in state ω_1 and the second action in state ω_2 . The cooperative strategy profile defines a Markov chain with the structure that is depicted in Figure 1.



Figure 1. The transition probabilities defined by cooperative strategy profile η^* .

The players' payoffs are (10, 10, 8) in state ω_1 and (7, 5, 7) in state ω_2 . We obtain that the maximal total payoff of the players in state games coincide with the payoff that players get in states implementing cooperative strategy profile η^* . However, Theorem 1 is also true for the case when this condition is not satisfied.

First, we calculate the characteristic function $v(\omega, S)$ *by Equation (7) and its approximation* $w(\omega, S)$ *by (11) for state games. The values of these functions are represented in Table 1.*

Table 1. Values of characteristic function v and approximated characteristic function w for states ω_1 and ω_2 .

S	{1,2,3}	{1,2}	{1,3}	{2,3}	{1}	{2}	{3}
$v(\omega_1, S)$	28	8	10	10	0	0	0
$v(\omega_2, S)$	19	12	6	8	2	1	1
$w(\omega_1, S)$	28	12	10	10	2	1	1
$w(\omega_2, S)$	19	12	10	10	2	1	1

The cores of state games $C(\omega)$ and $D(\omega)$ calculated with values of functions $v(\omega, S)$ and $w(\omega, S)$ by Formulae (8) and (12) are non-empty for any ω and represented on Figures 2 and 3 for ω_1 and ω_2 respectively.



Figure 2. The core $C(\omega_1)$ (gray region) and approximated core $D(\omega_1)$ (blue region inside gray region) for ω_1 state game.



Figure 3. The core $C(\omega_2)$ (gray region) and approximated core $D(\omega_2)$ (blue region inside gray region) for ω_2 state game.

2.3. New Approximated Characteristic Function for Stochastic Games

We propose a new method of determining characteristic function for stochastic games based on the values of approximated characteristic function defined in states and given by Formula (11).

We assume that coalition *S* at any state of the game may obtain $\hat{w}(S)$ as maximum. Accordingly, this value is the maximal value that the coalition can get, regardless of the state that currently appears. If we summarize this value over infinite horizon with discount factor δ , we can calculate the approximation or the upper bound of the payoff that coalition *S* can get in stochastic subgame starting from state ω , which is

$$\bar{w}(\omega,S) = \begin{cases} \hat{w}(S) + \delta \hat{w}(S) + \ldots = \frac{1}{1-\delta} \hat{w}(S), & \text{if } S \subset N, S \neq N, \\ \sum_{i \in N} E_i^{\omega}(\eta^*), & \text{if } S = N. \end{cases}$$
(15)

Mathematics 2020, 8, 1135

One should notice that, according to Equation (15), we save the value of characteristic function for grand coalition without approximation. The reason is that, when we define the allocation of a joint payoff, the players should redistribute the value that they obtain using cooperative strategy profile, but not the approximated one. The cooperative stochastic subgame is defined by the set of players N and function (15). In the following, we omit the set of players and refer the cooperative stochastic subgame as $\bar{w}(\omega, S)$ given by (15).

Let $\overline{D}(\omega)$ be the core calculated with the values of function (15), i.e.,

$$\bar{D}(\omega) = \left\{ (\alpha_1(\omega), \dots, \alpha_n(\omega)) : \sum_{i \in S} \alpha_i(\omega) \ge \bar{w}(\omega, S), \forall S \subset N, \sum_{i \in N} \alpha_i(\omega) = \bar{w}(\omega, N) \right\}.$$
(16)

Let $\overline{D}(\omega)$ be non-empty for any $\omega \in \Omega$. We can compare the core $\overline{D}(\omega)$ constructed with the values of approximated function (15) and the core defined with the values of characteristic function defined with the classical approach. For any subgame G^{ω} , we define characteristic function using the maxmin approach:

$$\bar{v}(\omega,S) = \max_{\eta_S \in \prod_{j \in S} H_j} \min_{\eta_{N \setminus S} \in \prod_{j \in N \setminus S} H_j} \sum_{i \in S} E_i^{\omega}(\eta_S, \eta_{N \setminus S}).$$
(17)

Let $\bar{C}(\omega)$ be a non-empty core of subgame G^{ω} constructed with the values of function (17).

Lemma 2. Let for any coalition $S \subset N$, $S \neq N$ the inequality

$$\hat{w}(S) < \min_{\omega \in \Omega} v(\omega, N) \tag{18}$$

hold, and $\overline{D}(\omega)$ is non-empty for any ω , then $\overline{D}(\omega) \subset \overline{C}(\omega)$.

Proof. If $\hat{w}(S) < \min_{\omega \in \Omega} v(\omega, N)$ is not satisfied, then the core $\bar{D}(\omega)$ is empty by construction. Consider any imputation $\bar{\alpha}(\omega) \in \bar{D}(\omega)$ and prove that it belongs to the set $\bar{C}(\omega)$.

First, $\sum_{i \in N} \bar{\alpha}_i(\omega) = \bar{w}(\omega, N) = \sum_{i \in N} E_i^{\omega}(\eta^*) = \bar{v}(\omega, N).$ Second, we prove that $\sum_{i \in S} \bar{\alpha}_i(\omega) \ge \bar{v}(\omega, S)$ taking into account that $\sum_{i \in S} \bar{\alpha}_i(\omega) \ge \bar{w}(\omega, S)$ for any $S \neq N$. We prove that $\bar{w}(\omega, S) \ge \bar{v}(\omega, S)$.

By definition, we have

$$\bar{v}(\omega, S) = \max_{\eta_S} \min_{\eta_{N\setminus S}} \sum_{i\in S} E_i^{\omega}(\eta_S, \eta_{N\setminus S}),$$

and we write the functional equation for the right-hand side of this equality and obtain the following

$$\max_{\eta_{S}} \min_{\eta_{N\setminus S}} \sum_{i\in S} E_{i}^{\omega}(\eta_{S}, \eta_{N\setminus S}) = \max_{\eta_{S}} \min_{\eta_{N\setminus S}} \Big\{ \sum_{i\in S} K_{i}^{\omega}(a_{S}^{\omega}, a_{N\setminus S}^{\omega}) + \delta p(\omega, a^{\omega}) \sum_{i\in S} E_{i}(\eta_{S}, \eta_{N\setminus S}) \Big\},\$$

where $p(\omega, a^{\omega})$ is a vector $(p(\omega'|\omega, a^{\omega}) : \omega' \in \Omega)$.

Let profile (η_S , $\eta_{N \setminus S}$) be such that maxmin is reached at this profile, we can write the functional equation, as follows:

$$\sum_{i\in S} E_i^{\omega}(\eta_S, \eta_{N\setminus S}) = (\mathbb{I}_k - \delta \Pi(\eta_S, \eta_{N\setminus S}))^{-1} \sum_{i\in S} K_i(a_S, a_{N\setminus S}) \leqslant \frac{1}{1-\delta} \max_{\omega\in\Omega} \max_{a_S} \min_{a_{N\setminus S}} \sum_{i\in S} K_i^{\omega}(a_S^{\omega}, a_{N\setminus S}^{\omega})$$
$$= \frac{1}{1-\delta} \max_{\omega\in\Omega} v(\omega, S) = \bar{w}(\omega, S).$$

In the last inequality, we use the property of stochastic matrices, i.e., the sum of the elements in any row of matrix $(\mathbb{I}_k - \delta \Pi(\eta_S, \eta_{N \setminus S}))^{-1}$ equal $1/(1-\delta)$, because $\Pi(\eta_S, \eta_{N \setminus S})$ is a stochastic matrix. The lemma is proved. \Box

Remark 4. We assume non-emptiness of the approximated core $\overline{D}(\omega)$ in stochastic game with any initial state ω . If condition (18) in Lemma 2 is satisfied, the non-emptiness of the approximated cores $D(\omega)$ for any ω implies the non-emptiness of approximated core $\overline{D}(\omega)$. It follows from definition of characteristic function $\overline{w}(\omega, S)$ and formula (10). Moreover, the non-emptiness of approximated core $\overline{D}(\omega)$ implies non-emptiness of core $\overline{C}(\omega)$.

Example 2. (continuation of Example 1) We continue calculations for stochastic game described in Example 1. Define characteristic function \bar{v} by (17) and approximated characteristic function \bar{w} by (15). The values of these functions are given in Table 2.

Table 2. Values of characteristic function \bar{v} and approximated characteristic function \bar{w} for stochastic game starting from states ω_1 and ω_2 .

S	{1,2,3}	{1,2}	{1,3}	{2,3}	{1}	{2}	{3}
$\bar{v}(\omega_1, S)$	252.07	92.41	62.01	64.00	9.47	9.00	9.00
$\bar{v}(\omega_2,S)$	245.86	95.17	77.87	60.00	10.52	10.00	10.00
$\bar{w}(\omega_1, S)$	252.07	120.00	100.00	100.00	20.00	10.00	10.00
$\bar{w}(\omega_2, S)$	245.86	120.00	100.00	100.00	20.00	10.00	10.00

The cores $\bar{C}(\omega)$ and $\bar{D}(\omega)$ constructed with the values of functions \bar{v} and \bar{w} , respectively, are non-empty and depicted on Figures 4 and 5 for initial states ω_1 and ω_2 , respectively. One can notice that $\bar{D}(\omega) \subset \bar{C}(\omega)$ for any ω .



Figure 4. The core $\bar{C}(\omega)$ (gray region) and approximated core $\bar{D}(\omega)$ (blue region inside gray region) in stochastic game with ω_1 initial state.

The approximated core $\overline{D}(\omega_1)$ *is defined as the set*

$$\bar{D}(\omega_1) = \left\{ (\bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3) : \bar{\alpha}_1 + \bar{\alpha}_2 + \bar{\alpha}_3 = 252.07, \bar{\alpha}_1 + \bar{\alpha}_2 \ge 120.00, \bar{\alpha}_1 + \bar{\alpha}_3 \ge 100.00, \\ \bar{\alpha}_2 + \bar{\alpha}_3 \ge 100.00, \bar{\alpha}_1 \ge 20.00, \bar{\alpha}_2 \ge 10.00, \bar{\alpha}_3 \ge 10.00 \right\}.$$

The approximated core $\overline{D}(\omega_2)$ *is defined as the set*

$$\bar{D}(\omega_2) = \left\{ (\bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3) : \bar{\alpha}_1 + \bar{\alpha}_2 + \bar{\alpha}_3 = 245.86, \bar{\alpha}_1 + \bar{\alpha}_2 \ge 120.00, \bar{\alpha}_1 + \bar{\alpha}_3 \ge 100.00, \bar{\alpha}_2 + \bar{\alpha}_3 \ge 100.00, \bar{\alpha}_1 \ge 20.00, \bar{\alpha}_2 \ge 10.00, \bar{\alpha}_3 \ge 10.00 \right\}.$$



Figure 5. The core $\overline{C}(\omega_2)$ (gray region) and approximated core $\overline{D}(\omega_2)$ (blue region inside gray region) in stochastic game with ω_2 initial state.

3. Strongly Subgame-Consistent Core in Stochastic Games

3.1. Imputation Distribution Procedure

In cooperation, players follow the cooperative strategy profile η^* and then agree on the core as a cooperative solution of the game or the set of possible imputations of the joint payoff in the game. We assume that the core for any subgame G^{ω} is calculated based on function (15), which is $\overline{D}(\omega)$. Consider an imputation $\bar{\alpha}(\omega) \in \bar{D}(\omega)$. Obviously, if the players are paid step by step according to initially given payoff functions K_i^{ω} , $i \in N$, we cannot guarantee that they will get the components of imputation $\bar{\alpha}(\omega)$ as an expected payoff in subgame G^{ω} . Therefore, we define the scheme of state payments that, in total, will give the players to obtain the components of imputation $\bar{\alpha}(\omega)$.

Definition 1. [10,11] We call the collection of vectors $(\beta_i : i \in N)$, where $\beta_i = (\beta_i(\omega_1), \dots, \beta_i(\omega_k)), \beta_i(\omega)$ is a payment to player i in state ω in cooperative stochastic game, an imputation distribution procedure (IDP) of imputation $\bar{\alpha}(\omega) \in \bar{D}(\omega)$ if

- 1. $\sum_{\substack{i \in N \\ \bar{\alpha}_i(\omega)}} \beta_i(\omega) = \sum_{\substack{i \in N \\ i}} K_i^{\omega}(a^{\omega*}) \text{ for any } \omega \in \Omega;$ 2. $\overline{\alpha}_i(\omega) = B_i^{\omega}, \text{ where } B_i^{\omega} \text{ is the expected discounted sum of payments to player } i \text{ in stochastic subgame}$ starting from state ω , according to procedure β .

The expected sum of payments to player *i* made according to IDP can be calculated by formula (see [14]):

$$B_i^{\omega} = \pi_0 (\mathbb{I} - \delta \Pi(\eta^*))^{-1} \beta_i,$$

where π_0 is such that $\pi_0^{\omega} = 1$ and $\pi_0^{\omega'} = 0$ for any $\omega' \neq \omega$.

Remark 5. The IDP determined in Definition 1 for an imputation $\bar{\alpha}(\omega) \in \bar{D}(\omega)$ may be non-unique.

In the following section, we describe a property of the imputations from the core and corresponding IDP, which allows to narrow the set of IDP.

3.2. Strongly Subgame-Consistent Core

We formulate the property of strongly subgame consistency of the core and propose sufficient conditions of strongly subgame consistency of the core in stochastic games with characteristic function (15). We suppose that the cores of stochastic game G and any subgame G^{ω} , $\omega \in \Omega$, are non-empty.

In cooperation, players agree on the joint implementation of cooperative strategy profile η^* and expect to obtain the components of the imputation belonging to the core $\bar{D}(\omega)$ in the subgame stating from ω . Reaching an intermediate state $\omega \in \Omega$, Player *i* chooses action $a_i^{\omega*}$ prescribed by cooperative strategy profile η^* and gets payoff $K_i^{\omega}(a^{\omega*})$. If the players recalculate the solution in the current subgame and find solution of cooperative subgame G^{ω} , we would assume that the cooperative solution is chosen from the core $\bar{D}(\omega)$. It would be reasonable to require that the payoff received by a player in state ω summarized with the expected sum of any imputations from the cores $\bar{D}(\omega')$, $\omega' \in \Omega$, following state ω , would be an imputation from the core $\bar{D}(\omega)$. If this property holds for any intermediate state $\omega \in \Omega$, then the core of cooperative stochastic game with characteristic function (15) is strongly subgame-consistent.

To determine a strongly subgame-consistent core, we need to define the so-called expected core at state ω , i.e., we define the set of expected imputations belonging to the cores, which are cooperative solutions of the following subgames. We determine the expected core of state $\omega \in \Omega$, as follows:

$$E\bar{D}(\omega) = \Big\{ \delta \sum_{\omega \in \Omega} p(\omega' | \omega, a^{\omega*}) \bar{\alpha}(\omega'), \quad \bar{\alpha}(\omega') \in \bar{D}(\omega') \Big\}.$$

Definition 2. We call the core $\overline{D}(\omega)$ strongly subgame consistent solution of cooperative stochastic game with approximated characteristic function $\overline{w}(\omega, S)$ starting from state ω if for any imputation $\overline{\alpha}(\omega) \in \overline{D}(\omega)$ there exists an IDP $\beta = (\beta_i : i \in N)$, where $\beta_i = (\beta_i(\omega) : \omega \in \Omega)$, satisfying condition:

$$\beta \oplus E\bar{D} \subset \bar{D},\tag{19}$$

where $E\overline{D}$ is the vector $(E\overline{D}(\omega_1), \dots, E\overline{D}(\omega_k))'$ of expected cores for states $\omega_1, \dots, \omega_k$ respectively, \overline{D} is a vector with elements which are sets, i.e., $\overline{D} = (\overline{D}(\omega_1), \dots, \overline{D}(\omega_k))'$.

Remark 6. The inclusion (19) is written in a vector form. To explain it, we write the first row of vector inclusion (19):

$$\beta(\omega_1) \oplus E\bar{D}(\omega_1) \subset \bar{D}(\omega_1)$$

where $\beta(\omega_1) \in \mathbb{R}^n$, $E\overline{D}(\omega_1) \subset \mathbb{R}^n$, $\overline{D}(\omega_1) \subset \mathbb{R}^n$. The operation $a \oplus C$, where $a \in \mathbb{R}^n$ and C is a set in \mathbb{R}^n , is defined as the set $\{a + c, \text{ for all } c \in C\}$.

Theorem 1. The core $\overline{D}(\omega)$, if it exists, is strongly subgame-consistent.

Proof. Following Definition 2 we need to prove that there exists an IDP of the elements from the core $\overline{D}(\omega)$ defined in (16) satisfying two properties from Definition 1, such that inclusion (19) is true.

Let for any imputation $\bar{\alpha}_i(\omega) \in \bar{D}(\omega)$, the IDP is calculated as

$$\beta_i = (\mathbb{I}_k - \delta \Pi(\eta^*)) \bar{\alpha}_i, \tag{20}$$

where $\beta_i = (\beta_i(\omega_1), \dots, \beta_i(\omega_k))'$ and $\bar{\alpha}_i = (\bar{\alpha}_i(\omega_1), \dots, \bar{\alpha}_i(\omega_k))'$.

First, we prove that β , defined in (20), satisfies properties 1 and 2 in Definition 1.

1. Find the sum of β_i over the set of players, we obtain

$$\sum_{i\in\mathbb{N}}\beta_i = (\mathbb{I}_k - \delta\Pi(\eta^*))\sum_{i\in\mathbb{N}}\bar{\alpha}_i = (\mathbb{I}_k - \delta\Pi(\eta^*))(\bar{v}(\omega_1, N), \dots, \bar{v}(\omega_k, N))'$$
$$= (\mathbb{I}_k - \delta\Pi(\eta^*))(\mathbb{I}_k - \delta\Pi(\eta^*))^{-1}\sum_{i\in\mathbb{N}}K_i(a^*) = \sum_{i\in\mathbb{N}}K_i(a^*),$$

or for any $\omega \in \Omega$ the equality $\sum_{i \in N} \beta_i(\omega) = \sum_{i \in N} K_i^{\omega}(a^{\omega*})$ is true.

2. We prove that $\bar{\alpha}_i(\omega) = B_i^{\omega}$ or in vector form $\bar{\alpha}_i = B_i$, where $B_i = (B_i(\omega_1), \dots, B_i(\omega_k))'$. We have

$$B_i = (\mathbb{I}_k - \delta \Pi(\eta^*))^{-1} \beta_i = (\mathbb{I}_k - \delta \Pi(\eta^*))^{-1} (\mathbb{I}_k - \delta \Pi(\eta^*)) \bar{\alpha}_i = \bar{\alpha}_i$$

Therefore, the payment vector β_i , $i \in N$, is the distribution procedure of imputation $\bar{\alpha}_i$.

Now, we prove that inclusion (19) holds. Let β_i be given by Equation (20), there $\bar{\alpha}_i = (\bar{\alpha}_i(\omega_1), \dots, \bar{\alpha}_i(\omega_k))'$ and $\bar{\alpha}_i(\omega_j) \in \bar{D}(\omega_j)$ for any $j = 1, \dots, k$. Consider the sum $\beta(\omega) + \varepsilon(\omega)$, where $\varepsilon(\omega)$ is any vector from the expected core $E\bar{D}(\omega)$. Substituting expressions of β_i from Equation (20) and element of the expected core into the sum, we get

$$\beta + \varepsilon = (\mathbb{I}_k - \delta \Pi(\eta^*))\bar{\alpha} + \delta \Pi(\eta^*)\bar{\alpha} = \bar{\alpha} \in \bar{D},$$

which proves the theorem. \Box

Theorem 1 gives the method of construction of payment scheme of any element from the core \bar{D} defined by (16) while using values of function (15).

Example 3. (continuation of Example 1 and 2) We demonstrate how to define IDP using a method from the proof of Theorem 1. Let for ω_1 and ω_2 the core imputations $\bar{\alpha}(\omega_1) = (100.00, 100.00, 52.07) \in \bar{D}(\omega_1)$ and $\bar{\alpha}(\omega_2) = (50.00, 95.86, 100.00) \in \bar{D}(\omega_2)$ be chosen. To calculate IDP by Formula (20), we need to define matrix $\Pi(\eta^*)$, which is

$$\Pi(\eta^*) = egin{pmatrix} 0.5 & 0.5 \ 0 & 1 \end{pmatrix}$$

for cooperative strategy profile η^* determined by (14).

Using formula (20) with $\bar{\alpha}_1 = (100.00, 50.00)$, $\bar{\alpha}_2 = (100.00, 95.86)$, $\bar{\alpha}_3 = (52.07, 100.00)$, we obtain

$$\begin{split} \beta_1 &= (32.50, -40.00), \\ \beta_2 &= (11.86, 5.86), \\ \beta_3 &= (-16.36, 53.14), \end{split}$$

where the first component of vector β_i is the payment to player *i* in state ω_1 and the second component is the payment in state ω_2 . We can easily check that collection of vectors ($\beta_i : i \in N$) satisfies conditions from Definition 1 of IDP.

The approximated cores $\tilde{D}(\omega_1)$ and $\tilde{D}(\omega_2)$ are strongly subgame-consistent, which is proved in Theorem 1.

Remark 7. The new method of construction of the characteristic function or the so-called approximated characteristic function proposed in the paper allows not only to find the strongly subgame-consistent subset of the core, but also simplifies calculations. In the example, each player has two actions in any state. Therefore, he has four pure stationary strategies in a stochastic game, and there are 64 strategy profiles in the game. The calculations of maxmin payoff of a coalition in such games is a complicated computational problem. The new approach allows for avoiding these calculations using the values of approximated characteristic function defined in state games to determine the function for a stochastic game.

4. Conclusions

We have proposed a new method of constructing the characteristic function in stochastic games. The method simplifies calculations in comparison with the previously introduced approaches. An additional advantage of the method is that the core calculated with the values of this characteristic function satisfies strongly subgame consistency. This property positively characterizes the realization of the imputations from the core in a dynamic game process. The property of strongly subgame consistency is applied for set-valued cooperative solutions, like the core. We can briefly

characterize the possible directions for future research in this area. We can also consider additional simplifications in characteristic function definitions, which allow not only to keep the strong subgame consistency properties of the core, but also to reduce the number of calculations defining cooperative stochastic game.

Author Contributions: Conceptualization, E.P. and L.P.; methodology, E.P. and L.P.; software, E.P. and L.P.; validation, E.P. and L.P.; formal analysis, E.P. and L.P.; investigation, E.P. and L.P.; resources, E.P. and L.P.; data curation, E.P. and L.P.; writing—original draft preparation, E.P. and L.P.; writing—review and editing, E.P. and L.P.; visualization, E.P. and L.P.; supervision, E.P. and L.P.; project administration, E.P. and L.P.; funding acquisition, E.P. and L.P. All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported by Russian Science Foundation, grant no. 17-11-01079.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Von Neumann, J.; Morgenstern, O. *Theory of Games and Economic Behavior*; Princeton University Press: Princeton, NJ, USA, 1944.
- 2. Aumann, R.J.; Peleg, B. Von Neumann-Morgenstern Solutions to Cooperative Games without Side Payments. *Bull. Am. Math.* **1960**, *66*, 173–179. [CrossRef]
- 3. Chander, P.; Tulkens, H. The core of an economy with multilateral environmental externalities. *Int. J. Game Theory* **1997**, *26*, 379–401. [CrossRef]
- 4. Chander, P. The gamma-core and coalition formation. Int. J. Game Theory 2007, 35, 539–556. [CrossRef]
- 5. Petrosyan, L.; Zaccour, G. Time-consistent Shapley value allocation of pollution cost reduction. *J. Econ. Dyn. Control.* **2003**, *27*, 381–398. [CrossRef]
- 6. Zaccour, G. Computation of Characteristic function values for linear-state differential games. *J. Optim. Theory Appl.* **2003**, *117*, 183–194. [CrossRef]
- 7. Reddy, P.V.; Zaccour, G. A friendly computable characteristic function. *Math. Soc. Sci.* **2016**, *82*, 18–25. [CrossRef]
- 8. Petrosyan, L.A.; Gromova, E.V. On an approach to constructing a characteristic function in cooperative differential games. *Autom. Remote Control.* **2017**, *78*, 1680–1692.
- 9. Pankratova, Y.B.; Petrosyan, L.A. New Characteristic Function for Multistage Dynamic Games. 2018. Available online: https://cyberleninka.ru/article/n/new-characteristic-function-for-multistage-dynamic-games (accessed on 10 June 2020).
- 10. Petrosjan, L.A. *Consistency of Solutions of Differential Games with Many Players;* Vestnik of Leningrad University, Series Mathematics, Mechanics, Astronomy: Saint Petersburg, Russia, 1977; pp. 46–52.
- 11. Petrosjan, L.A.; Danilov, N.A. *Time-Sonsitent Solutions of Non-Antagonistic Differential Games with Transferable Payoffs*; Vestnik Leningrad University: Saint Petersburg, Russia, 1979; pp. 52–59.
- 12. Petrosjan, L.A. Cooperative stochastic games. In *Advances in Dynamic Games. Annals of the International Society of Dynamic Games*; Haurie, A., Muto, S., Petrosjan, L.A., Raghavan, T.E.S., Eds.; Birkhauser: Boston, MA, USA, 2006; Volume 8; pp. 52–59.
- 13. Avrachenkov, K.; Cottatellucci, L.; Maggi, L. Cooperative Markov decision processes: Time consistency, greedy players satisfaction, and cooperation maintenance. *Int. J. Game Theory* **2013**, *42*, 39–262. [CrossRef]
- 14. Parilina, E.M. Stable cooperation in stochastic games. Autom. Remote. Control. 2015, 76, 1111–1122. [CrossRef]
- 15. Parilina, E.; Zaccour, G. Node-Consistent Core for Games Played over Event Trees. *Automatica* 2015, *53*, 304–311. [CrossRef]
- 16. Petrosjan, L.A. Construction of Strongly Time Consistent Solutions in Cooperative Differential Games; Birkhäuser: Cham, Switzerland, 1992; pp. 33–38.
- 17. Petrosyan, L. Strong Strategic Support of Cooperation in Multistage Games. *Int. Game Theory Rev.* **2019**, 21, 1940004. [CrossRef]
- Parilina, E.M.; Petrosyan, L.A. Strongly Subgame-Consistent Core in Stochastic Games. *Autom. Remote Control.* 2018, 79, 1515–1527. [CrossRef]
- 19. Petrosian, O.; Gromova, E.; Pogozhev, S. Strong Time-Consistent Subset of the Core in Cooperative Differential Games with Finite Time Horizon. *Autom. Remote Control.* **2018**, *79*, 1912–1928. [CrossRef]

- 20. Sedakov, A.A. On the Strong Time Consistency of the Core. *Autom. Remote Control.* **2018**, *79*, 757–767. [CrossRef]
- 21. Parilina, E.; Tampieri, A. Stability and Cooperative Solution in Stochastic Games. *Theory Decis.* **2018**, *84*, 601–625. [CrossRef]
- 22. Shapley, L.S. Stochastic Games. *Proc. Natl. Acad. Sci. USA* **1953**, *39*, 1095–1100. Available online: https://www.pnas.org/content/39/10/1095.short (accessed on 28 May 2020). [CrossRef] [PubMed]
- 23. Bondareva, O.N. Some applications of linear programming methods to the theory of cooperative games. *Problemy Kybernetiki* **1963**, *10*, 119–139.
- 24. Shapley, L.S. On balanced sets and cores. Nav. Res. Logist. Q. 1967, 14, 453–460. [CrossRef]
- 25. Peleg, B.; Sudhölter, P. Introduction to the Theory of Cooperative Games; Springer: Berlin/Heidelberg, Germany, 2007.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).