Contributions to Game Theory and Management, XI, 129–195

A Survey on Cooperative Stochastic Games with Finite and Infinite Duration*

Elena Parilina

St. Petersburg State University,
7/9 Universitetskaya nab., Saint Petersburg 199034, Russia
E-mail: e.parilina@spbu.ru

Abstract The paper is a survey on cooperative stochastic games with finite and infinite duration which based on the author's and coauthors' publications. We assume that the non-cooperative stochastic game is initially defined. The cooperative version of the game is constructed, the cooperative solutions are found. The properties of cooperative solutions of the game which are realised in dynamics are considered. Several numerical examples of stochastic games illustrate theoretical results.

Keywords: cooperative stochastic game, cooperative solution, imputation distribution procedure, subgame consistency.

1. Introduction

The paper is an overview of the results obtained in the theory of cooperative stochastic games by the author and her coauthors (Baranova, 2006, Parilina, 2014, Parilina, 2015, Parilina, 2016, Baranova and Petrosjan, 2006, Parilina and Petrosyan, 2017, Parilina and Tampieri, 2018, Petrosyan et al., 2004, Petrosjan and Baranova, 2005, Petrosyan and Baranova, 2003, Petrosyan and Baranova, 2005a, Petrosyan and Baranova, 2005b).

The starting point of stochastic game theory is a publication of L. Shapley (Shapley, 1953a), in which the existence of value of a zero-sum stochastic game with a finite set of players' strategies is proved. A generalization of this result for the case of n-person stochastic game was obtained in the papers (Fink, 1964) and (Takahashi, 1964), in which it was proved that equilibrium exists in stationary strategies in a stochastic game with a compact set of strategies and a finite set of states. Many papers are devoted to the proof of the existence of the Nash equilibrium in various classes of strategies, studying stochastic games with incomplete information, asymmetric players, stochastic games of a special structures (see the following publications: (Solan and Vieille, 2002, Vieille, 2000, Mertens and Neyman, 1981a, Mertens and Neyman, 1981b, Neyman, 2008, Neyman, 2013, Nowak, 1985, Nowak, 1999, Nowak and Radzik, Hörner et al., 2010, Solan, 1998, Jáskiewicz and Nowak, 2016, Neyman and Sorin, 2003, Solan, 2009, Solan and Vieille, 2015)).

The method of constructing a cooperative version of stochastic game realized on a finite tree was first proposed by L. A. Petrosyan in the paper (Petrosjan, 2006), where the problem of time consistency of the Shapley value was formulated and a method of regularization of time-inconsistent Shapley value is introduced. Then the method of constructing a cooperative version of stochastic game with infinite duration was proposed in the paper (Baranova and Petrosjan, 2006). Cooperative stochastic games of infinite duration with a finite set of strategies were later studied

* The work was supported by Russian Science Foundation, project no. 17-11-01079.

in (Kohlberg and Neyman, 2015, Parilina, 2015, Parilina and Tampieri, 2018). The principles of stable cooperation are formulated for dynamic and differential games in (Petrosyan and Zenkevich, 2015). The first principle is time consistency (or subgame consistency) of cooperative solutions which was initially proposed by L. A. Petrosyan (Petrosyan, 1977) for differential games.

The mechanism for determining payments to the players for regularization of time-inconsistent cooperative solutions using the so-called imputation distribution procedure was introduced by L. A. Petrosyan and V. V. Danilov (Petrosyan and Danilov, 1979). Further, the problem of constructing time-consistent cooperative solutions was studied in the paper (Petrosyan and Shevkoplyas, 2000) for differential games with random duration, and in (Yeung and Petrosyan, 2011) for dynamic games with random duration.

The second principle of stable cooperation in dynamic and differential games is strategic consistency of a cooperative solution which was initially proposed in (Petrosyan, 1998). This principle is relevant and can be adapted for various classes of differential and dynamic games (Shevkoplyas, 2010, Petrosjan and Grauer, 2002, Petrosyan and Chistyakov, 2013, Petrosyan and Sedakov, 2015).

The third principle of stable cooperation is irrational-behavior-proof which was formulated by D. W. K. Yeung (Yeung, 2006) and then was applied for linear-quadratic games (Tur, 2014, Markovkin, 2006). The conditions for stable cooperation with Markov processes, which allow players' cooperation, including irrational-behavior-proof condition, are formulated in (Avrachenkov et al., 2013).

Time consistency condition was extended for the case when the cooperative solution is a set (containing more than one imputation) in (Petrosyan, 1993) and was called strongly time consistency. Recently, this condition is investigated in various classes of games (Gromova and Petrosyan, 2015, Sedakov, 2015, Chistyakov and Petrosyan, 2011, Parilina and Pet ro syan, 2017).

The paper is organized as follows. Section 2 contains results on cooperative stochastic games with finite duration while Section 3 is devoted to cooperative stochastic games with infinite duration. We briefly conclude in Section 4.

2. Cooperative stochastic games with finite duration

2.1. Non-cooperative stochastic games

We define a finite stochastic game played on a graph. Let $\Psi = (Z, L)$ be a finite graph of a tree structure, where Z is the set of vertices of the graph, and $L : Z \longrightarrow Z$ is a point-set mapping defined on the set Z, with values in the set of the subsets of set Z. The vertex z_0 is the initial vertex of the tree graph Ψ. We denote the terminal vertices of graph Ψ by $Z^T \subset Z$, that is, the vertices z for which $L(z) = \varnothing$. The finite tree graph with initial vertex z_0 is denoted by $\Psi(z_0)$.

Let at each vertex $z \in Z$ of the graph $\Psi(z_0)$ the normal form game of n players

$$\Gamma(z) = \langle N, A_1^z, \dots, A_n^z, K_1^z, \dots, K_n^z \rangle ,$$

be given, and $N = \{1, 2, \dots, n\}$ is a finite set of players, the same for all vertices $z \in Z$; A_i^z is a finite set of actions of player $i \in N$, $K_i^z(a_1^z, \dots, a_n^z) : \prod_{j \in N} A_j^z \to \mathbb{R}$ is a payoff function of player i, $a_i^z \in A_i^z$. The collection of actions $a^z = (a_1^z, \dots, a_n^z)$, $a_i^z \in A_i^z$, $i \in N$, is called an action profile in the game $\Gamma(z)$. And $a^z \in A^z = \prod_{i \in N} A_i^z$, A^z is the set of action profiles in game $\Gamma(z)$.

For each vertex $z \in Z$ we define the transition probabilities to the vertices $y \in L(z)$ of the graph $\Psi(z_0)$ following the vertex z. These probabilities depend on the action profile a^z realized in the game $\Gamma(z)$. Thus, for each vertex $z \in Z$ we define a function $p(\cdot|z, a^z) : A^z \to \Delta(L(z))$, where $\Delta(L(z))$ is a probability distribution over the set $L(z)$:

$$p(y|z, a^z) \geqslant 0,$$
$$\sum_{y \in L(z)} p(y|z, a^z) = 1$$

for any action profile $a^z \in A^z$. The value $p(y|z, a^z)$ is the probability that at the next stage, the game $\Gamma(y)$ will be played, $y \in L(z)$, if at the previous stage in the game $\Gamma(z)$, the action profile $a^z = (a_1^z, \ldots, a_n^z)$ has been realized.

We also suppose that the duration of the game is random which values are 0, 1, ..., l, and l is the length of the game (by the length of the game we mean the number of stages in the game of maximal possible path). Define probabilities q_k of the event that the game will end at stage k. Notice that $0 \leqslant q_k \leqslant 1$, $k = 0, \ldots, l-1, q_l = 1$, where l is the length of the game (by the length of the game we mean the number of stages in the game of maximal possible path); stage k at vertex $z \in Z$ in a stochastic game with random duration is determined from the condition: $z \in (L(z_0))^k$.

Remark 1. Notice that the probabilities q_k, $k = 0, \ldots, l$ are conditional probabilities and do not form probability distribution of the game duration. In case when all paths in graph $\Psi(z_0)$ have the same length l, the discrete distribution of a random variable equal to the game duration, determined by the conditional probabilities q_k, is presented in Table 1, in which P_k is the probability that the game will end at stage k.

Table 1. Probability distribution of the game duration.

k	P_k
0	q_0
1	$(1-q_0)q_1$
2	$(1-q_0)(1-q_1)q_2$
$\vdots$	$\vdots$
l	$(1-q_0)(1-q_1)\cdot \ldots \cdot(1-q_{l-1})$

Definition 1. Stochastic game with random duration $G(z_0)$, where z_0 is an initial vertex of a tree graph $\Psi(z_0)$, is a set

$$G(z_0) = \left\langle N, \Psi(z_0), \{\Gamma(z)\}_{z \in Z}, \{q_k\}_{k=0}^{l}, \{p(\cdot|z, a^z)\}_{z \in Z, a^z \in A^z} \right\rangle. \tag{1}$$

From the definition of a stochastic game with a random duration it is clear that the transitions from some vertices of the graph $\Psi(z_0)$ to the others, as well as the final stage of the game are random.

Stochastic game with random duration $G(z_0)$ is played in the following way:

1. At vertex z_0 of the graph $\Psi(z_0)$, a simultaneous game $\Gamma(z_0)$ is played. Suppose that in this game action profile $a^{z_0} \in A^{z_0}$ is realized by the players. Each

player $i \in N$ receives a payoff $K_i^{z_0}(a^{z_0})$. The stochastic game $G(z_0)$ either terminates with probability q_0, $0 \leqslant q_0 \leqslant 1$, or continues with probability $1-q_0$ and transmits to the vertex $y \in L(z_0)$ of the graph $\Psi(z_0)$ with probability $p(y|z_0, a^{z_0})$, depending on the action profile a^{z_0} realized in the game $\Gamma(z_0)$. In case when the set $L(z_0)$ is empty, the game ends at the vertex z_0 with probability 1.

2. Suppose that at stage k the game process is at the vertex $z_k \in Z$, at which the game in a normal form $\Gamma(z_k)$ is given. Let the action profile $a^{z_k} \in A^{z_k}$ is realized in this game. Each player $i \in N$ receives a payoff $K_i^{z_k}(a^{z_k})$. Stochastic game either ends with probability q_k, $0 \leqslant q_k \leqslant 1$, or continues with probability $1-q_k$ and transits to the vertex $z_{k+1} \in L(z_k)$ with probability $p(z_{k+1}|z_k, a^{z_k})$, which depending on the action profile a^{z_k} realized in game $\Gamma(z_k)$. In case when the set $L(z_k)$ is empty, the game terminates at the vertex z_k with probability 1.
3. The stochastic game continues until the terminal vertex is reached or it may end according to the realizations of probabilities $q_0, \ldots, q_l$.

We denote by $G(z_k)$ the subgame (see (Kuhn, 1950, Kuhn, 1953)) of the game $G(z_0)$ starting at the vertex $z_k \in Z$ of graph the $\Psi(z_0)$ (starting with the game $\Gamma(z_k)$), which is also a stochastic game with random duration. Subgame $G(z_k)$ is defined on the subgraph $\Psi(z_k)$ with the set of vertices $Z(z_k) \subset Z$ and is given by the quintuple

$$G(z_k) = \left\langle N, \Psi(z_k), \{\Gamma(z)\}_{z \in Z'}, \{q_s\}_{s=k}^{l}, \{p(\cdot|z, a^z)\}_{z \in Z(z_k), a^z \in A^z} \right\rangle .$$

To solve the game you need to determine the set of players' strategies. We denote by $\varphi_i : Z \to \prod_{z \in Z} \Delta(A_i^z)$ the behavior strategy of player i in game $G(z_0)$, where $\Delta(A_i^z)$ is the set of mixed actions of the player i at the vertex $z \in Z$. The strategy profile in stochastic game $G(z_0)$ is a collection of the players' strategies given by $\varphi = (\varphi_i : i \in N)$. Denote by Σ_i the set of behavior strategies of player i in the stochastic game $G(z_0)$, then $\Sigma = \prod_{i \in N} \Sigma_i$ is the set of behavior strategy profiles in game $G(z_0)$. Obviously, the restriction of the strategy φ_i on subgraph $\Psi(z_k)$ of graph $\Psi(z_0)$ is a strategy in subgame $G(z_k)$. Denote this restriction of a strategy by $\varphi_i^{z_k}$.

2.2. Main functional equations

Assume that in stochastic game $G(z_0)$ players implement strategies φ_i, $i \in N$. Define the payoff of the player i as mathematical expectation of his payoff relative to a random variable equal to the game duration, and e. g., for the realized path $z_1 \in L(z_0)$, $z_2 \in L(z_1)$, $\ldots$, $z_l \in L(z_{l-1})$, $L(z_l) = \varnothing$, we obtain

$$E_i(z_0) = \sum_{k=0}^{l} P_k \sum_{j=0}^{k} K_i^{z_j}\left(a^{z_j}\right) = \sum_{k=0}^{l} q_k \left(\prod_{j=0}^{k-1} (1-q_j) \right) \left(\sum_{m=0}^{k} K_i^{z_m}\left(a^{z_m}\right) \right),$$

where $a^{z_0} \in A^{z_0}$, $a^{z_1} \in A^{z_1}$, $\ldots$, $a^{z_l} \in A^{z_l}$ is a sequence of realized action profiles when players adopt strategies $(\varphi_i : i \in N)$.

Since transitions from the vertices to the following vertices are stochastic, we consider mathematical expectation of the player's payoff relative to random transitions from vertices to the following vertices as a player's payoff in the stochastic game. The mathematical expectation $E_i(z_0, \varphi)$ of player i's payoff in the game satisfies the functional equation

$$E_i(z_0,\varphi) = q_0 K_i^{z_0}(a^{z_0}) + (1-q_0)\left(K_i^{z_0}(a^{z_0}) + \sum_{y\in L(z_0)} p(y|z_0,a^{z_0})E_i(y,\varphi^y)\right) \quad (2)$$
$$= K_i^{z_0}(a^{z_0}) + (1-q_0)\sum_{y\in L(z_0)} p(y|z_0,a^{z_0})E_i(y,\varphi^y),$$

where $E_i(y,\varphi^y)$ is the mathematical expectation of player i's payoff in the subgame $G(y)$ starting at the vertex $y \in L(z_0)$ of graph $G(z_0)$.

Assume that $z \in (L(z_0))^k$, that is, the game process enters the vertex $z \in Z$ at stage k, then the mathematical expectation of player i's payoff in the subgame $G(z)$ satisfies the functional equation

$$E_i(z,\varphi^z) = q_k K_i^{z}(a^{z}) + (1-q_k)\left(K_i^{z}(a^{z}) + \sum_{y\in L(z)} p(y|z,a^{z})E_i(y,\varphi^y)\right)$$
$$= K_i^{z}(a^{z}) + (1-q_k)\sum_{y\in L(z)} p(y|z,x^{z})E_i(y,\varphi^y).$$

To define a cooperative version of the game, it is necessary to determine a cooperative path (one of the cooperative paths, if there are several ones), that is, the path that maximizes the total players' payoffs. In the case of stochastic games, this is a subtree with the given transition probabilities, at which the maximum of the mathematical expectation of the total players' payoffs in the whole game is achieved. However, the maximum mathematical expectation of the total players' payoffs in mixed strategies is equal to the maximum mathematical expectation of the summarized players' payoffs in pure strategies. Therefore, we can restrict ourselves and consider the class of pure strategies to find cooperative strategies in the stochastic game.

2.3. Cooperative stochastic games with finite duration

Denote by $\bar{\varphi} = (\bar{\varphi}_1,\ldots,\bar{\varphi}_n)$ the pure strategy profile in game $G(z_0)$ which maximizes the total mathematical expectations of the players' payoffs:

$$V(N,z_0) = \max_{\varphi\in\Sigma}\left[\sum_{i\in N} E_i(z_0,\varphi)\right] = \sum_{i\in N} E_i(z_0,\bar{\varphi}).$$

We call this strategy profile as a cooperative one. Let strategy profile $\bar{\varphi}$ be such that $\bar{\varphi}_i(z) = \bar{a}_i^z$, $i \in N$, $z \in Z$. We can determine the cooperative strategy profile for any subgame $G(z)$, $z \in Z$, starting with simultaneous game $\Gamma(z)$.

We construct a cooperative version of a stochastic game on the basis of a non-cooperative stochastic game with random duration $G(z_0)$ described above. For this purpose it is necessary to define the characteristic function for each subset S (*coalition*) of the set of players N. The characteristic function calculated for the subgame $G(z)$, $z \in Z$, is denoted by $V(S,z)$, where $S \subset N$.

Characteristic function $V(S,z)$ shows which total payoff can be obtained by the players joining into coalition S. There are different approaches to defining the characteristic function that determines the cooperative game on the basis of a non-cooperative one. We introduce some of these approaches:

1. α-approach. In this case, $V(S,z)$ is the *maxmin* value of the zero-sum game between coalitions S and $N\backslash S$. Moreover, the maxmin is found in the pure strategies of coalition S. This approach can be described as "pessimistic", since $V(S,z)$ is equal to the minimum total payoff of coalition S which coalition S can obtain regardless of how coalition $N\backslash S$ behaves. This approach was proposed in the book of Neumann and Morgenstern (von Neumann and Morgenstern, 1944).
2. β-approach. Following this approach, $V(S,z)$ is the *minmax* value of the zero-sum game G_S between coalitions S and $N\backslash S$. Moreover, the minimax is found in pure strategies. This approach can be considered as "optimistic". Comparison of α- and β-approaches can be found in (Aumann and Peleg, 1960).
3. The value of game G_S. In this case, value $V(S)$ is equal to the value of the zero-sum game G_S game between coalitions S and $N\backslash S$. Moreover, this value always exists in mixed strategies, while it is equal to the maxmin and minimax of G_S. In case the minmax and maxmin are found in mixed strategies, the values of α- and β-characteristic functions coincide.
4. γ-approach. According to this approach, $V(S,z)$ is equal to the payoff of coalition S in the Nash equilibrium, when all the players who do not belong to coalition S play individually (Chander and Tulkens, 1997).
5. δ-approach. Value $V(S,z)$ is equal to the maximum payoff of coalition S in the strategy profile when the players who do not belong to coalition S adopt the Nash equilibrium strategies optimal in the n-person game when all players act individually. This approach was proposed in (Petrosjan and Zaccour, 2003) and further considered in detail in the paper (Reddy and Zaccour, 2016).
6. ζ-approach. In this case, $V(S,z)$ is equal to the payoff of coalition S in the strategy profile when the players from coalition S use strategies that maximize the total payoff of coalition N, and the players who do not belong to coalition S minimize the total payoff of the players from coalition S (the idea is proposed in (Gromova and Petrosyan, 2016)).

In this chapter we will use the α-approach and assume that the "power" of coalition S is equal to the maxmin value of a two-person zero-sum stochastic game G_S between coalitions S and $N\backslash S$. This approach was used in the paper (Petrosjan, 2006), in which for the first time a cooperative stochastic game was constructed on the basis of a non-cooperative one and the problem of time-inconsistency of the Shapley value is considered.

We determine the values of the characteristic function. First we consider the case when $S=N$ and find the maximum of the total payoff of the coalition N in stochastic game $G(z_0)$. For this purpose, we write Bellman's equation (see Bellman, 1957) for the maximum sum of the mathematical expectations of players' payoffs:

$$V(N,z_0) = \max_{\substack{a_i^{z_0}\in A_i^{z_0}\\ i\in N}} \left[\sum_{i\in N} K_i^{z_0}(a^{z_0}) + (1-q_0)\sum_{y\in L(z_0)} p(y|z_0,a^{z_0})V(N,y)\right] = \tag{3}$$
$$= \sum_{i\in N} K_i^{z_0}(\bar{a}^{z_0}) + (1-q_0)\sum_{y\in L(z_0)} p(y|z_0,\bar{a}^{z_0})V(N,y)$$

with boundary condition

$$V(N,z) = \max_{\substack{a_i^{z}\in A_i^{z}\\ i\in N}} \sum_{i\in N} K_i^{z}(a^{z}), \quad z\in\{z : L(z)=\varnothing \text{ or } q_k=1\}. \tag{4}$$

Later on in this chapter, we suppose that $z \in (L(z_0))^k$.

For the subgame of $G(z)$, $z \in Z$, the equation (3) with the initial condition (4) takes the form:

$$V(N,z) = \max_{\substack{a_i^z \in A_i^z \\ i \in N}} \left[\sum_{i \in N} K_i^z(a^z) + (1-q_k) \sum_{y \in L(z)} (p(y|z,a^z)V(N,y)) \right] = \quad (5)$$
$$= \sum_{i \in N} K_i^z(\bar{a}^z) + (1-q_k) \sum_{y \in L(z)} (p(y|z,\bar{a}^z)V(N,y))$$

with boundary condition

$$V(N,z) = \max_{\substack{a_i^z \in A_i^z \\ i \in N}} \sum_{i \in N} K_i^z(a^z), \quad z \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}. \quad (6)$$

Strategy profile $(\bar{\varphi}_i : i \in N)$ in stochastic game $G(z_0)$ generates the probability distributions over set Z of the vertices of graph $\Psi(z_0)$.

Definition 2. A subgraph of graph $\Psi(z_0)$, which consists of the vertices $z \in Z$ of the graph $\Psi(z_0)$, having positive realization probabilities, generated by the cooperative strategy profile $\bar{\varphi}(\cdot)$, is called a cooperative subtree and denoted by $\bar{\Psi}(z_0)$.

Obviously, subgraph $\bar{\Psi}(z_0)$ is a finite tree graph. The set of vertices in graph $\bar{\Psi}(z_0)$ is denoted by $CZ \subset Z$.

Let $S \subset N$, $S \neq N$. For each vertex $z \in CZ$ we define the auxiliary zero-sum game denoted by $G_S(z)$. It is a zero-sum game between coalition $S \subset N$ acting as a maximizing player, and coalition $N \setminus S$ acting as a minimizing player. In this case, the payoff of coalition S is calculated as the sum of the payoffs of the players belonging to coalition S. Then, the value of the characteristic function $V(S,z)$ is given by the lower value of zero-sum game $G_S(z)$ in pure strategies (similar to the lower value of the matrix game) [1].

Function $V(S,z)$, $z \in CZ$, satisfies the following functional equation

$$V(S,z) = \max_{a_S^z \in A_S^z} \min_{a_{N\setminus S}^z \in A_{N\setminus S}^z} \left[\sum_{i \in S} K_i^z(a_S^z, a_{N\setminus S}^z) + \right.$$
$$\left. +(1-q_k) \sum_{y \in L(z)} p(y|z,(a_S^z, a_{N\setminus S}^z))V(S,y) \right] \quad (7)$$

with boundary condition

$$V(S,z) = \max_{a_S^z \in A_S^z} \min_{a_{N\setminus S}^z \in A_{N\setminus S}^z} \sum_{i \in S} K_i^z(a_S^z, a_{N\setminus S}^z), \quad z \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}, \quad (8)$$

where $a_S^z = (a_i^z : i \in S)$ is an action of coalition S; $A_S^z = \prod_{i \in S} A_i^z$ is the action set of coalition S; $a_{N\setminus S}^z = (a_j^z : j \in N\setminus S)$ is an action of coalition $N\setminus S$; $A_{N\setminus S}^z = \prod_{j \in N\setminus S} A_j^z$ is the action set of coalition $N\setminus S$.

[1] In this chapter we use the α-approach for construction of the characteristic function, proposed by Neumann and Morgenstern (von Neumann and Morgenstern, 1944).

For all $z \in CZ$ it is natural to suppose that

$$V(\varnothing, z) = 0. \quad (9)$$

Thus, for each subgame $G(z)$, $z \in CZ$, we have determined the characteristic function $V(S, z)$, $S \subset N$. The characteristic function $V(S, z)$ is determined by the equation (5) with the boundary condition (6), and also the equation (7) with the boundary condition (8) and equation (9).

The characteristic function $V(S, z)$ defined by formulas (5) – (9) is superadditive on S, i. e., for any vertex $z \in CZ$ and any coalitions $S, P \subset N$, $S \cap P = \varnothing$, the inequality

$$V(S \cup P, z) \geqslant V(S, z) + V(P, z).$$

holds.

Definition 3. A cooperative stochastic game with random duration $\bar{G}(z_0)$ constructed on the basis of non-cooperative stochastic game $G(z_0)$ is a tuple $\langle N, V(S, z_0)\rangle$, where $V(S, z_0)$ is a characteristic function defined by equation (5) with boundary condition (6) for coalition N, by equation (7) with boundary condition (8) for coalition $S \neq N$, $S \neq \varnothing$, and by formula (9) for coalition $S = \varnothing$.

Definition 4. An imputation in cooperative stochastic game $\bar{G}(z_0)$ is a vector $\xi(z_0) = (\xi_1(z_0), \ldots, \xi_n(z_0))$, satisfying two properties:

1. Collective rationality: $\sum_{i \in N} \xi_i(z_0) = V(N, z_0)$;
2. Individual rationality: $\xi_i(z_0) \geqslant V(\{i\}, z_0)$ for any $i \in N$.

The set of imputations (see (Vilkas, 1990, Vorobiev, 1960, Vorobiev, 1967) and also (Vorobiev, 1985, Pecherski and Yanovskaya, 2004) for definitions of cooperative games) of cooperative stochastic game $\bar{G}(z_0)$ is denoted by $I(z_0)$.

Definition 5. A solution of cooperative stochastic game $\bar{G}(z_0)$ is a subset $C(z_0)$ of the set of imputations $I(z_0)$.

The solutions of a cooperative game can be conventionally divided into single-valued and multi-valued ones. The well-known single-valued solutions are the Shapley value (Shapley, 1953b), the nucleolus (Schmeidler, 1969). The most well-known multi-valued solution is the core (Gillies, 1959). Suppose that solution $C(z_0)$ of cooperative stochastic game $\bar{G}(z_0)$ is a non-empty subset of the imputation set $I(z_0)$.

Definition 6. A cooperative stochastic subgame $\bar{G}(z)$, $z \in Z$, of game $\bar{G}(z_0)$, constructed on the basis of non-cooperative stochastic subgame $G(z)$, is a pair $\langle N, V(S, z)\rangle$, where $V(S, z)$ is the characteristic function defined by equation (5) with boundary condition (6) for coalition N, by equation (7) with boundary condition (8) for coalition $S \neq N$, $S \neq \varnothing$, and by formula (9) for coalition $S = \varnothing$.

Determine the imputation, the imputation set and the solution for any cooperative subgame $\bar{G}(z)$, $z \in Z$.

Definition 7. An imputation in cooperative stochastic subgame $\bar{G}(z)$ is vector $\xi(z) = (\xi_1(z), \ldots, \xi_n(z))$, satisfying two properties:

1. $\sum_{i \in N} \xi_i(z) = V(N, z)$;

2. $\xi_i(z) \geqslant V(\{i\}, z)$ for any $i \in N$.

The set of imputations of cooperative stochastic subgame $\bar{G}(z)$ is denoted by $I(z)$.

Definition 8. The solution of a cooperative stochastic subgame $\bar{G}(z)$ is a subset $C(z)$ of the set of imputations $I(z)$.

Suppose that solution $C(z)$ of any cooperative subgame $\bar{G}(z)$ is non-empty subset of the imputation set $I(z)$ for any $z \in CZ$.

2.4. The Shapley value, core and nucleolus

In this section we define some cooperative solutions which will be used further in the work. The Shapley value of a cooperative stochastic game or subgame $\bar{G}(z)$, $z \in CZ$, is a vector $Sh(z) = (Sh_1(z), \ldots, Sh_n(z))$, where an element $Sh_i(z)$, $i \in N$, is calculated by the formula

$$Sh_i(z) = \sum_{\substack{S \subseteq N \\ S \ni i}} \frac{(|S|-1)!\,(n-|S|)!}{n!} \left(V(S,z) - V(S \setminus \{i\}, z)\right),$$

where $|S|$ is the cardinality of S. The definition of the Shapley value is introduced in Shapley's paper (Shapley, 1953b).

A core of a cooperative stochastic game or subgame $\bar{G}(z)$, $z \in CZ$, is a set denoted by $CO(z)$, and it is the set

$$CO(z) = \left\{ \xi(z) \in I(z) : \sum_{i \in S} \xi_i(z) \geqslant V(S,z) \text{ for } \forall S \subset N, \ \sum_{i \in N} \xi_i(z) = V(N,z) \right\}. \tag{10}$$

For the cooperative stochastic game or subgame $\bar{G}(z)$ and any vector $\xi(z) \in I(z)$, by $\theta(\xi(z))$ we denote the vector of the values of excesses $e(S, \xi(z)) = V(S,z) - \sum_{i \in S} \xi_i(z)$ located in a descending order:

$$\theta(\xi(z)) = (e(S_1, \xi(z)), e(S_2, \xi(z)), \ldots, e(S_{2^n-1}, \xi(z))),$$

where coalitions are numbers that $e(S_1, \xi(z)) \geqslant e(S_2, \xi(z)) \geqslant \ldots \geqslant e(S_{2^n-1}, \xi(z))$.

On the set of excesses $\{\theta(\xi(z)) : \xi(z) \in I(z)\}$ we consider the lexicographic ordering $\succ_{lex}$:

$$\theta(\xi(z)) \succ_{lex} \theta(\psi(z)) \Longleftrightarrow \exists \ \ l \in \{1, \ldots, 2^n\},$$

such that

$$\begin{cases} \theta_k(\xi(z)) = \theta_k(\psi(z)), & \text{for all } k = 1, \ldots, l-1; \\ \theta_l(\xi(z)) > \theta_l(\psi(z)), & \end{cases}$$

where $\psi(z) \in I(z)$.

The definition of the nucleolus is first introduced in (Schmeidler, 1969). The nucleolus of a cooperative stochastic game or subgame $\bar{G}(z)$, $z \in CZ$, is a subset of the imputation set on which $\min\limits_{\substack{\theta(\xi(z)) \\ \xi(z) \in I(z)}} \succ_{lex}$ is reached.

If $C(z_0)$ is the solution of cooperative stochastic game $\bar{G}(z_0)$, then later on in the work by solution $C(z)$ of cooperative subgame $\bar{G}(z)$ we mean a solution constructed according to the same "rules" as $C(z_0)$. For example, if $C(z_0)$ is the Shapley value in stochastic game $\bar{G}(z_0)$, then $C(z)$ is the Shapley value, calculated for cooperative subgame $\bar{G}(z)$, $z \in CZ$.

Here we assume that the players choose some fixed subset of the imputation set which contains the imputations satisfing "optimal" properties, i. e., the players forming coalition N, are going to follow some "rule" distributing the payoffs of coalition N throughout the game process. Set $C(z)$ may consist of a single imputation, if, e. g., the players have decided to use the Shapley value or the nucleolus, or it may be empty if, e. g., they have chosen the core and it is empty. The solution of the game or subgame $\bar{G}(z)$ can be any other imputations from the classical "static" cooperative theory, such as von Neumann-Morgenstern solution (or the so-called stable set), the kernel, $\mathcal{M}$-stable sets (see Pecherski and Yanovskaya, 2004).

Further in the work we will suppose that $C(z)$ is a nonempty subset of set $I(z)$ for any $z \in CZ$, that is, for each vertex $z \in CZ$ there exists at least one imputation

$$\xi(z) = (\xi_1(z), \ldots, \xi_n(z)) \in C(z) \subset I(z).$$

2.5. Imputation distribution procedure

In this section we introduce the definition of an imputation distribution procedure of the cooperative stochastic game solution, which has been chosen by the players. The imputation distribution procedure determines the payments to the players at each vertex of the cooperative subtree $\bar{\Psi}(z_0)$.

Definition 9. A path in a stochastic game is the sequence of action profiles a^{z_0}, a^{z_1}, ..., a^{z_l}, where a^{z_i} is the action profile realized in the game $\Gamma(z_i)$, $z_i \in L(z_{i-1})$, $i = 1, \ldots, l$.

Consider any vertex $z \in CZ$, $z \in (L(z_0))^k$, of the cooperative subtree. Each player receives some payments implementing a cooperative agreement [2]. Let at the vertex $z \in CZ$ the payment to player $i \in N$ be $\beta_i(z)$. In any cooperative subgame $\bar{G}(z)$, the player can calculate the sum of the payments along the path $\bar{a}^z, \ldots, \bar{a}^{z_l} = \bar{a}^{z,\ldots,z_l}$, and this sum is a random variable. We denote by $B_i(z)$ the mathematical expectation of the sum of such payments, calculated along the path segment $\bar{a}^{z,\ldots,z_l}$ in cooperative subgame $\bar{G}(z)$. The value $B_i(z)$ satisfies the following functional equation:

$$B_i(z) = \beta_i(z) + (1 - q_k) \sum_{y \in L(z)} p(y|z, x^z) B_i(y) \tag{11}$$

with boundary condition

$$B_i(z) = \beta_i(z) \text{ for } z \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}. \tag{12}$$

Now we define the distribution procedure of the imputation belonging to the cooperative solution $C(z_0)$, chosen by the players at the beginning of the game.

Definition 10. Let $\xi(z_0)$ be the vector $(\xi_1(z_0), \ldots, \xi_n(z_0)) \in C(z_0)$. The set of vectors $\{\beta(z) = (\beta_1(z), \ldots, \beta_n(z)) : z \in CZ\}$ is called a distribution procedure of the imputation $\xi(z_0)$ if the following conditions are satisfied:

[2] Obviously, all $z, \ldots, z_l \in CZ$, since CZ is the set of vertices of the cooperative subtree, and the strategy profile $\bar{\varphi}$ is determined.

1. For each vertex $z \in CZ$:

$$\sum_{i \in N} \beta_i(z) = \sum_{i \in N} K_i^z(\bar{a}^z).$$

2. The components $\xi_i(z_0)$, $i \in N$, of imputation ξ coincide with the mathematical expectation of the corresponding components of the imputation distribution procedure with respect to the probability distribution of transitions and the end of the game, i. e., $\xi_i(z_0) = B_i(z_0)$, where $B_i(z_0)$ satisfies the functional equation (11) with the boundary condition (12).

For each cooperative subgame $\bar{G}(z)$, $z \in CZ$, we write the functional equation for the components $\xi_i(z)$ of the imputation $\xi(z) \in C(z) \subset I(z)$ of type (11) and define the values $\gamma_i(z)$ from equation:

$$\xi_i(z) = \gamma_i(z) + (1 - q_k) \sum_{y \in L(z)} p(y|z, x^z)\xi_i(y), \tag{13}$$

where $\xi(y) = (\xi_i(y) : i \in N)$ is an imputation belonging to the solution $C(y)$ of the cooperative subgame $\bar{G}(y)$. The boundary condition for $\gamma_i(z)$ is as follows:

$$\gamma_i(z) = \xi_i(z) \text{ for } z \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}. \tag{14}$$

Lemma 1. *The vector $\gamma(z) = (\gamma_i(z) : i \in N)$ given by equation* (13) *with the boundary condition* (14) *is an imputation distribution procedure.*

Proof. It is obvious that for terminal vertices and the vertices at which the probability of the game end equals one, and the equality (14) holds, conditions 1 and 2 of Definition 10 are satisfied.

Now we prove that these conditions are satisfied for the remaining vertices of the cooperative subtree. From (13) we express the values $\gamma_i(z)$ and summing them up over $i \in N$, and obtain

$$\sum_{i \in N} \gamma_i(z) = \sum_{i \in N} \xi_i(z) - (1 - q_k) \sum_{i \in N} \left(\sum_{y \in L(z)} p(y|z, x^z)\xi_i(y) \right). \tag{15}$$

As we have

$$\xi(z) = (\xi_i(z) : i \in N) \in C(z) \subset I(z),$$
$$\xi(y) = (\xi_i(y) : i \in N) \in C(y) \subset I(y),$$

then from (15) we obtain:

$$\sum_{i \in N} \gamma_i(z) = V(N, z) - (1 - q_k) \sum_{y \in L(z)} p(y|z, x^z)V(N, y). \tag{16}$$

From (16) and (5) it follows that $\sum_{i \in N} \gamma_i(z) = \sum_{i \in N} K_i^z(\bar{a}^z)$ for action profile $\bar{a}^z = (\bar{a}_i^z : i \in N)$, which has been realized in game $\Gamma(z)$ when the players used a cooperative strategy profile $\bar{\varphi}$. Therefore, $\gamma_i(z)$ satisfies Condition 1 of Definition 10.

Now we verify if Condition 2 of Definition 10 is satisfied. Specifically, we find the mathematical expectation of the sums $\gamma_i(z)$, defined by formula (13), along the vertices of the cooperative subtree. For the vertices $z \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}$, Condition 2 is satisfied. Continue with the vertices of the cooperative subtree, from which the vertices mentioned above are reached with one stage. For these vertices, we obtain the equality:

$$B_i(z_l) = \xi_i(z_l) - (1 - q_l) \sum_{y \in L(z_l)} p(y|z_l, x^{z_l})\xi_i(y) + (1 - q_l) \sum_{y \in L(z_l)} p(y|z_l, x^{z_l})\gamma_i(y)$$
$$= \xi_i(z_l),$$

because $\xi_i(y) = \gamma_i(y)$. Following from the terminal vertices to the initial one, we prove that condition 2 of Definition 10 is satisfied. Lemma is proved.

2.6. Subgame consistency of cooperative stochastic game solution

Before the game starts, players come to an agreement about cooperation, i. e., they agree to maximize the mathematical expectation of the total payoff of coalition N and expect to receive the imputation $\xi(z_0) \in C(z_0)$. The game process takes place along the vertices of the cooperative subtree $\bar{\Psi}(z_0)$. But since the stochastic structure of the game implies uncertainty in realization of the vertices of the cooperative subtree, then moving along a certain path, that is, along the vertices of the cooperative subtree, does not yet ensure the support of cooperation. Indeed, players moving along the cooperative path get into cooperative subgames with the current initial states in which the same player may have different opportunities. Conditions of a conflict and players' opportunities involved in the conflict change over time. And it will be natural to require maintenance of the optimality principle or "approach" in the choice of solutions of cooperative subgames. But at some moment, at vertex $z \in CZ$, the sum of the remaining payments to player i may not be equal to the ith component of the imputation from solution $C(z)$ of a cooperative subgame $\bar{G}(z)$. Therefore, at vertex $z \in CZ$ player i may ask a question whether it is worth keeping the cooperative agreement to act "jointly optimally" proposed before the game starts. Thus, player i may wish to deviate from the cooperative strategy profile. If this deviation is beneficial for at least one player, it means subgame inconsistency of imputation $\xi(z_0) \in C(z_0)$ and, accordingly, the motion along the vertices of the cooperative subtree.

Definition 11. An imputation $\xi(z_0) \in C(z_0)$ is called subgame-consistent in cooperative stochastic game $\bar{G}(z_0)$ if for each vertex$z \in CZ \cap (L(z_0))^k$ there exists the imputation distribution procedure $\beta(z) = (\beta_i(z) : i \in N)$ such that

$$\xi_i(z) = \beta_i(z) + (1 - q_k) \sum_{y \in L(z)} p(y|z, x^z)\xi_i(y), \tag{17}$$

and

$$\xi_i(z) = \beta_i(z), \quad z \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}, \tag{18}$$

where $\xi(y) = (\xi_i(y) : i \in N)$ is an imputation belonging to solution $C(y)$ of cooperative subgame $\bar{G}(y)$.

Remark 2. If $C(z_0)$ consists of more than one imputation, then the choice of the imputation $\xi(z_0)$ is indefinite. If players have chosen a certain imputation

$\xi(z_0) \in C(z_0)$ and decided to verify if it is subgame consistent, first it is necessary to check the condition (17) for the vertex z_0. This means to verify if there exists the imputation distribution procedure $\beta(z_0) = (\beta_i(z_0) : i \in N)$, satisfying condition (17) for some imputation $\xi(y) \in C(y)$, where $y \in L(z_0)$. Obviously, there is indefiniteness in the choice of imputation $\xi(y) \in C(y)$, which in its turn should also be subgame consistent in cooperative subgame $\bar{G}(y)$. This means that condition (17) should be satisfied for imputation imputation $\xi(y) \in C(y)$. From Definition 11 it follows that this condition should be satisfied for all z from the set of vertices of the cooperative subtree.

Definition 12. We say that cooperative stochastic game $\bar{G}(z_0)$ has subgame consistent solution $C(z_0)$ if all imputations $\xi(z_0) \in C(z_0)$ are subgame consistent.

Obviously, if the payments to the players are made at the vertices of the cooperative subtree in accordance with the initially defined payoff functions, it is impossible in general to achieve subgame consistency of the cooperative solution. This may lead to the breakup of the cooperative agreement. In this connection, the problem of finding a scheme or procedure of payments to the players at the vertices of the cooperative subtree in order to satisfy the property of subgame consistency of a cooperative solution. For this we need to find such an imputation distribution procedure $(\beta_i(z) : i \in N)$ for all vertices $z \in CZ$, for which the conditions (17) and (18) are satisfied.

Theorem 1. *Let in the cooperative stochastic game $\bar{G}(z_0)$ and each subgame the cooperative solutions $C(z_0)$ and $C(z)$, $z \in CZ$, be nonempty. If for each $\xi(z) = (\xi_i(z) : i \in N) \in C(z)$ the imputation distribution procedure is defined by the formula*

$$\beta_i(z) = \xi_i(z) - (1 - q_k) \sum_{y \in L(z)} p(y|z, x^z)\xi_i(y), \tag{19}$$

for each $z \in CZ$, $z \notin \{z : L(z) = \varnothing\}$, where $\xi(y) = (\xi_i(y) : i \in N) \in C(y)$, and by formula (18) *for any $z \in \{z : L(z) = \varnothing\}$, then cooperative solution $C(z_0)$ is subgame consistent.*

Proof. To prove subgame consistency of the cooperative solution $C(z_0)$, it is required to prove that for each vector $\xi(z_0) \in C(z_0)$ conditions (17) and (18) are satisfied.

From Lemma 1 it follows that the payments, determined by formulas (19) and (18), are the components of the imputation distribution procedure. Condition (17) follows from (19) taking into account that $\xi(y) = (\xi_i(y) : i \in N)$ belongs to the cooperative solution of the subgame $\bar{G}(y)$.

The proposed method of implementing the imputation has an important property: at each vertex of the cooperative path, players are guided by the same "optimality principle" (property of subgame consistency) and, in this sense, have no reasons for interruption of the previously adopted cooperative agreement and deviation from the cooperative strategy profile. The sum of payments to the players at each vertex of the cooperative subtree is also equal to the sum of the payoffs received by the players at that vertex (condition 1 of Definition 10 of an imputation distribution procedure). The latter condition may be called a condition of attainability of the payments, since players redistribute the sum which they obtain in the game and do not take any funds outside.

Notice that Definition 11 does not require the nonnegativity of functions $\beta_i(z)$, where $z \in CZ$. All imputations belonging to the solution $C(z)$ will be subgame consistent if solution is such that $C(z) \neq \varnothing$ for all vertices $z \in CZ$. This is possible if the payments to the players are not made according to their initially defined payoffs in games along which the cooperative path realizes, but according to the imputation distribution procedure $\beta(z) = (\beta_1(z), \ldots, \beta_n(z))$ defined by (17), (18) for all $z \in CZ$, where $\beta_i(z)$ is the payment to player i at the vertex $z \in CZ$. Moreover, the mathematical expectation of all payments to player i coincides with the mathematical expectation of the ith component of the imputation belonging to the solution chosen by the players. It follows from Theorem 1. Thus, players can agree on getting negative payments at some vertices to ensure that the cooperation is supported throughout the whole game in order to guarantee receiving the components of initially selected imputation $\xi(z_0)$ partition belonging to the solution $C(z_0)$ of the cooperative stochastic game $\bar{G}(z_0)$.

2.7. Nonnegative components of imputation distribution procedure. Regularization of imputations

In this section, we consider the case when for any player $i \in N$ payoff function is non-negative: $K_i(x^z) \geqslant 0$ for all vertices $z \in CZ$. Assume that the players are interested in receiving non-negative payments at each vertex of the cooperative subtree and at the same time they want to guarantee subgame consistency of the cooperative solution. In case when non-negativity of $\beta_i(z)$ cannot be guaranteed for all vertices $z \in CZ$, one can construct new subgame-consistent solution based on the solution initially chosen by the players from the set $C(z_0)$. We present how this is done when the set $C(z_0) \subset I(z_0)$ is considered as the solution. Notice that this procedure can be applied to the imputations well-known in the classical "static" cooperative game theory (core, nucleolus, von Neumann-Morgenstern solution).

For each vertex $z \in CZ$ define new imputation distribution procedure by

$$\beta_i(z) = \frac{\sum\limits_{i \in N} K_i(\bar{a}_1^z, \ldots, \bar{a}_n^z)}{V(N, z)} \xi_i(z), \tag{20}$$

where $\xi(z) = (\xi_1(z), \ldots, \xi_n(z)) \in C(z)$, and $\bar{a}^z = (\bar{a}_1^z, \ldots, \bar{a}_n^z)$ is the realization of the cooperative strategy profile $\bar{\varphi} = (\bar{\varphi}_1(\cdot), \ldots, \bar{\varphi}_n(\cdot))$ at vertex $z \in CZ$ maximizing the sum of mathematical expectations of the players' payoffs in stochastic game $\bar{G}(z_0)$, $V(N, z)$ is the value of characteristic function of coalition N calculated for cooperative subgame $\bar{G}(z)$.

As $K_i(a^z) \geqslant 0$ for each vertex $z \in CZ$ and each player $i \in N$, then $\beta_i(z) \geqslant 0$ for each vertex $z \in CZ$. Taking into account equation (20) and equity $\sum_{i \in N} \xi_i(z) = V(N, z)$, we obtain that the current payment $\beta_i(z)$ to player i in game $\Gamma(z)$ should be proportional to the ith component of the imputation $\xi(z) \in C(z)$ in cooperative subgame $\bar{G}(z)$ of stochastic game $\bar{G}(z_0)$.

Determine a new imputation for cooperative subgame $\bar{G}(z)$, where $z \in CZ$, and $z \in (L(z_0))^k$ on the basis of the "old" imputation $\xi(z)$ as a solution of the functional equation

$$\hat{\xi}_i(z) = \frac{\sum\limits_{i \in N} K_i(\bar{a}^z)}{V(N, z)} \xi_i(z) + (1 - q_k) \sum_{y \in L(z)} p(y | z, \bar{a}^z) \hat{\xi}_i(y) \tag{21}$$

with boundary condition

$$\hat{\xi}_i(z) = \frac{\sum\limits_{i \in N} K_i(\bar{a}^z)}{V(N,z)} \xi_i(z) = \xi_i(z) \tag{22}$$

for $z \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}$.

Construct a new characteristic function $\hat{V}(S,z)$ for each cooperative subgame $\bar{G}(z)$ for all $z \in CZ$ using functional equation

$$\hat{V}(S,z) = \frac{\sum\limits_{i \in N} K_i(\bar{a}^z)}{V(N,z)} V(S,z) + (1 - q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) \hat{V}(S,y) \tag{23}$$

with boundary condition

$$\hat{V}(S,z) = V(S,z) \text{ for } z \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}\,. \tag{24}$$

Functions $\hat{V}(S,z)$ and $V(S,z)$ are superadditive, and $\hat{V}(N,z) = V(N,z)$ because $\hat{V}(N,z)$ and $V(N,z)$ satisfy the functional equation (5) with boundary condition (6).

For all vertices of $z \in CZ$ and all subgame-inconsistent imputations $\xi(z) \in C(z)$, we compute the regularized imputations $\hat{\xi}(z)$ and define the set of solutions $\hat{C}(z)$ as follows:

$$\hat{C}(z) = \Bigg\{ \hat{\xi}(z) : \hat{\xi}_i(z) = \frac{\sum\limits_{i \in N} K_i(\bar{a}^z)}{V(N,z)} \xi_i(z) + (1 - q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) \hat{\xi}_i(y), \tag{25}$$

$$\hat{\xi}_i(z) = \xi_i(z) \text{ for } z \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}\,, \xi(z) \in C(z) \Bigg\}.$$

Definition 13. The set $\hat{C}(z_0)$ defined by formula (25), is called the regularized solution of the cooperative stochastic game $\bar{G}(z_0)$.

Therefore, players have an opportunity to regularize the solution chosen at the beginning of the game so that at each vertex of the stochastic game $\bar{G}(z_0)$ "new" solution $\hat{C}(z_0)$ is subgame consistent. But the imputation belonging to the new regularized solution $\hat{C}(z_0)$, generally speaking, will not be an imputation for cooperative game with the characteristic function $V(S, z_0)$, defined by (7) and (8). It will be an imputation for a cooperative stochastic game with a new characteristic function $\hat{V}(S, z_0)$ defined by formulas (23), (24).

Theorem 2. *An imputation* $\hat{\xi}(z) = (\hat{\xi}_1(z), \ldots, \hat{\xi}_n(z))$*, defined by formula* (21) *with boundary condition* (22)*, is subgame consistent imputation in cooperative game* $\langle N, \hat{V} \rangle$ *where characteristic function* $\hat{V}(S,z)$ *is defined by functional equation* (23) *with boundary condition* (24).

Proof. Subgame consistency follows from the method of construction of a "new" imputation $\hat{\xi}(z)$. Comparing the functional equations (17) and (21), we obtain that for the proof it is necessary to show the non-negativity of the component

$$\frac{\sum\limits_{i \in N} K_i(\bar{a}^z)}{V(N,z)} \xi_i(z),$$

which is obvious because

$$K_i^z(a_1^z, \dots, a_n^z) \geqslant 0$$

for all $z \in Z$ and each player $i \in N$.

Now we prove that $\hat{\xi}(z) = (\hat{\xi}_1(z), \dots, \hat{\xi}_n(z))$ has the properties of an imputation in cooperative game with characteristic function $\hat{V}(S, z)$, which is given by the functional equation (23) with the boundary condition (24). To do this, for any player $i \in N$ and each vertex $z \in CZ$, it is necessary to prove satisfaction of two properties:

1. $\sum_{i \in N} \hat{\xi}_i(z) = \hat{V}(N, z)$,
2. $\hat{\xi}_i(z) \geqslant \hat{V}(\{i\}, z)$.

The first property is obviously satisfied for vertices $z \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}$ and $z \in CZ$. Now prove these properties for vertices $z \in \{z : L(z) \ni y \text{ and } L(y) = \varnothing\}$ and such that $z \in CZ$:

$$\sum_{i \in N} \hat{\xi}_i(z) = \frac{\sum_{i \in N} K_i(\bar{a}^z)}{V(N, z)} \sum_{i \in N} \xi_i(z) + (1 - q_k) \sum_{y \in L(z)} \left(p(y|z, \bar{a}^z) \sum_{i \in N} \hat{\xi}_i(y) \right) =$$

$$= \frac{\sum_{i \in N} K_i(\bar{a}^z)}{V(N, z)} V(N, z) + (1 - q_k) \sum_{y \in L(z)} \left(p(y|z, \bar{a}^z) \hat{V}(N, z) \right) =$$

$$= V(N, z) = \hat{V}(N, z),$$

because $y \in \{y : L(y) = \varnothing\}$.

The second property is also obviously satisfied for the vertices $z \in \{z : L(z) = \varnothing$ or $q_k = 1\}$. We show that $\hat{\xi}_i(z) - \hat{V}(\{i\}, z) \geqslant 0$ for the vertices $z \in \{z : L(z) \ni y$ and $L(y) = \varnothing\}$, using formulas (21) and (23):

$$\hat{\xi}_i(z) - \hat{V}(\{i\}, z) = \frac{\sum_{i \in N} K_i(\bar{a}^z)}{V(N, z)} \xi_i(z) + (1 - q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) \hat{\xi}_i(y) -$$

$$- \left\{ \frac{\sum_{i \in N} K_i(\bar{a}^z)}{V(N, z)} V(\{i\}, z) + (1 - q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) \hat{V}(\{i\}, y) \right\} =$$

$$= \frac{\sum_{i \in N} K_i(\bar{a}^z)}{V(N, z)} \left(\xi_i(z) - V(\{i\}, z) \right) +$$

$$+ (1 - q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) \left(\hat{\xi}_i(y) - \hat{V}(\{i\}, y) \right) \geqslant 0.$$

The first term is non-negative since $\xi(z)$ is an imputation of cooperative subgame $\bar{G}(z)$, and the second term is non-negative, because $y \in \{y : L(y) = \varnothing\}$. We prove recursively for the previous vertices $z \in CZ$ and so on until vertex z_0.

It is important to know in what relation the set $\hat{C}(z)$ which is a regularized solution defined by the formula (25), and the set $\tilde{C}(z)$ which is the solution found for the cooperative subgame $\bar{G}(z)$ with the characteristic function$\hat{V}(S,z)$ (i. e., the solution constructed using the same rules as the solution $C(z) \subset I(z)$ for the cooperative subgame $\bar{G}(z)$). Now we find the sets $\tilde{C}(z)$ and $\hat{C}(z)$ for the cooperative stochastic subgame $\bar{G}(z)$ if when the solutions of the stochastic game $\bar{G}(z_0)$ are the imputations (the Shapley value and the core) from the classical "static" theory of cooperative games.

2.8. Regularization of the Shapley value and the core

We start with the case when players choose the single-point optimality principle—Shapley value—as a cooperative solution. The Shapley value calculated in cooperative stochastic game $\bar{G}(z_0)$, is denoted by $Sh(z_0) = (Sh_i(z_0) : i \in N)$, and in cooperative subgame $\bar{G}(z)$, where $z \in CZ$, by $Sh(z) = (Sh_i(z) : i \in N)$.

Define the regularized Shapley value in cooperative subgame $\bar{G}(z)$, where $z \in CZ$, and $z \in (L(z_0))^k$ based on the Shapley value of the initially given game as a solution of the functional equation

$$\hat{Sh}_i(z) = \frac{\sum\limits_{i\in N} K_i(\bar{a}^z)}{V(N,z)} Sh_i(z) + (1-q_k) \sum_{y \in L(z)} p(y|z,\bar{a}^z)\hat{Sh}_i(y) \tag{26}$$

with boundary condition

$$\hat{Sh}_i(z) = Sh_i(z) \tag{27}$$

for $z \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}$.

The following theorem holds.

Theorem 3. *Vector satisfying the functional equation* (26) *with boundary condition* (27), *is subgame-consistent and it is the Shapley value of the cooperative subgame* $\langle N, \hat{V}(\cdot,z)\rangle$, $z \in CZ$ *of stochastic game* $\langle N, \hat{V}(\cdot,z_0)\rangle$, *where the values of characteristic function* $\hat{V}(\cdot,z)$ *are calculated by formulas* (23) *and* (24).

Remark 3. Theorem 3 provides the relation between the sets $\tilde{C}(z)$ and $\hat{C}(z)$, which are mentioned at the end of the previous paragraph. If the Shapley value is chosen as a solution of the stochastic game $\bar{G}(z_0)$, then $\tilde{C}(z) = \hat{C}(z)$ for any $z \in CZ$. Therefore, we may reformulate Theorem 2 in the following way.

Theorem 4. *Vector satisfying the functional equation* (26) *with boundary condition* (27), *is subgame-consistent, and* $\hat{Sh}(z_0) = \hat{C}(z_0) = \tilde{C}(z_0)$, *where* $\hat{C}(z_0)$ *is a regularized solution satisfying equation* (25), *and* $\tilde{C}(z_0)$ *is the Shapley value of the cooperative stochastic game* $\langle N, \hat{V}(\cdot,z_0)\rangle$ *with characteristic function given by formulas* (23), (24).

Proof. The fact that the vector satisfying the functional equation (26) with initial condition (27) is subgame-consistent, follows from Theorem 1 which is formulated for a general case, i. e. , for any solution $C(z)$.

Calculate the Shapley value of cooperative stochastic game $\langle N, \hat{V}(\cdot,z)\rangle$, $z \in CZ$, with regularized characteristic function given by formulas (23), (24):

$$\hat{Sh}_i(z) = \sum_{\substack{S\subset N\\ S\ni i}} \frac{(|S|-1)!\,(n-|S|)!}{n!} \left(\hat{V}(S,z) - \hat{V}(S\setminus\{i\},z)\right).$$

Rewrite (23) for coalition $S \setminus \{i\}$ and obtain

$$\hat{V}(S \setminus \{i\}, z) = \frac{\sum\limits_{i\in N} K_i(\bar{a}^z)}{V(N,z)} V(S \setminus \{i\}, z) + \\ + (1-q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) \hat{V}(S \setminus \{i\}, y). \quad (28)$$

Subtracting (28) from (23), multiplying by $\frac{(|S|-1)!(n-|S|)!}{n!}$ and summing up over the all possible coalitions $S \subset N$ such that $S \ni i$, we obtain

$$\sum_{\substack{S \subset N \\ S \ni i}} \frac{(|S|-1)!\,(n-|S|)!}{n!} \left[\hat{V}(S,z) - \hat{V}(S \setminus \{i\}, z)\right] = \quad (29)$$

$$= \left\{ \sum_{\substack{S \subset N \\ S \ni i}} \frac{(|S|-1)!\,(n-|S|)!}{n!} \left[V(S,z) - V(S \setminus \{i\}, z)\right] \right\} \frac{\sum\limits_{i\in N} K_i(\bar{a}^z)}{V(N,z)} +$$

$$+ (1-q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) \times$$

$$\times \left\{ \sum_{\substack{S \subset N \\ S \ni i}} \frac{(|S|-1)!\,(n-|S|)!}{n!} \left[\hat{V}(S,y) - \hat{V}(S \setminus \{i\}, y)\right] \right\} =$$

$$= Sh_i(z) \frac{\sum\limits_{i\in N} K_i(\bar{a}^z)}{V(N,z)} + (1-q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) \hat{Sh}_i(y).$$

The result of the theorem follows from (29) and (26).

Now we assume that the players choose the core as a solution of cooperative stochastic game $\bar{G}(z_0)$. As before, we suppose that $CO(z) \neq \varnothing$ for any vertex $z \in CZ$. We also assume that $CO(z_0)$ is not subgame-consistent, i. e., there exists at least one imputation $\xi(z_0) \in CO(z_0)$ for which the condition of subgame consistency is not satisfied.

Definition 14. The regularized core of stochastic game $\bar{G}(z_0)$ is the set:

$$\widehat{CO}(z_0) = \Bigg\{ \hat{\xi}(z_0) : \hat{\xi}_i(z_0) = \frac{\sum\limits_{i\in N} K_i(\bar{a}^{z_0})}{V(N,z_0)} \xi_i(z_0) + (1-q_0) \sum_{y \in L(z_0)} (p(y|z_0, \bar{a}^{z_0}) \hat{\xi}_i(y)), \\ \hat{\xi}_i(z_0) = \xi_i(z_0),\ z_0 \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}, \xi(z_0) \in CO(z_0) \Bigg\}. \quad (30)$$

Definition 15. The regularized core of cooperative subgame $\bar{G}(z)$ is the set $\widehat{CO}(z)$ defined as:

$$\widehat{CO}(z) = \Bigg\{ \hat{\xi}(z) : \hat{\xi}_i(z) = \frac{\sum\limits_{i\in N} K_i(\bar{a}^{z})}{V(N,z)} \xi_i(z) + (1-q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^{z}) \hat{\xi}_i(y) \\ \hat{\xi}_i(z) = \xi_i(z) \text{ for } z \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}, \xi(z) \in CO(z) \Bigg\}. \quad (31)$$

Denote by $\widetilde{CO}(z)$ the core calculated for cooperative subgame $\langle N, \hat{V}(\cdot, z)\rangle$, $z \in CZ$, with characteristic function $\hat{V}(S, z)$, defined by formulas (23), (24). We prove the theorem providing the relation between $\widetilde{CO}(z)$ and $\widehat{CO}(z)$.

Theorem 5. *The regularized core defined by formula* (30) *is subgame-consistent solution. Moreover,* $\widehat{CO}(z_0) \subset \widetilde{CO}(z_0)$, *where* $\widetilde{CO}(z_0)$ *is the core of cooperative stochastic game* $\langle N, \hat{V}(\cdot, z)\rangle$ *with characteristic function defined by formulas* (23), (24).

Proof. Subgame consistency of the core follows from Theorem 1. To prove that $\widehat{CO}(z_0) \subset \widetilde{CO}(z)$, we need to prove that any imputation $\hat{\xi}(z_0) \in \widehat{CO}(z_0)$ belongs to the set $\widetilde{CO}(z_0)$, which is equivalent to the following: for any $\hat{\xi}(z) \in \widehat{CO}(z)$, $z \in CZ$ and $S \subset N$ the inequality

$$\sum_{i \in S} \hat{\xi}_i(z) \geqslant \hat{V}(S, z) \tag{32}$$

is true.

The proof is obvious for the vertices $z \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}$. Now we prove this inequality for the vertices $z \in \{z : L(z) \ni y \text{ and } L(y) = \varnothing\}$:

$$\sum_{i \in S} \hat{\xi}_i(z) = \frac{\sum\limits_{i \in N} K_i(\bar{a}^z)}{V(N, z)} \sum_{i \in S} \xi_i(z) + (1 - q_k) \sum_{y \in L(z)} \left(p(y|z, \bar{a}^z) \sum_{i \in S} \hat{\xi}_i(y) \right) \geqslant$$

$$\geqslant \hat{V}(S, z),$$

which is true because $y \in \{z : L(z) = \varnothing \text{ or } q_k = 1\}$ and $\sum_{i \in S} \xi_i(z) \geqslant V(S, z)$, as $\xi(z)$ is the imputation belonging to the core $CO(z)$.

The following part of the proof is made for the next vertices up to the initial vertex z_0 like in the proof of Theorem 1.

Now we consider examples of construction and regularization of the solution in cooperative stochastic games defined on the graphs.

Example 1.1. (Petrosyan et al., 2004) Consider stochastic game $G(z_0)$ defined on graph $\Psi(z_0)$ which is represented on Fig. 1.

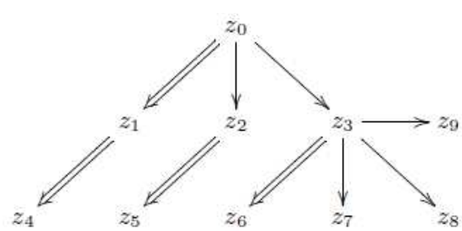


Fig. 1. Graph of Example 1.1.

The set of vertices of graph $\Psi(z_0)$ is $Z = \{z_0, \ldots, z_9\}$. The set of players is $N = \{1, 2\}$. In each vertex of graph $\Psi(z_0)$ two-player normal-form game $\Gamma(z)$,

$z \in Z$, is given and the payoffs in these games are the following:

$$\Gamma(z_0)\ :\begin{pmatrix}(5,5) & (0,8)\\ (8,0) & (1,1)\end{pmatrix},\qquad \Gamma(z_2)\ :\begin{pmatrix}(3,0) & (6,4)\\ (5,6) & (2,2)\end{pmatrix},$$
$$\Gamma(z_3):\begin{pmatrix}(1,11) & (4,2)\\ (1,3) & (1,1)\end{pmatrix},\qquad \Gamma(z_7):\begin{pmatrix}(1,1) & (0,2)\\ (2,0) & (1,2)\end{pmatrix},$$

$$\Gamma(z_8):\begin{pmatrix}(5,5) & (6,1)\\ (1,6) & (6,6)\end{pmatrix},\qquad \Gamma(z_9):\begin{pmatrix}(4,2) & (3,4)\\ (5,6) & (1,5)\end{pmatrix},$$
$$\Gamma(z_1),\Gamma(z_4),\Gamma(z_5),\Gamma(z_6)\ :\begin{pmatrix}(0,0) & (1,0)\\ (1,0) & (0,1)\end{pmatrix}.$$

To determine non-cooperative stochastic game $G(z_0)$ we need to define transition probabilities and probabilities of the game duration. First, define the transition probabilities from the vertices of the graph to the next vertices. If in game $\Gamma(z_0)$ the action profile $(2,2)$ is realised, then stochastic game $G(z_0)$ transits to the vertex z_2 with probability $1/3$ and to the vertex z_3 with probability $2/3$. If any other action profile different from $(2,2)$ is realised (arrow $\Longrightarrow$ means the deterministic transition), then the game $G(z_0)$ transits to vertex z_1. At vertices z_1, z_2 when any action profile is played, stochastic game $G(z_0)$ transits to vertices z_4 and z_5 respectively. If in the game $\Gamma(z_3)$ the action profile $(2,2)$ is played, then stochastic game $G(z_0)$ transits to vertices z_8 and z_9 with equal probabilities $1/2$. And if in the game $\Gamma(z_3)$ the action profile $(2,1)$ is played, the game $G(z_0)$ transits to vertex z_7 with probability 1. The deterministic transition (with probability 1) is made from other action profiles to vertex z_6 (arrow $\Longrightarrow$ means the deterministic transition).

Let probabilities q_k that the game ends at stage k be given:

$$q_1 = \frac{1}{8},\quad q_2 = 0,\quad q_3 = 1.$$

Let players choose the Shapley value as the cooperative solution of the game. For two-player game, it is calculated by formulas:

$$Sh_1(z) = V(\{1\},z) + \frac{V(\{1,2\},z) - V(\{1\},z) - V(\{2\},z)}{2},$$

$$Sh_2(z) = V(\{2\},z) + \frac{V(\{1,2\},z) - V(\{1\},z) - V(\{2\},z)}{2},$$

where $V(\{1\},z)$ and $V(\{2\},z)$ are the values of characteristic function for the subgame beginning at vertex z calculated for coalitions $\{1\}$ and $\{2\}$ respectively.

The above described sets and values determine stochastic game with random duration $G(z_0)$ (see (1)).

We start to find the solution of the cooperative game from the terminal vertices of the graph, i. e., the vertices from which it is impossible to transmit to any other vertices of the graph. First, calculate $V(\{1\},z_9)$ and $V(\{2\},z_9)$ as maximum guaranteed players' payoffs in the game $\Gamma(z_9)$ using formula (8):

$$V(\{1\},z_9) = 3,\quad V(\{2\},z_9) = 4,\quad V(\{1,2\},z_9) = 11.$$

Then, we may calculate the Shapley value of the subgame $\bar{G}(z_9)$ of the game $\bar{G}(z_0)$ starting from game $\Gamma(z_9)$:

$$Sh_1(z_9) = 5, \quad Sh_2(z_9) = 6.$$

We make the similar calculations for the subgames starting from the games $\Gamma(z_4)$, $\Gamma(z_5)$, $\Gamma(z_6)$, $\Gamma(z_7)$ and $\Gamma(z_8)$ using formula (8) while these games are realised at the vertices belonging to the set $\{z : L(z) = \varnothing\}$. The values of characteristic functions for these subgames and corresponding Shapley values are given in the Table 2.

Table 2. Characteristic functions and the Shapley values of subgames $\bar{G}(z)$, $z \in \{z_4, z_5, z_6, z_7, z_8, z_9\}$.

Vertex z	$V(\{1\}, z)$	$V(\{2\}, z)$	$V(\{1,2\}, z)$	$Sh_1(z)$	$Sh_2(z)$
z_4	0	0	1	1/2	1/2
z_5	0	0	1	1/2	1/2
z_6	0	0	1	1/2	1/2
z_7	1	2	3	1	2
z_8	5	5	12	6	6
z_9	3	4	11	5	6

Now consider the vertices from the set $\{z : (L(z))^2 = \varnothing\}$. We start from vertex z_3. As stochastic game may transit to the other vertices of the graph, we need to transform the payoff matrix of the game to calculate the Shapley value of cooperative subgame $\bar{G}(z_3)$. With action profile (2,2) the mathematical expectations of the players' payoffs we find in the following way:

- for Player 1:

$$1 + (1 - q_2)\left(\frac{1}{2}V(\{1\}, z_8) + \frac{1}{2}V(\{1\}, z_9)\right) = 5,$$

- for Player 2:

$$1 + (1 - q_2)\left(\frac{1}{2}V(\{2\}, z_8) + \frac{1}{2}V(\{2\}, z_9)\right) = 5.5.$$

With action profile (2,1) they are

- for Player 1:

$$1 + (1 - q_2)V(\{1\}, z_7) = 2,$$

- for Player 2:

$$1 + (1 - q_2)V(\{2\}, z_7) = 3.$$

Similarly, with action profile (1,1) the mathematical expectations of the players' payoffs are

- for Player 1:

$$1 + (1 - q_2)V(\{1\}, z_6) = 1,$$

- for Player 2:

$$11 + (1 - q_2)V(\{2\}, z_6) = 11;$$

and with action profile (1,2) the mathematical expectations of the players' payoffs are

- for Player 1:
$$4+(1-q_2)V(\{1\},z_6)=4,$$
- for Player 2:
$$2+(1-q_2)V(\{2\},z_6)=2.$$

Then the bi-matrix game written for the calculations of the values of characteristic functions $V(\{1\},z_3)$ and $V(\{2\},z_3)$ looks like

$$\begin{pmatrix}(1,11) & (4,2)\\ (2,5) & (5,5.5)\end{pmatrix}.$$

The values of characteristic function of cooperative subgame $\bar{G}(z_3)$ of the game $\bar{G}(z_0)$ for coalitions $\{1\}$, $\{2\}$ are

$$V(\{1\},z_3)=2,\quad V(\{2\},z_3)=5.$$

To calculate $V(\{1,2\},z_3)$ we use formula (7) and obtain the bi-matrix game:

$$\begin{pmatrix}12+(1-q_2)V(\{1,2\},z_6) & 6+(1-q_2)V(\{1,2\},z_6)\\ 4+(1-q_2)V(\{1,2\},z_7) & 2+(1-q_2)(0.5V(\{1,2\},z_8)+0.5V(\{1,2\},z_9))\end{pmatrix}$$

or in numeric form:

$$\begin{pmatrix}13 & 7\\ 7 & 13.5\end{pmatrix}.$$

Therefore,

$$V(\{1,2\},z_3)=13,5,$$
$$Sh_1(z_3)=5.25,\quad Sh_2(z_3)=8.25.$$

We make similar calculations for the cooperative subgame $\bar{G}(z_1)$:

$$V(\{1\},z_1)=0,\quad V(\{2\},z_1)=0,\quad V(\{1,2\},z_1)=2,$$
$$Sh_1(z_1)=Sh_2(z_1)=1,$$

and for subgame $\bar{G}(z_2)$:

$$V(\{1\},z_2)=3,\quad V(\{2\},z_2)=2,\quad V(\{1,2\},z_2)=12,$$
$$Sh_1(z_2)=6.5,\quad Sh_2(z_2)=5.5.$$

For cooperative stochastic game $\bar{G}(z_0)$, the matrix game for the calculation of the values of characteristic function for coalitions $\{1\}$, $\{2\}$ can be found by formula (7). With action profile (2,2) the mathematical expectations of the players' payoffs are

- for Player 1:
$$1+(1-q_1)\left(\frac{1}{3}V(\{1\},z_2)+\frac{2}{3}V(\{1\},z_3)\right)=3\frac{1}{24},$$

- for Player 2:

$$1 + (1 - q_1)\left(\frac{1}{3}V(\{2\}, z_2) + \frac{2}{3}V(\{2\}, z_3)\right) = 4\frac{1}{2}.$$

With action profile (2,1) the mathematical expectations of the players' payoffs are

- for Player 1:

$$8 + (1 - q_1)V(\{1\}, z_1) = 8,$$

- for Player 2:

$$0 + (1 - q_1)V(\{2\}, z_1) = 0.$$

Similarly, with action profile (1,1) the mathematical expectations of the players' payoffs are

- for Player 1:

$$5 + (1 - q_1)V(\{1\}, z_1) = 5,$$

- for Player 2:

$$5 + (1 - q_1)V(\{2\}, z_1) = 5.$$

With action profile (1,2) the mathematical expectations of the players' payoffs are

- for Player 1:

$$0 + (1 - q_1)V(\{1\}, z_1) = 0,$$

- for Player 2:

$$8 + (1 - q_1)V(\{2\}, z_1) = 8.$$

Finally, we obtain the matrix:

$$\begin{pmatrix} (5,5) & (0,8) \\ (8,0) & (3\frac{1}{24}, 4\frac{1}{2}) \end{pmatrix},$$

$$V(\{1\}, z_0) = 3\frac{1}{24}, V(\{2\}, z_0) = 4\frac{1}{2}.$$

For the calculation of $V(\{1,2\}, z_0)$ we form matrix game using formula (7):

$$\begin{pmatrix} 10 + (1 - q_1)V(\{1,2\}, z_1) & 8 + (1 - q_1)V(\{1,2\}, z_1) \\ 8 + (1 - q_1)V(\{1,2\}, z_1) & 2 + (1 - q_1)(\frac{1}{3}V(\{1,2\}, z_2) + \frac{2}{3}V(\{1,2\}, z_3)) \end{pmatrix}$$

or in a numeric form:

$$\begin{pmatrix} 11\frac{3}{4} & 9\frac{3}{4} \\ 9\frac{3}{4} & 13\frac{3}{8} \end{pmatrix}.$$

Calculating $V(\{1,2\}, z_0)$ and $Sh_1(z_0)$, $Sh_2(z_0)$, we obtain:

$$V(\{1,2\}, z_0) = 13\frac{3}{8}, \qquad Sh_1(z_0) = 5\frac{23}{24}, \qquad Sh_2(z_0) = 7\frac{5}{12}.$$

The set of vertices forming the cooperative subtree consists of the vertices z_0, z_2, z_3, z_5, z_8, z_9.

Now we verify if the imputation distribution procedure is non-negative. It is negative at vertex z_3 that follows from equation (17), in which the vertex z_3 is used:

$$Sh_1(z_3) = \beta_1(z_3) + (1 - q_1)\left[\frac{1}{2} \cdot Sh_1(z_8) + \frac{1}{2} \cdot Sh_1(z_9)\right],$$
$$5.25 = \beta_1(z_3) + (1 - 0)\left[\frac{1}{2} \cdot 5 + \frac{1}{2} \cdot 6\right],$$
$$\beta_1(z_3) = -0.25.$$

As $\beta_1(z_3)$ is negative, we make the regularization of the Shapley value to construct a «new» non-negative Shapley value.

Determine the new Shapley value for the vertices of the cooperative subtree with vertices z_0, z_2, z_3, z_5, z_8, z_9 by formulas (26) and (27):

$$\hat{Sh}_1(z_5) = 0.5, \hat{Sh}_1(z_8) = 6, \hat{Sh}_1(z_9) = 5,$$
$$\hat{Sh}_2(z_5) = 0.5, \hat{Sh}_2(z_8) = 6, \hat{Sh}_2(z_9) = 6,$$
$$\hat{Sh}_1(z_2) = \frac{11}{12} \cdot 6.5 + \frac{1}{2} = 6\frac{11}{24},$$
$$\hat{Sh}_2(z_2) = \frac{11}{12} \cdot 5.5 + \frac{1}{2} = 5\frac{13}{24},$$
$$\hat{Sh}_1(z_3) = \frac{2}{13.5} \cdot 5.25 + \left[\frac{1}{2} \cdot 5 + \frac{1}{2} \cdot 6\right] = 6\frac{5}{18},$$
$$\hat{Sh}_2(z_3) = \frac{2}{13.5} \cdot 8.25 + \left[\frac{1}{2} \cdot 6 + \frac{1}{2} \cdot 6\right] = 7\frac{2}{9},$$
$$\hat{Sh}_1(z_0) = \frac{2}{13\frac{3}{8}} \cdot 5\frac{23}{24} + \left(1 - \frac{1}{8}\right)\left[\frac{1}{3} \cdot 6\frac{11}{24} + \frac{2}{3} \cdot 6\frac{5}{18}\right] = 6\frac{80741}{184896} \approx 6.437,$$
$$\hat{Sh}_2(z_0) = \frac{2}{13\frac{3}{8}} \cdot 7\frac{5}{12} + \left(1 - \frac{1}{8}\right)\left[\frac{1}{3} \cdot 5\frac{13}{24} + \frac{2}{3} \cdot 7\frac{2}{9}\right] = 6\frac{173491}{184896} \approx 6.938.$$

The «new» vector is the Shapley value of the cooperative game with characteristic function defined by formulas (23), (24). It is subgame-consistent which follows from Theorem 4.

For the games $\Gamma(z_5)$, $\Gamma(z_8)$ and $\Gamma(z_9)$ the new characteristic functions are presented in Table 3.

Table 3. «New» characteristic functions.

Vertex z	$\hat{V}(\{1\}, z)$	$\hat{V}(\{2\}, z)$	$\hat{V}(\{1,2\}, z)$
z_0	3.763	4.265	13.375
z_2	2.750	1.833	12.000
z_3	4.296	5.574	13.500
z_5	0.000	0.000	1.000
z_8	5.000	5.000	12.000
z_9	3.000	4.000	11.000

Remark 4. The nucleolus may be chosen by the players as a solution of the cooperative game (see Schmeidler, 1969). Notice that the nucleolus consists of one

vector, so there are no problems with the choice of a unique imputation from the imputation set. We also notice that the nucleolus belongs to the core when the latter is non-empty.

Example 1.2. Consider stochastic game $G(z_0)$ defined on the graph $\Psi(z_0)$ which is presented on Fig. 3. The set of vertices of graph $\Psi(z_0)$ is $Z = \{z_0, \ldots, z_5\}$. The set

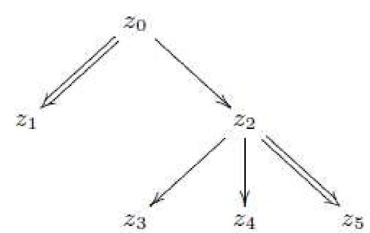


Fig. 2. Graph of Example 1.2.

of players is $N = \{1, 2, 3\}$. At each vertex of graph $G(z_0)$ three-player normal-form game $\Gamma(z)$, $z \in Z$, is given. The payoff matrix are the following:

$$\Gamma(z_0), \Gamma(z_2) : \left(\begin{pmatrix} (1,1,1) & (2,2,0) \\ (2,2,0) & (0,0,3) \end{pmatrix} \begin{pmatrix} (1,1,2) & (2,2,1) \\ (1,3,1) & (3,0,1) \end{pmatrix} \right),$$

$$\Gamma(z_1), \Gamma(z_5) : \left(\begin{pmatrix} (2,0,1) & (1,0,1) \\ (3,1,2) & (2,2,2) \end{pmatrix} \begin{pmatrix} (2,2,1) & (1,1,3) \\ (2,1,1) & (2,1,2) \end{pmatrix} \right),$$

$$\Gamma(z_3) : \left(\begin{pmatrix} (1,1,1) & (2,2,2) \\ (3,2,0) & (1,4,1) \end{pmatrix} \begin{pmatrix} (2,0,1) & (2,1,1) \\ (4,0,1) & (0,4,1) \end{pmatrix} \right),$$

$$\Gamma(z_4) : \left(\begin{pmatrix} (2,1,0) & (2,1,3) \\ (3,1,2) & (3,6,4) \end{pmatrix} \begin{pmatrix} (4,5,0) & (0,5,4) \\ (2,8,0) & (0,8,2) \end{pmatrix} \right).$$

In each game defined above, Player 1 chooses rows, Player 2 chooses columns and Player 3 chooses matrices.

First, we define the transition probabilities from the vetrices to the other vertices of the graph. If in game $\Gamma(z_0)$ action profile (1,1,1) is played, then stochastic game $G(z_0)$ transits to the vertex z_1 with probability 1/3 and to the vertex z_2 with probability 2/3. Otherwise, if any action profile different from (1,1,1) is played (arrow $\Longrightarrow$ means the deterministic transition), then the game $G(z_0)$ transits to the vertex z_1. If action profile (2,1,2) is realised at vertex z_2, stochastic game $G(z_0)$ transits to the vertex z_3 and z_4 with probabilities 1/3, 2/3 respectively. If any other action profile different from (2,1,2) is realised, game $G(z_0)$ transits to vertex z_5 with probability 1.

The probabilities q_k that stochastic game $G(z_0)$ ends at stage k are given:

$$q_1 = 0.5, \quad q_2 = 0, \quad q_3 = 1.$$

Let players choose the Shapley value as a solution of the game. We start solving the game with the vertices of the graph which belong to the set $\{z : L(z) = \varnothing\}$. We calculate the values of characteristic function and the Shapley value for subgame

$\bar{G}(z_3)$. Similar calculations are made for the vertices z_1, z_5, z_4, and then for the vertices z_2 and z_0 using formula (7). The calculations are presented in Tables 4 and 5.

Table 4. Characteristic functions for subgames $\bar{G}(z)$, $z \in \{z_0, z_1, z_2, z_3, z_4, z_5\}$.

z	$V(\{1\},z)$	$V(\{2\},z)$	$V(\{3\},z)$	$V(\{1,2\},z)$	$V(\{2,3\},z)$	$V(\{1,3\},z)$	$V(\{1,2,3\},z)$
z_0	2	1	3/2	11/2	9/2	2	83/9
z_1	2	0	1	3	4	3	6
z_2	3	1	4/3	7	6	7	47/3
z_3	1	1	1	4	4	3	6
z_4	0	1	0	8	9	5	13
z_5	2	0	1	3	4	3	6

Table 5. The Shapley values of subgames $\bar{G}(z)$, $z \in \{z_0, z_1, z_2, z_3, z_4, z_5\}$.

z	$Sh_1(z)$	$Sh_2(z)$	$Sh_3(z)$
z_0	193/54	305/108	305/108
z_1	8/3	7/6	13/6
z_2	37/6	14/3	29/6
z_3	11/6	7/3	11/6
z_4	10/3	35/6	23/6
z_5	8/3	7/6	13/6

The set of vertices of the cooperative subtree is $CZ = \{z_0, z_1, z_2, z_3, z_4\}$. We regularize the Shapley value:

$$Sh(z_0) = \left(3\frac{31}{54}, 2\frac{89}{108}, 2\frac{89}{108}\right)$$

and verify if the imputation distribution procedure is non-negative. For this, we find values $\beta_i(z)$ for vertices $z_0 \in CZ$ and $z_2 \in CZ$ using formula (19) and verify if imputation distribution procedure $\beta_i(z)$ is non-negative:

$$\beta_i(z_2) = Sh_i(z_2) - (1 - q_2)\left(\frac{1}{3}Sh_i(z_3) + \frac{2}{3}Sh_i(z_4)\right),$$

$$\beta_i(z_0) = Sh_i(z_0) - (1 - q_1)\left(\frac{1}{3}Sh_i(z_1) + \frac{2}{3}Sh_i(z_2)\right),$$

obtaining

$$\beta_1(z_2) = 3\frac{1}{3}, \qquad \beta_2(z_2) = 0, \qquad \beta_3(z_2) = 1\frac{2}{3};$$

$$\beta_1(z_0) = 1\frac{2}{27}, \qquad \beta_2(z_0) = 1\frac{2}{27}, \qquad \beta_3(z_0) = \frac{23}{27}.$$

For $z \in \{z_0, z_2\}$ the following conditions: $\beta_i(z) \geqslant 0$ and $\sum_{i \in N} \beta_i(z_2) = 5$ are satisfied, and $\sum_{i \in N} \beta_i(z_0) = 3$.

In all vertices of the cooperative subtree, conditions of subgame consistency and non-negativity of the Shapley value are satisfied. Therefore, we state that the Shapley value is subgame-consistent imputation in game $\bar{G}(z_0)$.

Now we repeat calculations assuming that players adopt the nucleolus as a solution of the game $\bar{G}(z_0)$. The nucleolus was initially proposed by D. Schmeidler (Schmeidler, 1969). The definition and some usefull theorems and lemmas about the properties of the nucleolus may be found in (Pecherski and Yanovskaya, 2004, Driessen et al., 1992, Kohlberg, 1971). The works (Kohlberg, 1972, Montero, 2005) are devoted to the calculation of the nucleolus which contains the unique vector. For the calculation of the nucleolus, one may use Matlab (Mathworks, 2017) and program TUGlab (TUGlab), written for calculation in cooperative game theory, or Mathematica (MATHEMATICA) and program TUGames (Meinhardt) written for the same tasks.

The characteristic function was calculated above. The nucleolus of the subgame $\bar{G}(z)$, $z \in CZ$, is denoted by $n(z) = (n_i(z) : i \in N)$.

We calculate the nucleoli for all subgames of the game $\bar{G}(z_0)$. The results are presented in Table 6.

Table 6. The nucleoli of subgames $\bar{G}(z)$, $z \in \{z_0, z_1, z_2, z_3, z_4, z_5\}$.

z	$n_1(z)$	$n_2(z)$	$n_3(z)$
z_0	$3\frac{5}{9}$	$3\frac{1}{18}$	$2\frac{11}{18}$
z_1	$2\frac{1}{2}$	$1\frac{1}{4}$	$2\frac{1}{4}$
z_2	$6\frac{1}{3}$	$4\frac{1}{2}$	$4\frac{2}{3}$
z_3	$1\frac{2}{3}$	$2\frac{2}{3}$	$1\frac{2}{3}$
z_4	$2\frac{2}{3}$	$6\frac{2}{3}$	$3\frac{2}{3}$
z_5	$2\frac{1}{2}$	$1\frac{1}{4}$	$2\frac{1}{4}$

Now we verify the subgame consistency of the nucleolus using formula (17) and calculate $\beta_i(z_2)$ for vertex z_2 by formula:

$$n_i(z_2) = \beta_i(z_2) + (1 - q_2)(p(z_3|z_2, \bar{a}^{z_2})n_i(z_3) + p(z_4|z_2, \bar{a}^{z_2})n_i(z_4)).$$

We obtain:

$$\beta_1(z_2) = 4, \quad \beta_2(z_2) = -\frac{5}{6}, \quad \beta_3(z_2) = 1\frac{5}{6}.$$

The nucleolus of the cooperative stochastic game $\bar{G}(z_0)$ is not subgame-consistent if the non-negativity of the imputation distribution procedure is required. For example, $\beta_2(z_2) < 0$. We won't verify existence of non-negative imputation distribution procedure (17), but we will make the regularization of the nucleolus.

Calculate «new» nucleolus for each vertex $z \in CZ$ by formula (21) with initial condition (22). For vertex z_2, we use the following formula:

$$\hat{n}_i(z_2) = \frac{\sum\limits_{i \in N} K_i(\bar{a}^{z_2})}{V(N, z_2)} n_i(z_2) + (1 - q_2) \sum_{y \in L(z_2)} p(y|z_2, \bar{a}^{z_2})\hat{n}_i(y),$$

for vertex z_0, we use formula:

$$\hat{n}_i(z_0) = \frac{\sum\limits_{i \in N} K_i(\bar{a}^{z_0})}{V(N, z_0)} N_i(z_0) + (1 - q_1) \sum_{y \in L(z_0)} p(y|z_0, \bar{a}^{z_0})\hat{n}_i(y).$$

Table 7. The nucleoli of subgames $\bar{G}(z)$, $z \in \{z_0, z_1, z_2, z_3, z_4, z_5\}$.

z	$\hat{n}_1(z)$	$\hat{n}_2(z)$	$\hat{n}_3(z)$
z_0	$3\frac{3487}{140436}$	$3\frac{128867}{280872}$	$2\frac{69149}{93624}$
z_1	$2\frac{1}{2}$	$1\frac{1}{4}$	$2\frac{1}{4}$
z_2	$4\frac{50}{141}$	$6\frac{217}{282}$	$4\frac{51}{94}$
z_3	$1\frac{2}{3}$	$2\frac{2}{3}$	$1\frac{2}{3}$
z_4	$2\frac{2}{3}$	$6\frac{2}{3}$	$3\frac{2}{3}$
z_5	$2\frac{1}{2}$	$1\frac{1}{4}$	$2\frac{1}{4}$

«New» nucleoli for the vertices of set CZ are given in Table 7.

Calculate characteristic function $\hat{V}(S, z)$ for each vertex $z \in CZ$ by formulas (23) and (24). Moreover, $\hat{V}(S, z_3) = V(S, z_3)$, $\hat{V}(S, z_4) = V(S, z_4)$, $\hat{V}(S, z_1) = \hat{V}(S, z_5) = V(S, z_1) = V(S, z_5)$. For the calculation of $\hat{V}(S, z_2)$ we use formula:

$$\hat{V}(S, z_2) = \frac{\sum\limits_{i \in N} K_i(\bar{a}^{z_2})}{V(N, z_2)} V(S, z_2) + (1 - q_2) \sum_{y \in L(z_2)} p(y|z_2, \bar{a}^{z_2}) \hat{V}(S, y),$$

and for $\hat{V}(S, z_0)$:

$$\hat{V}(S, z_0) = \frac{\sum\limits_{i \in N} K_i(\bar{a}^{z_0})}{V(N, z_0)} V(S, z_0) + \\ + (1 - q_1) \Big(p(z_1|z_0, \bar{a}^{z_0}) \hat{V}(S, z_1) + p(z_2|z_0, \bar{a}^{z_0}) \hat{V}(S, z_2) \Big).$$

The values of the function $\hat{V}(S, \cdot)$ are given in Table 8.

Table 8. Characteristic function $\hat{V}(S, z)$, $z \in \{z_0, z_1, z_2, z_3, z_4, z_5\}$.

z	$\hat{V}(\{1\}, z)$	$\hat{V}(\{2\}, z)$	$\hat{V}(\{3\}, z)$	$\hat{V}(\{1,2\}, z)$	$\hat{V}(\{2,3\}, z)$	$\hat{V}(\{1,3\}, z)$	$\hat{V}(\{1,2,3\}, z)$
z_0	$1\frac{245}{249}$	$\frac{164}{249}$	$1\frac{74}{747}$	$4\frac{155}{249}$	$3\frac{80}{83}$	$4\frac{52}{83}$	$9\frac{2}{9}$
z_1	2	0	1	3	4	3	6
z_2	$1\frac{41}{141}$	$1\frac{15}{47}$	$\frac{107}{141}$	$8\frac{127}{141}$	$9\frac{35}{141}$	$6\frac{80}{141}$	$15\frac{2}{3}$
z_3	1	1	1	4	4	3	6
z_4	0	1	0	8	9	5	13
z_5	2	0	1	3	4	3	6

Notice that the "new" nucleolus $\hat{n}(z_2)$ of subgame $\bar{G}(z_2)$ belongs to the imputation set with characteristic function $\hat{V}(S, z_2)$ (the nucleolus $\hat{n}(z_2)$ also belongs to the set $I(z_2)$, which is not true in general), but it is not the nucleolus of the cooperative game. The nucleolus of cooperative game defined by characteristic function $\hat{V}(S, z_2)$, is denoted by $\tilde{n}(z_2) = (\tilde{n}_1(z_2), \tilde{n}_2(z_2), \tilde{n}_3(z_2))$. It equals to the following one:

$$\tilde{n}(z_2) \approx (4.213, 6.894, 4.560) \neq \hat{n}(z_2).$$

The «new» nucleolus $\hat{n}(z_0)$ calculated for the game $\bar{G}(z_0)$, belongs to the imputation set of the cooperative game defined by characteristic function $\hat{V}(S, z_0)$ ($\hat{n}(z_0)$ also belongs to the imputation set $I(z_0)$), but it does not coincide with the nucleolus of this cooperative game. The nucleolus of the cooperative game defined by characteristic function $\hat{V}(S, z_0)$, given above, is denoted by $\tilde{n}(z_0)$ and it equals

$$\tilde{n}(z_0) \approx (3.621, 2.720, 2.881) \neq \hat{n}(z_0).$$

2.9. Strongly subgame consistency of the core

In this section we consider the case when solution of the cooperative stochastic game is the set and contains more than one point. As an example of such a solution we examine the core. First, we describe the problem of subgame consistency and then find the sufficient conditions of strongly subgame consistency of the core. This problem was initially examined by Leon Petrosyan for differential games (Petrosyan, 1992) and then for multicriteria problems of optimal control (Petrosyan, 1993).

Suppose that the cores of stochastic game $\bar{G}(z_0)$ and any subgame $\bar{G}(z)$, $z \in CZ$, are non-empty. When players cooperate they come to an agreement about the realization of the cooperative strategy profile $\bar{\varphi}$ and expect to receive the components of the imputation belonging to the core $CO(z_0)$. Reaching the intermediate vertex $z \in CZ \setminus \{z_0\}$ of the cooperative subtree, player $i \in N$ chooses an action $\bar{a}_i^z$ in accordance with the cooperative strategy $\bar{\varphi}_i$ and receives the payoff $K_i^z(\bar{a}^z)$. If the players recalculate the cooperative solution, i.e., find the solution of the cooperative subgame starting from vertex z, the recalculated solution will be the core $CO(z)$. It will be rational to require that the payoff received by the player in vertex z summarized with the expected sum of any imputations from solutions $CO(y)$, $y \in L(z)$, of the games of the cooperative subtrees following game $\Gamma(z)$, is equal to the imputation from solution $CO(z)$. If this property is satisfied for any vertex z of the cooperative subtree, the core of cooperative stochastic game $\bar{G}(z_0)$ is strongly subgame-consistent.

To introduce the mathematically strict definition of strongly subgame-consistent core, it is necessary to define the so-called *expected core*. For any non-terminal vertex of the cooperative subtree we define the set of expected imputations belonging to the cores which are the solutions of the subgames following the considered vertex. For any vertex $z \in CZ$, $L(z) \neq \varnothing$, define the *expected core*:

$$EC(L(z)) = \left\{ \alpha(L(z)) = \sum_{y \in L(z)} p(y|z, \bar{a}^z)\alpha(y) \mid \alpha(y) \in CO(y) \right\}. \tag{33}$$

The set $EC(L(z))$ consists of the vectors $\alpha(L(z))$ which are the mathematical expectations of the possible collection of the imputations from the cores of the subgames beginning from the vertices following vertex z with respect to the probability distribution $\{p(y|z, \bar{a}^z), y \in L(z)\}$.

We also define the distribution procedure of the players' payoffs in the vertices of the cooperative subtree. Refine Definition 10 of the imputation distribution procedure. The first condition in Definition 10 maybe called the condition of "feasibility of the imputation distribution procedure" because it guarantees that in any vertex of the cooperative subtree the sum of the payments to the players equals the sum of the payoffs received by the players when they realize cooperative strategies. The second condition guarantees to the players that they receive the components of the initially chosen imputation from the core of cooperative game $\bar{G}(z_0)$ in the sense of mathematical expectation, if the payments to the players along the game are realized in accordance with imputation distribution procedure $\{\beta(z) : z \in CZ\}$.

Now we need to define the distribution procedure of the imputation $\alpha(z_0)$ from the core $CO(z_0)$ in a way that the core is strongly subgame-consistent.

Definition 16. We call the core $CO(z_0)$ of the cooperative stochastic game $\bar{G}(z_0)$ strongly subgame-consistent if there exists the distribution procedure $\{\beta(z)\}_{z\in CZ}$ of the imputation from the core $CO(z_0)$ such that for each vertex $z \in CZ$ the inclusions take place:

$$\beta(z) \oplus (1-q_k)EC(L(z)) \subset CO(z), \tag{34}$$

$$B(z_0) \in CO(z_0), \tag{35}$$

where

$$\beta(z) \oplus (1-q_k)EC(L(z)) = \Big\{\beta(z) + (1-q_k)\alpha(L(z)) : \alpha(L(z)) \in EC(L(z))\Big\}.$$

And the imputation distribution procedure $\{\beta(z)\}_{z\in CZ}$ is called strongly subgame-consistent. [3]

Condition (34) means that the set of vectors which are equal to the sum of the imputation distribution procedure of the player at vertex z and the imputation from the expected core of the vertex z, belongs to the core of the subgame beginning from vertex z. This condition provides the restrictions on the payments to the players in the games defined at vertices and often it is not satisfied for any game if the payments to the players are realised in accordance with initially defined payoff functions.

We impose additional restrictions on characteristic functions of subgames starting from the vertices of the cooperative subtree to obtain sufficient conditions of strongly subgame consistency of the core. Denote by $EV(S, L(z))$ the expected values of characteristic function calculated for coalition $S \subseteq N$ at the vertices following the vertex z:

$$EV(S, L(z)) = \sum_{y\in L(z)} p(y|z, \bar{a}^z)V(S, y).$$

Denote by

$$\Delta V(S, z) = V(S, z) - (1-q_k)EV(S, L(z))$$

the difference between the value of characteristic function at vertex z and expected value of characteristic function on condition that the game does not finish at vertex z. Denote by $\Delta CO(z)$ analogue of the core calculated using function $\Delta V(S, z)$. Now define sufficient condition of strongly subgame consistency of the imputation distribution procedure and the core $CO(z_0)$.

Theorem 6. *Let for each vertex $z \in CZ$ the core $CO(z)$ and the set $\Delta CO(z)$ be non-empty. For each vertex $z \in CZ$ distribution procedure $\{\beta(z) : z \in CZ\}$ of the imputation from the core $CO(z_0)$ satisfies the conditions:*

$$\beta(z) \in \Delta CO(z), \tag{36}$$

$$B(z_0) \in CO(z_0). \tag{37}$$

then the core $CO(z_0)$ and distribution procedure $\{\beta(z) : z \in CZ\}$ are strongly subgame-consistent.

[3] The sum denoted by sign $\oplus$ is called Minkowski sum (see (Schneider), in which some properties of this operator are proved).

Proof. We need to prove that any vector $\beta(z) \in \Delta CO(z)$ satisfying conditions (36) and (37) is strongly subgame-consistent distribution procedure of the imputation $\alpha(z_0) \in CO(z_0)$. So, the conditions (34) and (35) from Definition 16 hold. Condition (37) coincides with (35), therefore, it remains to show that the inclusion (34) holds for any vertex $z \in CZ$. Consider any vector $\alpha(L(z)) \in EC(L(z))$ for vertex z and calculate the sum $\beta(z) + (1 - q_k)\alpha(L(z))$. Verify if the latter vector belongs to the core $CO(z)$. Now calculate the sum of all components of the vector:

$$\sum_{i \in N} \beta_i(z) + (1 - q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) \sum_{i \in N} \alpha_i(y) =$$

$$= V(N, z) - (1 - q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) V(N, y) +$$
$$+ (1 - q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) \sum_{i \in N} \alpha_i(y) = V(N, z),$$

which carries out the property of collective rationality.

Now consider $S \subset N$, $S \neq N$:

$$\sum_{i \in S} \beta_i(z) + (1 - q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) \sum_{i \in S} \alpha_i(y) \geqslant$$
$$\geqslant V(S, z) + (1 - q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) V(S, y) -$$
$$- (1 - q_k) \sum_{y \in L(z)} p(y|z, \bar{a}^z) V(S, y) = V(S, z).$$

By virtue of the arbitrariness of vertex $z \in CZ$, we make a conclusion that the core of cooperative game $\bar{G}(z_0)$ and procedure $\{\beta(z) : z \in CZ\}$ are strongly subgame-consistent.

When analogue of the core $\Delta CO(z)$ is non-empty for each vertex z of the cooperative subtree, Theorem 6 provides the method of construction of strongly subgame-consistent distribution procedure of the imputations from the core, equal $B_i(z_0)$ by condition (37). Notice that in a general case not all the imputations from the core can be realised using distribution procedure $\{\beta(z) : z \in CZ\}$ defined above.

Example 1.3 Consider stochastic game $G(z_0)$ defined on graph $\Psi(z_0)$ depicted on Fig. 3.

The set of the vertices of graph $\Psi(z_0)$ is $Z = \{z_0, \dots, z_5\}$. The set of the players is $N = \{1, 2, 3\}$. In each vertex of graph $G(z_0)$ the three-person normal-form game

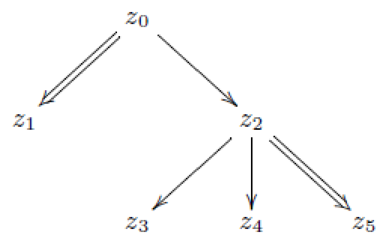


Fig. 3. The tree $\Psi(z_0)$.

$\Gamma(z)$, $z \in Z$, is defined. The payoff matrices are

$$\Gamma(z_0) : \left(\begin{pmatrix} (2,2,2) & (2,2,0) \\ (2,2,0) & (0,0,3) \end{pmatrix} \begin{pmatrix} (1,1,2) & (2,2,1) \\ (1,3,1) & (3,0,1) \end{pmatrix} \right),$$

$$\Gamma(z_2) : \left(\begin{pmatrix} (1,1,1) & (2,2,0) \\ (2,2,0) & (0,0,3) \end{pmatrix} \begin{pmatrix} (1,1,2) & (2,2,1) \\ (1,3,1) & (3,0,1) \end{pmatrix} \right),$$

$$\Gamma(z_1), \Gamma(z_5) : \left(\begin{pmatrix} (2,0,1) & (1,0,1) \\ (3,1,2) & (2,2,2) \end{pmatrix} \begin{pmatrix} (2,2,1) & (1,1,3) \\ (2,1,1) & (2,1,2) \end{pmatrix} \right),$$

$$\Gamma(z_3) : \left(\begin{pmatrix} (1,1,1) & (2,2,2) \\ (3,2,0) & (1,4,1) \end{pmatrix} \begin{pmatrix} (2,0,1) & (2,1,1) \\ (4,0,1) & (0,4,1) \end{pmatrix} \right),$$

$$\Gamma(z_4) : \left(\begin{pmatrix} (2,1,0) & (2,1,3) \\ (3,1,2) & (3,6,4) \end{pmatrix} \begin{pmatrix} (4,5,0) & (0,5,4) \\ (2,8,0) & (0,8,2) \end{pmatrix} \right).$$

In each game the first player chooses rows, the second one chooses columns, the third one chooses matrices. The strategy set of player $i \in N$ in game $\Gamma(z)$ is $A_i^z = \{1, 2\}$.

Define the probabilities of transition from all vertices to the following ones. If in game $\Gamma(z_0)$ the action profile $(1,1,1)$ is realised, stochastic game $G(z_0)$ transits to vertex z_1 with a probability of $1/3$ and to vertex z_2 with a probability of $2/3$. If any action profile different from (1,1,1) is realised (arrow $\Longrightarrow$ means the deterministic transition), the game $G(z_0)$ transits to vertex z_1. If at vertex z_2 action profile (2,1,2) is realised, stochastic game $G(z_0)$ transits to vertices z_3 and z_4 with probabilities of $1/3$, $2/3$ respectively. The game $G(z_0)$ transits to vertex z_5 with a probability of 1 from any other vertices.

The probabilities q_k that stochastic game $G(z_0)$ ends at stage k are given:

$$q_1 = 0.5, \quad q_2 = 0, \quad q_3 = 1.$$

To construct the cooperative version of stochastic game we find the cooperative strategy profile $\bar{\varphi}$. This profile $\bar{\varphi}$ prescribes to play action profile $(1,1,1)$ at vertex z_0. The game ends at stage z_0 with probability 0.5 and transits to the next stage with a probability of 0.5. If the game does not end, it transits to stage z_1 with a probability of $1/3$, at which the players should realise any of action profiles $(2,1,1)$ or $(2,2,1)$, or with a probability of $2/3$ the game transits to vertex z_2, at which the players should play action profile $(2,1,2)$. At vertex z_2 the game does not end because $q_1 = 0$ and transits to the vertices z_3 and z_4 with probabilities of $1/3$ and $2/3$ respectively. At vertices z_3 and z_4 the game terminates. Therefore, the set of the vertices of the cooperative subtree represented on Fig. 4 is $\bar{\Psi}(z_0) = \{z_0, z_1, z_2, z_3, z_4\}$.

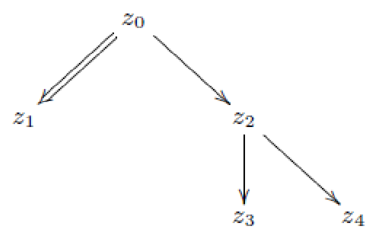


Fig. 4. Cooperative subtree $\bar{\Psi}(z_0)$ of the game $\bar{G}(z_0)$.

Find the values of characteristic function using formulas (5) with boundary condition (6) for $S = N$, (7) with boundary condition (8) for $S \subset N$ and (9) for $S = \varnothing$. Calculations are given in Table 9. For further calculations we use package `TUGlab` of program Matlab [16].

Table 9. Characteristic functions $v(S, z)$ for $\bar{G}(z)$, $z \in \{z_0, z_1, z_2, z_3, z_4, z_5\}$.

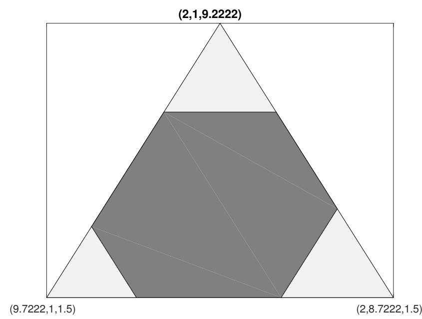
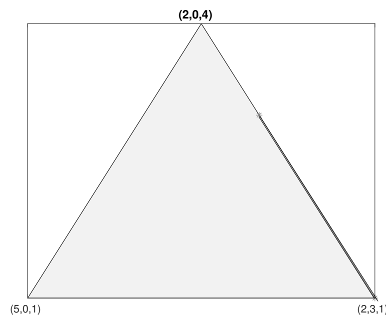
$z \setminus S$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{2,3\}$	$\{1,3\}$	$\{1,2,3\}$
z_0	2	1	1.5	5.5	4.5	6	110/9
z_1	2	0	1	3	4	3	6
z_2	3	1	4/3	7	6	7	47/3
z_3	1	1	1	4	4	3	6
z_4	0	1	0	8	9	5	13
z_5	2	0	1	3	4	3	6

Now we define the cores of subgames beginning from the vertices of cooperative subtree $\bar{\Psi}(z_0)$. We also assure that they all are non-empty to use the core as a cooperative solution of a stochastic game. The systems of linear inequalities and equities which determines the cores and their graphical representations are given in Tables 10 and 11. On the figures, the imputation set is depicted as a light-gray triangle and the cores are dark-grey sets. Notice that at vertices z_1 and z_5 condition $\alpha_1 = 2$ holds for each element of the core. And the core is the segment connecting points $(2, 1, 3)$ and $(2, 3, 1)$.

For each vertex of the cooperative subtree $\bar{\Psi}(z_0)$ we define the analogues of the cores denoted by $\Delta CO(z)$. Remind that for terminal vertices z_1, z_3, z_4 of set $\Delta CO(\cdot)$ coincide with the core $CO(\cdot)$. Systems of linear inequalities and equities determining sets $\Delta CO(z_0)$ and $\Delta CO(z_2)$ and also their graphs are presented in Tables 12. Notice that analogues of the cores $\Delta CO(\cdot)$ are non-empty for all vertices of the cooperative subtree. First, verify if the core is strongly subgame-consistent if the payments to the players are realised according to initially defined payoff functions, i. e., verify if payoff vectors in the vertices of the cooperative subtree belong to the corresponding sets $\Delta CO(\cdot)$ when the players realise cooperative strategy profile:

$$K^{z_0}(1,1,1) = (2,2,2) \in \Delta CO(z_0),$$
$$K^{z_1}(2,2,1) = (2,2,2) \in CO(z_1) = \Delta CO(z_1),$$

Table 10. The core for vertices $z_0, z_1, z_5 \in CZ$.

z	Core	Graph of the core
z_0	$\begin{cases} \alpha_1 \geqslant 2 \\ \alpha_2 \geqslant 1 \\ \alpha_3 \geqslant 1.5 \\ \alpha_1+\alpha_2 \geqslant 5.5 \\ \alpha_1+\alpha_3 \geqslant 6 \\ \alpha_2+\alpha_3 \geqslant 4.5 \\ \alpha_1+\alpha_2+\alpha_3 = 110/9 \end{cases}$	
z_1, z_5	$\begin{cases} \alpha_1 \geqslant 2 \\ \alpha_2 \geqslant 0 \\ \alpha_3 \geqslant 1 \\ \alpha_1+\alpha_2 \geqslant 3 \\ \alpha_1+\alpha_3 \geqslant 3 \\ \alpha_2+\alpha_3 \geqslant 4 \\ \alpha_1+\alpha_2+\alpha_3 = 6 \end{cases}$	

$$K^{z_2}(2,1,2) = (1,3,1) \notin \Delta CO(z_2),$$
$$K^{z_3}(1,2,1) = (2,2,2) \in CO(z_3) = \Delta CO(z_3),$$
$$K^{z_4}(2,2,1) = (3,6,4) \in CO(z_4) = \Delta CO(z_4).$$

We can easily see that at vertex z_2 the condition of inclusion is not satisfied and we can't guarantee strongly subgame consistency of an imputation from the core if the payments to the players are realised according to initially defined payoff functions.

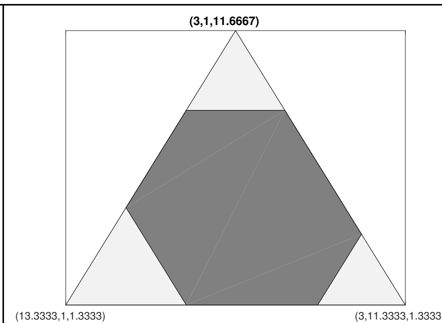
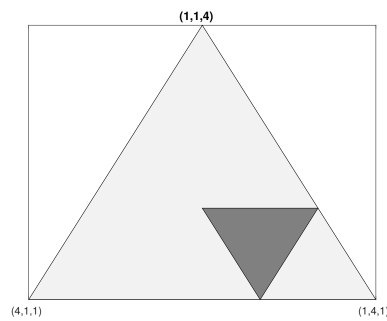
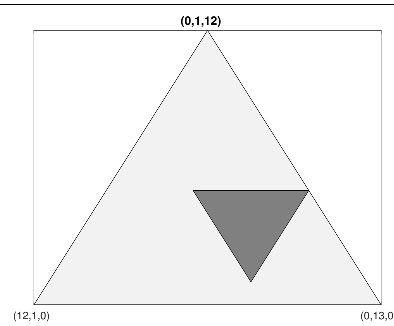
We show that condition (34) does not hold at vertex z_2. Following Definition 16 players may choose any imputation from expected core of vertex z. Let they choose the imputations: $(1.5, 3, 1.5) \in CO(z_3)$ and $(0, 8, 5) \in CO(z_4)$, then the sum at the left-handed term of inclusion (34) takes form:

$$(1,3,1) + \frac{1}{3}(1.5, 3, 1.5) + \frac{2}{3}(0, 8, 5) = \left(\frac{3}{2}, \frac{28}{3}, \frac{29}{6}\right),$$

and this vector does not belong to the core $CO(z_2)$, which means that condition (34) does not hold and the core is not subgame-consistent.

Following Theorem 6, the set of vectors $\beta(z)$ belonging to $\Delta CO(z)$, $z \in CZ$, is the distribution procedure of an imputation from the core $CO(z_0)$ of initially defined game. By Theorem 6 we may also conclude that collection of vectors $(\beta(z) : z \in CZ)$ is not strongly subgame-consistent. For example, consider element from the set $\Delta C(z)$, $z \in CZ$: $\beta(z_0) = (4, 1, 1)$, $\beta(z_1) = (2, 2, 2)$, $\beta(z_2) = (3, 1, 1)$, $\beta(z_3) = (2, 2, 2)$, $\beta(z_4) = (3, 6, 4)$. Calculate the mathematical expectations of the players' payoffs if in the vertices of cooperative subtree they are paid in accordance with

Table 11. The core for the vertices $z_2, z_3, z_4 \in CZ$.

z_2	$\begin{cases} \alpha_1 \geqslant 3 \\ \alpha_2 \geqslant 1 \\ \alpha_3 \geqslant 4/3 \\ \alpha_1+\alpha_2 \geqslant 7 \\ \alpha_1+\alpha_3 \geqslant 7 \\ \alpha_2+\alpha_3 \geqslant 6 \\ \alpha_1+\alpha_2+\alpha_3 = 47/3 \end{cases}$	
z_3	$\begin{cases} \alpha_1 \geqslant 1 \\ \alpha_2 \geqslant 1 \\ \alpha_3 \geqslant 1 \\ \alpha_1+\alpha_2 \geqslant 4 \\ \alpha_1+\alpha_3 \geqslant 3 \\ \alpha_2+\alpha_3 \geqslant 4 \\ \alpha_1+\alpha_2+\alpha_3 = 6 \end{cases}$	
z_4	$\begin{cases} \alpha_1 \geqslant 0 \\ \alpha_2 \geqslant 1 \\ \alpha_3 \geqslant 0 \\ \alpha_1+\alpha_2 \geqslant 8 \\ \alpha_1+\alpha_3 \geqslant 5 \\ \alpha_2+\alpha_3 \geqslant 9 \\ \alpha_1+\alpha_2+\alpha_3 = 13 \end{cases}$	

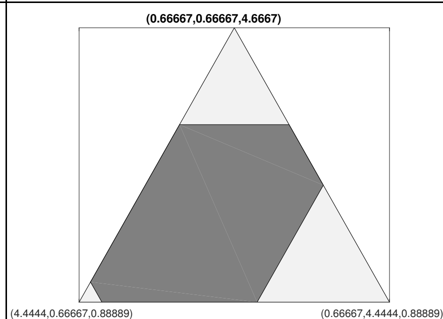
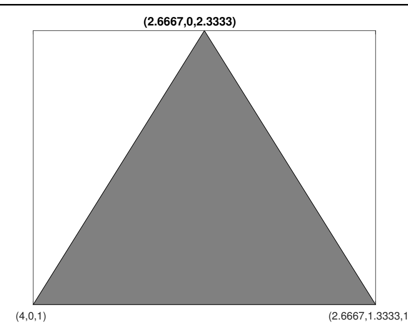
$\{\beta(\cdot)\}$:

$$B(z_0) = (4,1,1) + 0.5\left\{\frac{1}{3}(2,2,2) + \frac{2}{3}\left((3,1,1) + \frac{1}{3}(2,2,2) + \frac{2}{3}(3,6,4)\right)\right\} =$$
$$= \left(\frac{56}{9}, \frac{29}{9}, \frac{25}{9}\right).$$

Obviously, $B(z_0) \in CO(z_0)$.

So, we have proposed a method of construction of strongly subgame-consistent imputation distribution procedure when the core is chosen by the players as a set-valued optimality principle.

Table 12. Sets $\Delta CO(z)$ for vertices z_0 and z_2.

z	$\Delta C(z)$	Graphs of $\Delta CO(z)$
z_0	$\begin{cases} \alpha_1 \geqslant 2/3 \\ \alpha_2 \geqslant 2/3 \\ \alpha_3 \geqslant 8/9 \\ \alpha_1+\alpha_2 \geqslant 8/3 \\ \alpha_1+\alpha_3 \geqslant 19/6 \\ \alpha_2+\alpha_3 \geqslant 11/6 \\ \alpha_1+\alpha_2+\alpha_3 = 6 \end{cases}$	
z_2	$\begin{cases} \alpha_1 \geqslant 8/3 \\ \alpha_2 \geqslant 0 \\ \alpha_3 \geqslant 1 \\ \alpha_1+\alpha_2 \geqslant 1/3 \\ \alpha_1+\alpha_3 \geqslant 8/3 \\ \alpha_2+\alpha_3 \geqslant -4/3 \\ \alpha_1+\alpha_2+\alpha_3 = 5 \end{cases}$	

3. Cooperative stochastic games with infinite duration

3.1. Noncooperative stochastic games with infinite duration

In this section we consider stochastic games with infinite duration defined by Shapley in the paper (Shapley, 1953a). The main classical results on noncooperative stochastic games are presented in (Filar and Vrieze, 1997, Neyman and Sorin, 2003). Similar to the previous section, the game is realised in a discrete time. The significant difference of this stochastic game from the game considered in Section 2 is that now the game has an infinite duration, the set of states which can be realised at any stage is finite and does not change over time. We define first a stochastic game and then describe the set of strategies and the payoff function of the player. Notice that the notations of this section which are widely used in modern literature on stochastic games are a bit different from the notations of Section 2.

Consider stochastic game G defined by

1. The finite set of players $N = \{1, \dots, n\}$.
2. The finite non-empty set of states $\Omega = \{1, \dots, \bar{\omega}\}$;
3. The finite, non-empty set of available actions A_i^ω of player $i \in N$ in state $\omega \in \Omega$. The action of player $i \in N$ in state $\omega \in \Omega$ is element $a_i^\omega \in A_i^\omega$. The action profile in state $\omega \in \Omega$ is a vector of players' actions $a^\omega = (a_i^\omega : i \in N)$. The set of action profiles in state ω is $A^\omega = A_1^\omega \times \ldots \times A_n^\omega$.
4. The finite payoff function $K_i^\omega : \prod_{k \in N} A_k^\omega \to \mathbb{R}$, for every player $i \in N$ and every state $\omega \in \Omega$.

5. The transition function $p(\cdot|\omega, a^\omega) : \Omega \times A^\omega \to \Delta(\Omega)$ from state $\omega \in \Omega$ and action profile $a^\omega \in \prod_{i \in N} A_i^\omega$. Here $\Delta(\Omega)$ is probability distribution over set Ω.
6. The initial state is determined by probability distribution

$$\pi_0 = (\pi_0^1, \ldots, \pi_0^\omega, \ldots, \pi_0^{\bar{\omega}}),$$

where π_0^ω is the probability that state ω is realised at the first stage of the game, $\sum_{\omega \in \Omega} \pi_0^\omega = 1$.

Time is discrete and game G lasts for an infinite number of stages denoted by t. Stochastic game G is realised in the following way:

1. Prior to the game, an initial state ω' is chosen along the probability distribution π_0, i. e., with probability π_0^ω stochastic game starts with state ω.
2. At the first stage, state ω is realised and players simultaneously choose their actions. Player i chooses action $a_i^\omega \in A_i^\omega$, $i \in N$. Thus the action profile $a^\omega = (a_i^\omega : i \in N) \in A_1^\omega \times \ldots \times A_n^\omega$ is realised at the first stage. Player i receives payoff $K_i^\omega(a^\omega)$. Once a^ω is announced for all players, then the game transits to the next state $\omega' \in \Omega$ with probability $p(\omega'|\omega, a^\omega)$.
3. At the second stage, player $i \in N$ chooses action $a_i^{\omega'} \in A_i^{\omega'}$. Thus, at the second stage the action profile $a^{\omega'} = (a_i^{\omega'} : i \in N) \in A_1^{\omega'} \times \ldots \times A_n^{\omega'}$ is played and player i receives payoff $K_i^{\omega'}(a^{\omega'})$.
4. The game further is played in the way described above.

Finally, let $\hat{a}_i^\omega \in \Delta(A_i^\omega)$ be a mixed action of player i in state ω, where $\Delta(A_i^\omega)$ is a probability measure over A_i^ω.

Definition 17. A discounted stochastic game G is defined as

$$G = \left\langle N, \Omega, \{A_i^\omega\}_{\substack{i \in N \\ \omega \in \Omega}}, \{K_i^\omega\}_{\substack{i \in N \\ \omega \in \Omega}}, \pi_0, \left\{p(\omega''|\omega', a^{\omega'})\right\}_{\substack{\omega', \omega'' \in \Omega \\ a^{\omega'} \in \prod_{i \in N} A_i^{\omega'}}}, \delta \right\rangle, \tag{38}$$

where $\delta \in (0, 1)$ is a discount factor, the same for all players.

Every state ω is determined by n-person normal-form game

$$\langle N, \{A_i^\omega\}_{i \in N}, \{K_i^\omega\}_{i \in N} \rangle.$$

A change of state may correspond to the presence of (positive or negative) shocks of different size. They will be reflected on the players' payoffs.

The subgame of noncooperative stochastic game G beginning from stage k is denoted by $G(k)$.

To solve a stochastic game, we need to define the class of players' strategies and the calculation method of players' payoffs in the whole game. First, define players' strategies and distinguish two classes of strategies:

- The behavior strategy of player $i \in N$ is a function $\varphi_i = \{\varphi_i(k)\}_{k=1}^\infty$ and $\varphi_i(k) : h(k) \times \Omega \longmapsto \Delta(A_i^\omega)$, where $h(k)$ is a history of stage k, which is given by a collection of pairs consisting of states and action profiles which were realised at the previous stages until stage k: $((\omega(1), a(1)), (\omega(1), a(2)), \ldots, (\omega(k-1), a(k-1)))$. Denote the set of behavior strategies of player i by Φ_i and behavior strategy profile in stochastic game by $\varphi = (\varphi_i : i \in N)$.

- We also consider the subset of behavior strategies set, that is, the set of stationary strategies. A stationary strategy prescribes a player to choose the same strategy in the same state independently of the history of the stage. Denote a stationary strategy to distinguish behavior (not necessarily stationary) and stationary strategies. Denote a stationary strategy of player i by $\eta_i = \{\eta_i(k)\}_{k=1}^{\infty}$, $\eta_i(k) : \Omega \longmapsto \Delta(A_i^{\omega})$. Denote the profile of stationary strategies in a stochastic game by $\eta = (\eta_i : i \in N)$, and the set of stationary strategies of player i by H_i, while $H_i \subset \Phi_i$.

Now we determine players' payoffs in stochastic game (1):

- For the finite number of stages t a payoff of player i in a stochastic game is determined as a mathematical expectation:

$$E_i(\varphi) = E^{\omega(1),\varphi} \frac{1}{t} \sum_{k=1}^{t} K_i^{\omega(k)}(a(k)),$$

  i. e., a mathematical expectation of a payoff with respect to the initial state $\omega(1)$ and strategy profile φ, while $K_i^{\omega(k)}(a(k))$ is a payoff of player i in state $\omega(k)$ realised at stage k, $a(k)$ is a strategy profile in state $\omega(k)$ realised at stage k in accordance with strategy profile φ.
- In case of infinite game G, a discounted payoff of player i is given by

$$E_i(\varphi) = E^{\omega(1),\varphi} \sum_{k=1}^{\infty} \delta^{k-1} K_i^{\omega(k)}(a(k)) \tag{39}$$

  as a mathematical expectation of the payoff with respect to the initial state $\omega(1)$ and profile φ.

We formulate the main results on the existence of the values of stochastic games with two and more than two players which are used in the present work.

Theorem 7. *(Shapley, 1953a) A two-person zero-sum stochastic game with discount factor $\delta \in (0,1)$ has a value for any initial state. Moreover, players' optimal strategies are stationary.*

This result was extended on the case of nonzero-sum games with more than two players by Fink and Takahashi in 1964:

Theorem 8. *(Fink, 1964, Takahashi, 1964) A nonzero-sum stochastic game with many players with discount factor $\delta \in (0,1)$ and finite set of states and strategies has a value for any initial state. Moreover, there exist optimal stationary strategies of the players.*

3.2. Stochastic games in stationary strategies

In this section we provide formulas to calculate players' payoffs in a stochastic game when players use stationary strategies. Since the set of states Ω is finite, there are only $\bar{\omega}$ subgames $G^{\omega_1}, \ldots, G^{\bar{\omega}}$, each with initial states $\omega_1, \ldots, \bar{\omega}$ respectively, because stationary strategies prescribe the same behavior in the same states even with different histories of the current stage. We denote a non-cooperative stochastic subgame in stationary strategies with initial state $\omega \in \Omega$ by G^{ω}.

We now define the $\bar{\omega} \times \bar{\omega}$-matrix of transition probabilities in G:

$$\Pi(\eta) = \begin{pmatrix} p(\omega_1|\omega_1, a^{\omega_1}) \dots p(\bar{\omega}|\omega_1, a^{\omega_1}) \\ p(\omega_1|\omega_2, a^{\omega_2}) \dots p(\bar{\omega}|\omega_2, a^{\omega_2}) \\ \dots \quad \dots \quad \dots \\ p(\omega_1|\bar{\omega}, a^{\bar{\omega}}) \dots p(\bar{\omega}|\bar{\omega}, a^{\bar{\omega}}) \end{pmatrix} \tag{40}$$

which is a function $p(\omega'|\omega, a^{\omega})$ of a stationary strategy profile $\eta = (\eta_i : i \in N)$ such that $\eta_i(\omega) = a_i^{\omega} \in \Delta(A_i^{\omega})$, $\omega \in \Omega$, $i \in N$, and $a^{\omega} = (a_1^{\omega}, \dots, a_n^{\omega})$ for any state $\omega \in \Omega$. Matrix entry (40) which is the element of the j^{th} row and the j'^{th} column is the probability to transit from state j^{th} to state j'^{th} when players use strategy profile $\eta = (\eta_i : i \in N)$.

We simplify equation (39) for player i's payoff, i.e., we calculate his expected payoff in an explicit form. Let $E_i^{\omega}(\eta)$ be the expected payoff of player i in subgame G^{ω} when profile $\eta = (\eta_1, \dots, \eta_n)$ in stationary strategies is adopted. The vectorial form of the expected payoffs is $E_i(\eta) = (E_i^{\omega_1}(\eta), \dots, E_i^{\bar{\omega}}(\eta))^T$.

Hence a player i's indirect utility function in subgame G^{ω} satisfies the following recurrent equation:

$$E_i^{\omega}(\eta) = K_i^{\omega}(a^{\omega}) + \delta \sum_{\omega' \in \Omega} p(\omega'|\omega, a^{\omega}) E_i^{\omega'}(\eta). \tag{41}$$

Given a matrix form of transition probabilities (40), rewrite equation (41) in a matrix form:

$$E_i(\eta) = K_i(a) + \delta \Pi(\eta) E_i(\eta), \tag{42}$$

where $K_i(a) = (K_i^{\omega_1}(a^1), \dots, K_i^{\bar{\omega}}(a^{\bar{\omega}}))^T$. Equation (3) is equivalent to the equation

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

where $\mathbb{I}$ is an identity matrix of size $\bar{\omega} \times \bar{\omega}$. Matrix $(\mathbb{I} - \delta \Pi(\eta))^{-1}$ always exists for discount factor $\delta \in (0, 1)$. The payoff of player i in game G taking into account the initial state with distributed with π_0 in stationary strategies is

$$\bar{E}_i(\eta) = \pi_0 E_i(\eta) = \pi_0 (\mathbb{I} - \delta \Pi(\eta))^{-1} K_i(a). \tag{43}$$

3.3. Cooperative stochastic games with infinite duration

We now develop the cooperative version of stochastic game G. Suppose that players decide to cooperate by forming a grand coalition N with the aim to maximise total payoff. The existence of maximum of the discounted joint payoff follows from theorem proved in (Shapley, 1953a), according to which the cooperative strategy of the grand coalition that yields the maximal payoff is stationary. Denote the profile of pure stationary strategies of player i as $\eta_i \in H_i$, where $H_i \subset \Phi_i$.[4] The mixed stationary strategy is denoted as $\hat{\eta}_i \in \hat{H}_i$, with $H_i \subset \hat{H}_i$.

A *cooperative strategy profile* or *cooperative solution* maximising the sum of the expected players' payoffs in G is denoted as $\eta^* = (\eta_1^*, \dots, \eta_n^*)$, where[5]

$$\max_{\eta \in \prod_{i \in N} H_i} \sum_{i \in N} \bar{E}_i(\eta) = \sum_{i \in N} \bar{E}_i(\eta^*). \tag{44}$$

[4] From now on we use the notation η_i if player i uses the stationary strategy in the game. When a player i uses a behaviour strategy (not necessarily stationary), we use the notation φ_i.

[5] Without loss of generality we may find the maximum in equation (6) over the set of pure actions of coalition N.

In order to define the cooperative solution of the stochastic game, we determine the values of a characteristic function for any coalition $S \subseteq N$. This function describes how much collective payoff players can gain by forming a coalition. We denote the characteristic function as $V(S) = (V^{\omega_1}(S), \ldots, V^{\bar{\omega}}(S))$. Following (Kohlberg and Neyman, 2015), let $V(S)$ be the minmax value of two-person zero-sum game G_S between coalition S and coalition $N \backslash S$.[6] Before introducing characteristic function, we first define the pure stationary strategies of coalition S and $N \backslash S$ as $\eta_S \in H_S = \prod_{i \in S} H_i$ and $\eta_{N \backslash S} \in H_{N \backslash S} = \prod_{i \in N \backslash S} H_i$, respectively.

Remark 5. When we determine the characteristic function V(S), $S \subseteq N$, we assume that players in S play in the interests of the coalition. Therefore, the actions and strategies of the players in S are correlated (Aumann, 1974).

In state $\omega \in \Omega$, the correlated actions of the players from coalition S are $\hat{a}^{\omega}_S \in \Delta(A^{\omega}_S)$ where $A^{\omega}_S = \prod_{i \in S} A^{\omega}_i$. The correlated stationary strategy of players from coalition S and $N \backslash S$ are $\hat{\eta}_S(\omega) \in \Delta(A^{\omega}_S)$ and $\hat{\eta}_{N \backslash S}(\omega) \in \Delta(A^{\omega}_{N \backslash S})$, respectively. Let the set of correlated stationary strategies of coalition S and $N \backslash S$ be $\hat{H}_S$ and $\hat{H}_{N \backslash S}$, respectively.

Begin the construction of the characteristic function by examining the grand coalition, $S = N$. The Bellman equation for the characteristic function $V(N)$ represents the discounted payoff of N:

$$V(N) = \max_{\eta \in \prod\limits_{i \in N} H_i} \sum_{i \in N} \bar{E}_i(\eta) = \sum_{i \in N} K_i(a^*) + \delta \Pi(\eta^*) V(N), \tag{45}$$

where η^* is the cooperative strategy profile satisfying condition (6) and $\eta^*(\omega) = a^{\omega *}$, $\omega \in \Omega$, and $K_i(a^*) = (K_i^{\omega_1}(a^{\omega_1 *}), \ldots, K_i^{\bar{\omega}}(a^{\bar{\omega} *}))^T$. From (4), we can infer the matrix form of $V(N)$:

$$V(N) = (\mathbb{I} - \delta \Pi(\eta^*))^{-1} \sum_{i \in N} K_i(a^*), \tag{46}$$

where $\mathbb{I}$ is an identity $\bar{\omega} \times \bar{\omega}$-matrix and $\Pi(\eta^*)$ is the $\bar{\omega} \times \bar{\omega}$-matrix of transition probabilities in G when players use the strategy profile η^*. Matrix $\Pi(\eta^*)$ is described in details by (40).

We define next the value of $V^{\omega}(S)$ of coalition S as the minmax payoff in the subgame G^{ω}_S starting from state ω:

$$V^{\omega}(S) = \min_{\hat{\eta}_{N \backslash S}} \max_{\eta_S} \sum_{i \in S} E^{\omega}_i(\eta_S, \hat{\eta}_{N \backslash S}) = \max_{\hat{\eta}_S} \min_{\eta_{N \backslash S}} \sum_{i \in S} E^{\omega}_i(\hat{\eta}_S, \eta_{N \backslash S}). \tag{47}$$

In equation (9), the maximum in $\min\limits_{\hat{\eta}_{N \backslash S}} \max\limits_{\eta_S} \sum\limits_{i \in S} E^{\omega}_i(\eta_S, \eta_{N \backslash S})$ is found over the set of pure strategies of coalition S, while the minimum in $\max\limits_{\hat{\eta}_S} \min\limits_{\eta_{N \backslash S}} \sum\limits_{i \in S} E^{\omega}_i(\eta_S, \eta_{N \backslash S})$ is found over the set of pure strategies of coalition $N \backslash S$.

The Bellman equation for the characteristic function $V^{\omega}(S)$ is

$$\begin{aligned} V^{\omega}(S) &= \min_{\hat{\eta}_{N \backslash S} \in \hat{H}_{N \backslash S}} \max_{\eta_S \in H_S} \sum_{i \in S} E^{\omega}_i(\eta_S, \hat{\eta}_{N \backslash S}) = \sum_{i \in S} E^{\omega}_i(\eta_S, \hat{\eta}_{N \backslash S}) \\ &= \sum_{i \in S} K^{\omega}_i(\mathbf{a}^{\omega}_S, \hat{\mathbf{a}}^{\omega}_{N \backslash S}) + \delta \sum_{\omega' \in \Omega} p\left(\omega' | \omega, (\mathbf{a}^{\omega}_S, \hat{\mathbf{a}}^{\omega}_{N \backslash S})\right) V^{\omega'}(S), \end{aligned} \tag{48}$$

[6] The existence of the minmax value of two-player discounted stochastic game is proved by Shapley (1953a).

where $(\mathbf{a}_S^\omega, \hat{\mathbf{a}}_{N\setminus S}^\omega)$ is a profile in correlated actions in state $\omega \in \Omega$ such that $\eta_S(\omega) = \mathbf{a}_S^\omega$, $\hat{\eta}_{N\setminus S}(\omega) = \hat{\mathbf{a}}_{N\setminus S}^\omega$, and $K_i(\mathbf{a}_S^\omega, \hat{\mathbf{a}}_{N\setminus S}^\omega) = (K_i^{\omega_1}(\mathbf{a}_S^{\omega_1}, \hat{\mathbf{a}}_{N\setminus S}^{\omega_1}), \ldots, K_i^{\bar{\omega}}(\mathbf{a}_S^{\bar{\omega}}, \hat{\mathbf{a}}_{N\setminus S}^{\bar{\omega}}))$. We then rewrite equation (48) in a matrix form:

$$V(S) = \left(\mathbb{I} - \delta \Pi(\eta_S, \hat{\eta}_{N\setminus S})\right)^{-1} \sum_{i\in S} K_i(\mathbf{a}_S^\omega, \hat{\mathbf{a}}_{N\setminus S}^\omega). \tag{49}$$

Finally, we define the characteristic function $\bar{V}(S)$ for the whole stochastic game as:

$$\bar{V}(S) = \pi_0 V(S), \tag{50}$$

for any coalition $S \subseteq N$, where $V(S) = (V^{\omega_1}(S), \ldots, V^{\bar{\omega}}(S))$, and $V^\omega(S)$ is the value of the characteristic function of subgame G^ω for S.

The characteristic function satisfies two properties. First, for any state $\omega \in \Omega$:

$$V^\omega(\varnothing) = 0. \tag{51}$$

Second, the characteristic functions $\bar{V}(S)$ and $V^\omega(S)$ determined by (10) and (7)-(51), respectively, are superadditive (Aumann and Peleg, 1960). In other words, for any disjoint coalitions $S, T \subset N$, and $S \cap T = \varnothing$, the inequality $V(S) + V(T) \leqslant V(S \cup T)$ holds. Superadditivity implies that the value of two disjoint coalitions is at least as great when they play together as when they act non-cooperatively. If superadditivity is not satisfied, then the coalition $S \cup T$ is not profitable, thus it will not be formed.[7]

We are now in a position to define the cooperative version of stochastic game 17 and its subgames.

Definition 18. A cooperative stochastic game G_c, corresponding to a stochastic game G, is a set $\langle N, \bar{V}\rangle$, where N is the set of players and $\bar{V} : 2^N \longrightarrow \mathbb{R}$ is the characteristic function calculated by (10). A cooperative stochastic subgame G_c^ω *starting from state* ω is a set $\langle N, V^\omega\rangle$, where $V^\omega : 2^N \longrightarrow \mathbb{R}$ is the characteristic function calculated by (7), (9) and (51).

When forming the grand coalition, players should decide not only what strategies to use to maximise the joint payoff but also how to allocate the total payoff. The next definitions display the allocation rule or solution (also called *imputation*) of G_c^ω and G_c, respectively. To determine an imputation of the joint payoff (6) we need to determine the values of the characteristic function for any coalition $S \subset N$.

Definition 19. An imputation in the subgame G_c^ω, $\omega \in \Omega$, is a vector $\sigma^\omega = (\sigma_1^\omega, \ldots, \sigma_n^\omega)$ satisfying: (i) $\sum_{i\in N} \sigma_i^\omega = V^\omega(N)$, and (ii) $\sigma_i^\omega \geqslant V^\omega(\{i\})$ for any $i \in N$. The set of imputations in G_c^ω is denoted as Σ^ω.

Definition 20. An imputation in the game G_c is a vector $\bar{\sigma} = (\bar{\sigma}_1, \ldots, \bar{\sigma}_n)$, where $\bar{\sigma}_i = \pi_0 \sigma_i$, $\sigma_i = (\sigma_i^{\omega_1}, \ldots, \sigma_i^{\bar{\omega}})^T$, and $(\sigma_1^\omega, \ldots, \sigma_n^\omega) = \sigma^\omega \in \Sigma^\omega$, $\omega \in \Omega$. The set of imputations in G_c is denoted as $\bar{\Sigma}$.

[7] The property of superadditivity is not needed and it is often omitted in cooperative game theory, because in real life there are a lot of motivations to consider both profitable and non-profitable coalitions. As Aumann and Dreze (1974, p. 233) note, there are arguments for superadditivity that are quite persuasive, but, as they also note, superadditivity is quite problematic in some economic applications.

By Definition 19, an imputation satisfies the following conditions: (i) any player should obtain no less than she may get by non-cooperative play (*individual rationality condition*) and (ii) the sum of components of the imputation equals the value of the characteristic function corresponding to grand coalition (*group rationality condition*). The set of imputations is non-empty in any subgame G_c^ω, $\omega \in \Omega$ and in the whole cooperative stochastic game G_c, since the characteristic function determined by equations (4)-(51) is superadditive.

3.4. Principles of stable cooperation

In cooperative games, the solution of a game is determined by an optimality principle. The optimality principle is assumed to be the subset of the imputation set. Therefore, the optimality principle contains one or more than one imputations or solutions of a cooperative game but sometimes it maybe empty. For example, the core may be empty, then the solution of a cooperative game does not exist according to this optimality principle. The Shapley value as an optimality principle always exists and contains a unique imputation. Therefore, the solution of a cooperative game always exists and it is unique according to this optimality principle. The solution of cooperative stochastic game means an imputation.[8] Now we do not consider the problem of choosing a unique imputation from the set but assume that the optimality principle contains the only one imputation. The examples of one-point solutions are the Shapley value (Shapley, 1953b), the Von Neumann-Morgenstern solution (von Neumann and Morgenstern, 1944) and the nucleolus (Schmeidler, 1969). The realisation of an imputation in a cooperative stochastic game requires the satisfaction of some principles, which in turn ensure *stable cooperation* in a game. Following (Petrosyan and Zenkevich, 2015), we formulate the main principles of stable cooperation including subgame consistency, strategic support (or strategic stability) and irrational-behaviour-proof of the solution of a cooperative stochastic game. Each principle of stable cooperation is defined and analysed separately.

Subgame consistency. The principle of subgame consistency ensures that in any subgame cooperative solution is determined according to the initially chosen allocation rule. This concept deserves a detailed explanation. Players agree on cooperation before the game and adopt an imputation following the allocation mechanism. During the game, they play a cooperative strategy profile a_i^*, $i \in N$ which maximises their total payoff. In any subgame beginning in a certain state, a player is able to derive her expected payoff for the remainder of the game. If at some intermediate stage of the game players decide to calculate their expected payoffs in the subgame according to the initially defined payoff functions, then most often these expected discounted payoffs do not coincide with an imputation calculated in accordance with the initially chosen optimality principle. This means subgame inconsistency of a cooperative solution (or optimality principle). If for any subgame discounted players' payoffs coincide with the imputations calculated in accordance with initial optimality principle, cooperative solution (or optimality principle) is subgame consistent (see Petrosyan, 1977). To make cooperative solution subgame consistent, we

[8] We further consider the case when the solution of a cooperative stochastic game is an imputation set consisting of more than one imputation.

propose the transfer mechanism, called *imputation distribution procedure* (IDP).[9] Originally, the idea of IDP was proposed by L. A. Petrosyan for differential games (Petrosyan and Danilov, 1979).

This mechanism leads to a modification of the players' payoffs in a dynamic game. We call the modified game as *σ-regularisation*, where σ is an initially chosen imputation in cooperative game G_c. This modified game ensures several advantages to the players. First, subgame consistency is ensured through the "new" payoff functions. Second, the expected payoffs in the regularised game will be equal to the components of the chosen imputation σ. Moreover, the sum of the stage payoffs in the regularised game is equal to the sum of the payoffs in the correspondent state of the initial game. For instance, suppose that players choose the Shapley value at the beginning of the game as an allocation rule. In this case, subgame consistency guarantees that, in each subgame, the vector of the players' payoff for the remaining stages is the Shapley value calculated for this subgame.

Let players adopt cooperative solution in stochastic game, i.e., they choose imputation $\sigma^\omega = (\sigma_1^\omega, \dots, \sigma_n^\omega)^T \in \Sigma^\omega$ for every subgame G_c^ω. The problem is to determine the transfers that ensure the expected payoff σ_i^ω for player i in every subgame G_c^ω. If transfers are based on the payoff functions in every state, then players can hardly expect to get the payoff based on the initially chosen allocation rule. To overcome this, we propose a rule to transfer the players' total payoff, based on the method for differential games (Petrosyan and Danilov, 1979).

Since strategies are stationary, the number of states corresponds to the number of relevant "different" histories. In turn, when players implement cooperative strategies in the stochastic game (1), the number of relevant subgames is equal to the number of possible states. Therefore, we need to determine a vector of transfers $\beta_i = (\beta_i^{\omega_1}, \dots, \beta_i^{\bar{\omega}})^T$ for where β_i^ω is the transfer of player $i \in N$ in state $\omega \in \Omega$.

Definition 21. The set of transfers $\{\beta_i\}_{i\in N}$ is IDP if the following conditions are satisfied:

1. In each state $\omega \in \Omega$, the sum of the transfers is equal to the sum of players' payoffs in cooperative strategy profile η^*:
$$\sum_{i\in N} \beta_i^\omega = \sum_{i\in N} K_i^\omega(a^{\omega *}). \tag{52}$$
2. The expected sum of transfers to player $i \in N$ in the game $\bar{G}$ is equal to the i^{th} component of the initially chosen imputation $\bar{\sigma}$.

We then define the conditions of subgame consistency for the imputation and IDP.

Definition 22. Imputation $\bar{\sigma} = (\bar{\sigma}_1, \dots, \bar{\sigma}_n)$ and corresponding IDP $\{\beta_i\}_{i\in N}$ are called subgame consistent if the expected sum of transfers to player i in each subgame G^ω is equal to the i^{th} component of the initially chosen imputation in subgame G_c^ω (in accordance with the principle imputation $\bar{\sigma}$ of the whole game is calculated).

[9] Imputation distribution procedure was adapted for the class of discounted stochastic games in (Baranova and Petrosjan, 2006). See Petrosjan and Danilov (1979), and Baranova and Petrosjan (2006).

The following statement suggests the method of IDP construction for imputation $\bar{\sigma}$.

Lemma 2. *Let imputation* $\bar{\sigma}$ *be such that* $(\bar{\sigma}_1, \dots, \bar{\sigma}_n) \in \bar{\Sigma}$ *where* $\bar{\sigma}_i = \pi_0 \sigma_i$, $\sigma_i = (\sigma_i^{\omega_1}, \dots, \sigma_i^{\bar{\omega}})^T$ *and* $(\sigma_1^{\omega}, \dots, \sigma_n^{\omega}) = \sigma^{\omega} \in \Sigma^{\omega}$. *Then the collection* $\{\beta_i\}_{i \in N}$ *where* β_i *calculated by*

$$\beta_i = (\mathbb{I} - \delta \Pi(\eta^*))\sigma_i, \tag{53}$$

is an imputation distribution procedure[10] *in game* G.

Proof. Verify the IDP condition:

$$\sum_{i \in N} \beta_i^{\omega} = \sum_{i \in N} K_i^{\omega}(a^{\omega *}),$$

where $a^{\omega *}$ is an action profile adopted under cooperative profile η^* in state ω. It is easy to show that β_i from (53) satisfies (52). Since $\sum_{i \in N} \beta_i^{\omega}$ is equal to $(\mathbb{I} - \delta \Pi(\eta^*)) \sum_{i \in N} \sigma_i = (\mathbb{I} - \delta \Pi(\eta^*)) V(N)$, and $V(N)$ is determined by (7), then equation (52) holds.

The second IDP condition is satisfied since the expected total payoff of player i, denoted as B_i, with new transfer β_i^{ω} in state $\omega \in \Omega$ satisfies the recurrent equation:

$$B_i^{\omega} = \beta_i^{\omega} + \delta \sum_{\omega' \in \Omega} p(\omega' | \omega, a^{\omega *}) B_i^{\omega'},$$

or, in vectorial form:

$$B_i = \beta_i + \delta \Pi(\eta^*) B_i, \tag{54}$$

where $B_i = (B_i^{\omega_1}, \dots, B_i^{\bar{\omega}})^T$. Equation (54) is equivalent to:

$$B_i = (\mathbb{I} - \delta \Pi(\eta^*))^{-1} \beta_i. \tag{55}$$

Given the second condition of IDP and equation (55) we obtain:

$$\sigma_i = (\mathbb{I} - \delta \Pi(\eta^*))^{-1} \beta_i, \tag{56}$$

where $\sigma_i = (\sigma_i^{\omega_1}, \dots, \sigma_i^{\bar{\omega}})^T$, $(\sigma_1^{\omega}, \dots, \sigma_n^{\omega}) = \sigma^{\omega} \in \Sigma^{\omega}$. Equation (56) can be rewritten equivalently as:

$$\beta_i = (\mathbb{I} - \delta \Pi(\eta^*))\sigma_i. \tag{57}$$

Finally, equation (53) equals to:

$$\sigma_i = \beta_i + \delta \Pi(\eta^*)\sigma_i. \tag{58}$$

The second item in the right part of (16) is the expected value of the transfers calculated for the subgame from the next stage onwards. Suppose that the imputation for each subgame is chosen following the same allocation rule that has been chosen by the players at the beginning of the game. If players maintain cooperative strategy profile η^*, then the expected payoff of player i with new transfers is equal to the correspondent component of imputation $\bar{\sigma}$ in cooperative stochastic game G_c.

[10] Notice that IDP is uniquely defined by formula (53) if optimality principle provides unique cooperative solution $\bar{\sigma}$ (e.g., if the solution is nucleolus, the Shapley value or another single-valued solution). If the cooperative solution is the set of imputations containing more than one imputation, the method of IDP construction should be modified (see Parilina and Zaccour, 2015).

Given Definition 21, for every imputation $\bar{\sigma} = (\bar{\sigma}_1, \dots, \bar{\sigma}_n) \in \bar{\Sigma}$, where $\bar{\sigma}_i = \pi_0 \sigma_i$, $\sigma_i = (\sigma_i^{\omega_1}, \dots, \sigma_i^{\bar{\omega}})^T$, $(\sigma_1^{\omega}, \dots, \sigma_n^{\omega}) = \sigma^{\omega} \in \Sigma^{\omega}$, we can define the regularization of stochastic game G as follows:

Definition 23. A σ-regularisation of stochastic game G (subgame G^{ω}, $\omega \in \Omega$) is non-cooperative stochastic game G_{σ} (subgame G_{σ}^{ω}) if, for any player $i \in N$ in state ω, payoff function $K_i^{\sigma,\omega}(a^{\omega})$ is defined as:

$$K_i^{\sigma,\omega}(a^{\omega}) = \begin{cases} \beta_i^{\omega}, & \text{if } a^{\omega} = a^{\omega *}, \\ K_i^{\omega}(a^{\omega}), & \text{if } a^{\omega} \neq a^{\omega *}, \end{cases} \tag{59}$$

where β_i^{ω} is a component of PDP of player i defined by (53) and $a^{\omega *} = \eta^*(\omega)$.

Equation (59) determines the modified payoff function for game G.

Remark 6. The σ-regularisation changes the payoff functions in any state $\omega \in \Omega$ only when action profiles $a^{\omega *} = \eta^*(\omega)$ are adopted. We may expect that players agree to modify the initial payoff functions to be sure that their cooperative solution satisfies the principle of subgame consistency.

The following theorem shows that the players' payoffs in σ- regularization of initial stochastic game G satisfy the principle of subgame consistency.

Theorem 9. *Let $\bar{\sigma} = (\bar{\sigma}_1, \dots, \bar{\sigma}_n) \in \bar{\Sigma}$ be the initially chosen imputation in game G, where $\bar{\sigma}_i = \pi_0 \sigma_i$, $\sigma_i = (\sigma_i^{\omega_1}, \dots, \sigma_i^{\bar{\omega}})^T$, $(\sigma_1^{\omega}, \dots, \sigma_n^{\omega}) = \sigma^{\omega} \in \Sigma^{\omega}$, then σ-regularization of stochastic game G satisfies the principle of subgame consistency, i.e., the cooperative solution $\bar{\sigma}$ is subgame consistent in game G_{σ}.*

Proof. At the beginning of the game, players choose the following imputation: $\bar{\sigma} = (\bar{\sigma}_1, \dots, \bar{\sigma}_n) \in \bar{\Sigma}$, where $\bar{\sigma}_i = \pi_0 \sigma_i$, $\sigma_i = (\sigma_i^{\omega_1}, \dots, \sigma_i^{\bar{\omega}})^T$, $(\sigma_1^{\omega}, \dots, \sigma_n^{\omega}) = \sigma^{\omega} \in \Sigma^{\omega}$. A cooperative strategy profile is η^*. Consider the σ-regularization of game G determined by Definition 23, thus the set of transfers $\{\beta_i\}_{i \in N}$ defined by (53) is a IDP which follows from Lemma 2. To prove that the σ-regularisation of the game G satisfies the principle of subgame consistency, we need to calculate the discounted payoffs in every subgame of the game G_{σ} when a cooperative strategy profile η^* occurs. Consider any subgame G_{σ}^{ω} starting from state $\omega \in \Omega$. The discounted payoff of player i in this subgame is:

$$E_i^{\omega}(\eta^*) = \beta_i^{\omega} + \delta \sum_{\omega' \in \Omega} p(\omega' | \omega, a^{\omega *}) E_i(\eta^*), \tag{60}$$

where $E_i(\eta^*) = (E_i^{\omega_1}(\eta^*), \dots, E_i^{\bar{\omega}}(\eta^*))^T$ and $E_i^{\omega}(\eta^*)$ is the discounted payoff of player i in subgame G_{σ}^{ω} starting from state ω when players adopt η^*. Equation (60) can be rewritten in a vector form:

$$E_i(\eta^*) = \beta_i + \delta \Pi(\eta^*) E_i(\eta^*),$$

or

$$E_i(\eta^*) = (\mathbb{I} - \delta \Pi(\eta^*))^{-1} \beta_i.$$

Since β_i satisfies (53), we obtain

$$E_i(\eta^*) = (\mathbb{I} - \delta \Pi(\eta^*))^{-1} (\mathbb{I} - \delta \Pi(\eta^*)) \sigma_i = \sigma_i.$$

This equation proves that σ-regularization of game G satisfies the principle of subgame consistency.

Definition 23 and Theorem 9 provide a method of constructing subgame consistent transfers in every state of a stochastic game. The imputation distribution procedure $\beta_i^{\omega_1}, \ldots, \beta_i^{\bar{\omega}}$ in states $\omega_1, \ldots, \bar{\omega}$ ensures that a player i receives the same expected payoff in game G_σ (subgame G_σ^ω), as she planned to receive in cooperative stochastic game G_c (subgame G_c^ω). Moreover, the expected payoff from future transfers is in line with the same allocation rule chosen by the players at the beginning of the game.

Strategic support. The principle of strategic support ensures that, along the whole game, an individual deviation from cooperative strategy profile in a regularized game does not yield a higher payoff than cooperation. In other words, it guarantees the existence of the Nash equilibrium in a regularized game with the same payoffs that players expect to receive with the cooperative solution (which was the basis of regularization). This principle was proposed in (Petrosyan, 1998). We reformulate the principle and then find conditions under which Nash equilibrium is subgame perfect (see Selten, 1975) in a regularized game with the payoffs described above.

The subgame perfectness is important for dynamic games because it allows to guarantee the existence of the Nash equilibrium in any subgame with which the players' payoffs coincide with the cooperative ones. Comparing our approach with the standard analysis of deterministic (repeated) games, the condition of strategic support for stochastic (or dynamic) games corresponds to the condition of the existence of subgame perfect Nash equilibrium in grim-trigger strategies. The main difference is that, in our setting, players first regularize the initial game by adapting the IDP to achieve subgame consistency.

Suppose players come to a cooperative agreement, i. e., find a cooperative strategy profile η^* that maximises the expected total payoff in the whole game. If a player deviates from the cooperative strategy profile, then the other players switch to trigger strategy from the next stage until forever to punish the deviating player. The strict definition of a behavior strategy used by players in Nash equilibrium is given below (see formula (63)). Here we assume that a stochastic game is the game with perfect monitoring, that is, all players know the state of a current stage and the history of the stage.

To begin with, we define the Nash equilibrium in a regularized stochastic game. Denote the expected payoff of player i in σ-regularisation of subgame G^ω starting from state ω as $E_i^{\omega,\sigma}$.

Definition 24. A Nash equilibrium in the regularised game G_σ is a behaviour strategy profile $\varphi^* = (\varphi_1^*, \ldots, \varphi_n^*)$ such that, for any player $i \in N$ and for any state $\omega \in \Omega$, the condition

$$E_i^{\omega,\sigma}(\varphi_i^*, \varphi_{N\setminus i}^*) \geqslant E_i^{\omega,\sigma}(\varphi_i, \varphi_{N\setminus i}^*) \tag{61}$$

holds for any behaviour strategy of player i: $\varphi_i \in \Phi_i$.

We assume that the behaviour strategy exhibits the following structure. If, in the history of stage k, all players use their cooperative strategies, then they implement the cooperative correlated actions also in stage k. Conversely, if before stage k the individual deviation of a player $z \in N$ is observed, then the coalition $N\setminus z$ punishes player z. We assume that the punishment ensures that player z's payoff is at most her minimax value in any subgame.[11] Notice that, since we focus on a Nash equilibrium, we need to consider only individual deviations from this profile.[12] If deviation occurs by more than one member of the coalition, the player may implement any strategy from the her set of strategies.

We now outline the condition under which the Nash equilibrium with players' payoffs equal the cooperative ones exists. For convenience, define

$$F(\{i\}) \equiv (F^{\omega_1}(\{i\}), \ldots, F^{\bar{\omega}}(\{i\}))^T,$$
$$F^{\omega}(\{i\}) = \max_{\hat{a}_i^{\omega} \in \Delta(A_i^{\omega})} \left\{ K_i^{\omega}(\hat{a}_i^{\omega}, a_{N\setminus i}^{\omega *}) + \delta \sum_{\omega' \in \Omega} p(\omega'|\omega, (\hat{a}_i^{\omega}, a_{N\setminus i}^{\omega *}))V^{\omega'}(\{i\}) \right\}.$$

The following inequality:

$$\sigma_i = (\mathbb{I} - \delta \Pi(\eta^*))^{-1}\beta_i \geqslant F(\{i\}), \tag{62}$$

compares two payoffs for each subgame: (i) the payoff when players adopt the cooperative strategy profile in the left hand side, and (ii) the payoff of deviation plus future punishment in the right hand side. If the first payoff is greater or equal to the second one, the player gets no benefit from deviation. If this is true for any player and any state, then the principle of strategic stability is satisfied. This result is summarised in the following proposition.

Proposition 1. *If in an σ-regularisation G_σ such that $\bar{\sigma} = \pi_0\sigma$, inequality* (14) *holds for any player $i \in N$, then there exists behaviour strategy profile $\widehat{\varphi}$ such that it is the Nash equilibrium with players' payoffs $(\bar{\sigma}_1, \ldots, \bar{\sigma}_n)$.*

Proof. We determine the behaviour strategy profile $\widehat{\varphi} = (\widehat{\varphi}_1, \ldots, \widehat{\varphi}_n)$ where strategies $\widehat{\varphi}_i$, $i \in N$ are:

$$\widehat{\varphi}_i(h(k)) = \begin{cases} a_i^{\omega *}, & \text{if } \omega(k) = \omega, h(k) \subset h^*; \\ \hat{a}_i^{\omega}(z), & \text{if } \omega(k) = \omega, \text{ and } \exists\, l \in [1, k-1], \\ & z \in N,\ i \neq z\text{: } h(l) \subset h^*, \text{ and} \\ & (\omega(l), a(l)) \notin h^*, \text{ but} \\ & (\omega(l), (a_z^*(l), a_{N\setminus z}(l)) \in h^*; \\ \text{any} & \text{otherwise,} \end{cases} \tag{63}$$

where $a_i^{\omega *}$ corresponds to the player i's cooperative action, while $\hat{a}_i^{\omega}(z) \in \Delta(A_i^{\omega})$ is the player i's punishment that, together with actions $\hat{a}_{i'}^{\omega}(z) \in \Delta(A_{i'}^{\omega})$, of players $i' \neq i$, $i' \in N\setminus z$, forms the action (either in pure or mixed strategies) of coalition $N\setminus z$ against player z.[13] The proof of the proposition follows from the folk theorem for stochastic games (Dutta, 1995) using the structure of the behaviour strategy (63). Notice that we do not define the reaction of players when they observe the deviations of more than one player. This because we focus here on the

[11] The strict definition of the behaviour strategy is given in the proof of Proposition 1.

[12] Things change for subgame perfectness. In this case, we need to prove that eq. (13) holds for all possible histories and all stages. Therefore, we need to determine the strategy of a player even if more than one player deviates. Strategy (71) defines the behaviour of the player given any history.

[13] Notice that the actions of the players from coalition $N\setminus z$ are correlated.

Nash equilibrium (not subgame perfect). When more than one player deviates, the player chooses any strategy from the player's set of strategies. We now prove that $\widehat{\varphi}(\cdot) = (\widehat{\varphi}_1(\cdot), \dots, \widehat{\varphi}_n(\cdot))$ determined in (63) is a NE in the stochastic game G_σ. Given strategy (63) and provided that all players do not deviate from a cooperative strategy profile η^*, the discounted payoff of player i in the subgame G_σ^ω, $\omega \in \Omega$, is:

$$E_i^\omega(\widehat{\varphi}) = E_i^\omega(\eta^*).$$

Let $E_i(\widehat{\varphi})$ be equal to the vector $(E_i^{\omega_1}(\widehat{\varphi}), \dots, E_i^{\bar{\omega}}(\widehat{\varphi}))^T$. Then for any player $i \in N$ the next equation holds:

$$E_i(\widehat{\varphi}) = (\mathbb{I} - \delta \Pi(\eta^*))^{-1} \beta_i. \tag{64}$$

Consider next the profile of strategies $(\varphi_z, \widehat{\varphi}_{N\setminus z})$, when some player z deviates from strategy $\widehat{\varphi}_z$. For any k, there exists $l \in [1, k-1]$ such that $h(l) \subset h^*$ but $(\omega(k), a(k)) \notin h^*$ and $(\omega(k), (a_z^*(k), a_{N\setminus z}(k))) \in h^*$. Without loss of generality, we simplify $\omega(k) = \omega$. In words, the first individual deviation of player z occurs at stage k. We are now able to determine the total payoff of player z in the game G_σ with strategy profiles $(\varphi_z, \widehat{\varphi}_{N\setminus z})$ by

$$\bar{E}_z^\sigma(\varphi_z, \widehat{\varphi}_{N\setminus z}) = \pi_0 E_z^\sigma(\varphi_z, \widehat{\varphi}_{N\setminus z}),$$

where

$$E_z^\sigma(\varphi_z, \widehat{\varphi}_{N\setminus z}) = E_z^{\sigma,[1,k-1]}(\varphi_z, \widehat{\varphi}_{N\setminus z}) + \delta^{k-1} \Pi^{k-1}(\varphi_z, \widehat{\varphi}_{N\setminus z}) E_z^{\sigma,[k,\infty)}(\varphi_z, \widehat{\varphi}_{N\setminus z}). \tag{65}$$

The first term in the right hand side of (65) is the expected payoff of player z in the first $k-1$ stages of the game G_σ, the second term is the expected payoff of player z in the subgame of G_σ beginning from stage k, where $E_z^{\sigma,[k,\infty)}(\varphi_z, \widehat{\varphi}_{N\setminus z})$ is the vector $(E_z^{\sigma,1}(\varphi_z, \widehat{\varphi}_{N\setminus z}), \dots, E_z^{\sigma,\bar{\omega}}(\varphi_z, \widehat{\varphi}_{N\setminus z}))^T$, with $E_z^{\sigma,\omega}(\varphi_z, \widehat{\varphi}_{N\setminus z})$ being the player z's expected payoff in the regularised subgame G_σ^ω beginning at state ω. Since there are no deviations from a cooperative strategy profile η^* up to stage $k-1$, the following equalities hold:

$$E_z^{\sigma,[1,k-1]}(\varphi_z, \widehat{\varphi}_{N\setminus z}) = E_z^{\sigma,[1,k-1]}(\eta^*),$$

$$\Pi^{k-1}(\varphi_z, \widehat{\varphi}_{N\setminus z}) = \Pi^{k-1}(\eta^*).$$

We now find the discounted payoff of player z in the subgame G_σ^ω beginning with stage k and when state $\omega(k)$ is equal to ω. The following formula takes place:

$$E_z^{\sigma,\omega}(\varphi_z, \widehat{\varphi}_{N\setminus z}) = K_z^\omega(\hat{a}_z^\omega, a_{N\setminus z}^{\omega*}) + \delta \sum_{\omega' \in \Omega} p(\omega'|\omega, (\hat{a}_z^\omega, a_{N\setminus z}^{\omega*})) V^{\omega'}(\{z\}), \tag{66}$$

where $\hat{a}_z^\omega \in \Delta(A_z^\omega)$. Players from the coalition $N \setminus z$ punish player z by playing the strategies which allow player z to obtain her minmax payoff according to the definition of strategy profile $\widehat{\varphi}$. In (66), the value of the characteristic function $V^{\omega'}(\{z\})$ is determined by (9). Since the expected payoffs of player z in the strategy profiles $\widehat{\varphi}$ and $(\varphi_z, \widehat{\varphi}_{N\setminus z})$ do not change up to stage $k-1$, then a deviation may increase player z's payoff only at the expenses of the expected payoff in the subgame

G^{ω}_{σ}, $\omega \in \Omega$. In particular, the strategy profile $(\varphi_z, \widehat{\varphi}_{N\setminus z})$ ensures the following expected payoff of player z from stage k:

$$F(\{z\}) = \max_{\hat{a}^{\omega}_z \in \Delta(A^{\omega}_z)} \left\{ K^{\omega}_z(\hat{a}^{\omega}_z, a^{\omega *}_{N\setminus z}) + \delta \sum_{\omega' \in \Omega} p(\omega'|\omega, (\hat{a}^{\omega}_z, a^{\omega *}_{N\setminus z}))V^{\omega'}(\{z\}) \right\}. \quad (67)$$

According to the definition of PDP, the expected payoff of player z in the regularised subgame G^{ω}_{σ} with a profile of strategies $\widehat{\varphi}(\cdot)$ can be found from:

$$E^{\sigma}_z(\widehat{\varphi}) = (\mathbb{I} - \delta\Pi(\eta^*))^{-1}\beta_z = \sigma_z, \quad (68)$$

where $E^{\sigma}_z(\widehat{\varphi}) = (E^{\sigma,\omega_1}_z(\widehat{\varphi}), \dots, E^{\sigma,\bar{\omega}}_z(\widehat{\varphi}))^T$. Taking into account (14) from (67), (68) and the above discussion we get

$$E^{\sigma}_z(\widehat{\varphi}) \geqslant E^{\sigma}_z(\varphi_z, \widehat{\varphi}_{N\setminus z}),$$

which is satisfied when inequality

$$\sigma_z = \left(\mathbb{I} - \delta\Pi\left(\eta^*\right)\right)^{-1} \beta_z \geq F(\{z\}) \quad (69)$$

is true. In inequality (69) is satisfied for any player $z \in N$, a player is not willing to deviate from the cooperative strategy profile in any subgame of the σ-regularisation of game G.

Thus the behaviour strategy profile (63) is a NE in the σ-regularisation of game G. The discounted payoff of i in the game G_{σ} with profile of strategies $\widehat{\varphi}$ is equal to $\bar{\sigma}_i$, where $\bar{\sigma}_i = \pi_0\sigma_i$, while $\sigma_i = (\sigma^{\omega_1}_i, \dots, \sigma^{\bar{\omega}}_i)^T$ consists of i^{th} components of imputations σ^{ω_1}, …, $\sigma^{\bar{\omega}}$ derived from the cooperative subgames G^1, …, $G^{\bar{\omega}}$ accordingly.

Notice that the players' strategies used in a punishing regime of the behaviour strategies (63) are not individually rational, i.e., player i punishing the deviated player z needs to implement the strategies minimizing the payoff of player z in a subgame which may be not profitable for player i and may motivate player i to deviate from strategy profile formed by (63). Therefore, the strategy profile determined by strategies (63) is not subgame perfect.

We investigate now the conditions to obtain a subgame perfect Nash equilibrium (SPNE) of the σ-regularisation of G. To do so, we need to determine the behaviour strategy profile such that, for any state occurring in any period with any history, individual deviation is not profitable.

We assume that, if the history of the stage differs from the cooperative history, then all players implement a Nash equilibrium of the game G denoted by $\hat{\eta}^{ne} = (\hat{\eta}^{ne}_1, \dots, \hat{\eta}^{ne}_n)$ such that $\hat{\eta}^{ne}_i(\omega) \in \Delta(A^{\omega}_i)$.[14] Again for convenience, define

$$Q(\{i\}) \equiv (Q^{\omega_1}(\{i\}), \dots, Q^{\bar{\omega}}(\{i\}))^T,$$

$$Q^{\omega}(\{i\}) = \max_{\hat{a}^{\omega}_i \in \Delta(A^{\omega}_i)} \left\{ K^{\omega}_i(\hat{a}^{\omega}_i, a^{\omega *}_{N\setminus i}) + \delta \sum_{\omega' \in \Omega} p(\omega'|\omega, (\hat{a}^{\omega}_i, a^{\omega *}_{N\setminus i}))E^{\omega'}_i(\eta^{ne}) \right\},$$

and

$$\sigma_i = (\mathbb{I} - \delta\Pi(\eta^*))^{-1}\beta_i \geqslant Q(\{i\}). \quad (70)$$

[14] In the case of multiple Nash equilibria, one of them should be chosen for the realisation of the punishment. Notice that this can be implemented because players use correlated strategies.

The condition of existence of a SPNE are summarised in the following proposition. The validity of inequality (70) implies that the principle of strategic stability holds when the Nash equilibrium is subgame perfect.

Proposition 2. *If, in an σ-regularisation G_σ such that $\bar{\sigma} = \pi_0\sigma$, inequality* (70) *holds for any player $i \in N$, then there exists behaviour strategy profile $\tilde{\varphi}$ which is a SPNE with players' payoffs $(\bar{\sigma}_1,\ldots,\bar{\sigma}_n)$.*

Proof. The proof is similar to the proof of Proposition 1 using the structure of the "new" strategy profile. Determine this behaviour strategy profile as $\tilde{\varphi} = (\tilde{\varphi}_1, \ldots, \tilde{\varphi}_n)$ where strategies $\tilde{\varphi}_i$, $i \in N$ are:

$$\tilde{\varphi}_i(h(k)) = \begin{cases} a_i^{\omega *}, & \text{if } \omega(k) = \omega, h(k) \subset h^*; \\ \hat{a}_i^{\omega,ne}, & \text{if } h(k) \not\subseteq h^*, \end{cases} \tag{71}$$

where $\hat{a}_i^{\omega,ne} \in \Delta(A_i^\omega)$ is player's i's punishment, which can be either in pure or mixed strategies. Notice that, if a multi-player deviation is observed in the history, all players implement $\hat{\eta}^{ne}$.

Irrational-behaviour-proof. Subgame consistency and strategic support assume that the players are fully rational. However, in reality cooperation may be broken down by irrational reasons. For instance, a player may use irrational acts to extort additional gains if some circumstances allow it. Refusal of other players to yield to his extortion would result in the dissolution of the cooperative scheme. Thus in this case, a deviation would imply an "irrational behaviour."[15]

D.W.K. Yeung proposed a condition[16] under which, even if an irrational behaviour emerges in the game, a player is certain to obtain at least her individual payoff (Yeung, 2006). This procedure can be explained as follows. Suppose two different scenarios. In the first scenario, a player cooperates until a certain period, and then the cooperation breaks up. In the second scenario, a player plays individually during the whole game. If the payoff in the first scenario is not less than the payoff in the second scenario, then the principle of irrational behaviour proof is satisfied. The following definition provides the condition to satisfy this principle.

Definition 25. Cooperative solution $\bar{\sigma}$ and the corresponding IDP satisfy the principle of irrational-behaviour-proof if

$$E_i^\sigma[1,k] + \delta^k \Pi^k(\eta^*)V(\{i\}) \geqslant V(\{i\}), \text{ for every } i \in N \text{ and any } k = 1, 2, \ldots, \tag{72}$$

where $E_i^\sigma[1,k]$ is the expected player i's payoff at the first k stages in σ-regularisation G_σ.

The underlying assumption is that, before the beginning of each stage, players know if the cooperation has broken down or not, so that the information is not delayed. In the left hand side of inequality (15), the first term is equal to the expected value of player i's payoff if, in the first k stages, players play cooperative strategy profile η^* and σ-regularization of game G_σ is made. The second term is the expected payoff of player i from stage $k+1$, when the cooperation breaks up. The right hand side of (15) is the payoff of player i if she plays individually from the start onwards.

[15] Note that it is possible to formulate an analogous condition for repeated games.

[16] The so-called Yeung's condition or principle of irrational- behaviour-proof was adopted for linear-quadratic games in (Tur, 2014, Markovkin, 2006).

Theorem 10. *If inequality*

$$(\mathbb{I} - \delta\Pi(\eta^*))(\sigma_i - V(\{i\})) \geqslant 0 \tag{73}$$

holds for any $i \in N$*, then the cooperative solution* $\bar{\sigma}$ *and the corresponding IDP* $\{\beta_i\}_{i\in N}$ *satisfy the principle of irrational-behaviour-proof.*

Proof. In what follows, we show that condition (73) is sufficient for inequality (15) to hold for any $k = 1, 2, \ldots$. The proof is based on the mathematical induction method. First, we rewrite (15) for $k = 1$. Then we transform (73) by considering definition σ_i and using IDP (56). We get

$$V(\{i\}) \leqslant \beta_i + \delta\Pi(\eta^*)V(\{i\}). \tag{74}$$

The inequalities coincide and it proves Theorem for $k = 1$.

Suppose that (73) implies (15) for $k = l$. Rewriting (15) for $k = l$ we yield:

$$V(\{i\}) \leqslant \beta_i + \ldots + \delta^{l-1}\Pi^{l-1}(\eta^*)\beta_i + \delta^l\Pi^l(\eta^*)V(\{i\}). \tag{75}$$

We adopt the same procedure for $k = l + 1$. Inequality (15) for $k = l + 1$ is:

$$V(\{i\}) \leqslant \beta_i + \ldots + \delta^l\Pi^l(\eta^*)\beta_i + \delta^{l+1}\Pi^{l+1}(\eta^*)V(\{i\}). \tag{76}$$

Next we need to prove that, if (73) holds, then (15) holds for $k = l + 1$. After transformation the right hand side of (76) is:

$$\beta_i + \delta\Pi(\eta^*)\left\{\beta_i + \delta\Pi(\eta^*)\beta_i + \ldots + \delta^{l-1}\Pi^{l-1}(\eta^*)\beta_i + \delta^l\Pi^l(\eta^*)V(\{i\})\right\}.$$

Taking into account (75), the expression in braces is not less than $V(\{i\})$. Therefore the right part of (76) is not less than $\beta_i + \delta\Pi(\eta^*)V(\{i\})$. From equation (53) and (73), we get (15) for $k = l + 1$, which proves the theorem.

Corollary 1. *For irrational-behaviour-proof principle it is sufficient that for each* $i \in N$ *the following inequality is true:*

$$K_i(\tilde{a}) - \beta_i \leqslant \delta\left(\sigma_i^{\min} - V^{\max}(\{i\})\right), \tag{77}$$

where $K_i(\tilde{a}) = \left(\max\limits_{a_i^{\omega_1}\in A_i^{\omega_1}} K_i^{\omega_1}(a^{\omega_1 *}\|a_i^{\omega_1}), \ldots, \max\limits_{a_i^{\bar{\omega}}\in A_i^{\bar{\omega}}} K_i^{\bar{\omega}}(a^{\bar{\omega}*}\|a_i^{\bar{\omega}})\right)^T$*, and* $\max\limits_{a_i^{\omega}\in A_i^{\omega}} K_i^{\omega}(a^{\omega *}\|a_i^{\omega})$ *is the maximal payoff of player* i *which he obtains deviating from action profile* $a^{\omega *}$ *which is the part of cooperative strategy profile* η^* *satisfying condition* (6)*, and* $\tilde{a} = \arg\max\limits_{a_i^{\omega}\in A_i^{\omega}} K_i^{\omega}(a^{\omega *}\|a_i^{\omega})$ *for each state* $\omega \in \Omega$ *and each player* $i \in N$*:*

$$\sigma_i^{\min} = \left(\min_{\omega\in\Omega}\sigma_i^{\omega}, \ldots, \min_{\omega\in\Omega}\sigma_i^{\omega}\right)^T,$$

$$V^{\max}(\{i\}) = \left(\max_{\omega\in\Omega} V^{\omega}(\{i\}), \ldots, \max_{\omega\in\Omega} V^{\omega}(\{i\})\right)^T.$$

Proof. Let sufficient condition (77) be satisfied. It can be rewritten in the following way:

$$\beta_i + \delta\sigma_i^{\min} \geqslant K_i(\tilde{a}) + \delta V^{\max}(\{i\}). \tag{78}$$

Estimate the left- and right- hand parts of inequality (78). As matrix of transition probabilities $\Pi(\eta^*)$ is stochastic, we obtain:

$$\beta_i + \delta\sigma_i^{\min} = \beta_i + \delta\Pi(\eta^*)\sigma_i^{\min} \leqslant \beta_i + \delta\Pi(\eta^*)\sigma_i. \tag{79}$$

For the right-hand side of inequality (78), the equity is true:

$$K_i(\tilde{a}) + \delta V^{\max}(\{i\}) = K_i(\tilde{a}) + \delta\Pi(\tilde{\tilde{\eta}})V^{\max}(\{i\}), \tag{80}$$

where $\Pi(\tilde{\tilde{a}})$ is a stochastic matrix, and $\tilde{\tilde{\eta}} = (\tilde{\tilde{\eta}}_i : i \in N)$ is a profile in stationary strategies such that

$$\tilde{\tilde{\eta}}_j = \begin{cases} \arg\max\limits_{\eta_i \in H_i} \Pi(\eta^* \| \eta_i)V(\{i\}), & \text{if } j = i \\ \eta_j^*, & \text{if } j \neq i \end{cases}$$

Therefore, we have the inequality:

$$K_i(\tilde{a}) + \delta\Pi(\tilde{\tilde{\eta}})V^{\max}(\{i\}) = \max_{a_i \in A_i} K_i(a^* \| a_i) + \delta \max_{\eta_i \in \Xi_i} \{\Pi(\eta^* \| \eta_i)V(\{i\})\} \geqslant \max_{\eta_i \in H_i} \{K_i(a^* \| a_i) + \delta\Pi(\eta^* \| \eta_i)V(\{i\})\}. \tag{81}$$

The inequalities (78), (79), (80) and (81) implies condition (73). Therefore, by Theorem 1 the principle of irrational-behaviour-proof is satisfied.

3.5. Existence of stable cooperative solution

In this section we discuss the conditions guaranteeing the existence of a stable cooperative solution. First, we need to mention that the allocation rule adopted should give a non-empty subset of the imputation set. Cooperative solutions such as the Shapley value or the nucleolus always exist and we may calculate them for any subgame using the values of the characteristic function given by (7), (9) and (51).

The existence of a subgame consistent cooperative solution follows from Theorem 9 and the method of construction of IDP for $\bar{\sigma}$. For a given cooperative solution $\bar{\sigma}$, the regularisation of a stochastic game determines new payoff functions to players in order to satisfy the principle of subgame consistency. Hence, the players' discounted payoffs in σ-regularisation of the initial game are equal to the components of cooperative solution $\bar{\sigma}$, which is subgame consistent.

Thus, if the payments to the players are modified through σ-regularisation, then subgame consistent cooperative solution $\bar{\sigma}$ exists in general.

To verify whether cooperative solution $\bar{\sigma}$ satisfies the principle of strategic stability and irrational-behaviour-proof, we need to check that the following system of inequalities holds:

$$\begin{cases} \sigma_i = (\mathbb{I} - \delta\Pi(\eta^*))^{-1}\beta_i \geqslant F(\{i\}),\ i \in N, \\ (\mathbb{I} - \delta\Pi(\eta^*))(\sigma_i - V(\{i\})) \geqslant 0,\ i \in N. \end{cases} \tag{82}$$

These conditions on discount factor δ are similar to those necessary to prove that a cooperative strategy profile is SPNE in repeated games. This system is non-linear with respect to δ and the solution of the system cannot be obtained in an explicit form.

However, we may state the existence of a stable cooperative solution for the class of stochastic games in which the cooperative strategy profile coincides with the Nash equilibrium and the players are symmetric. In this case, the Shapley value satisfies the principles of stable cooperation. Further, we examine the solution of system (82) on a specific class of stochastic games with two states and two players.

Example 3. Stochastic game of competition between asymmetric firms.

Noncooperative game. Consider Cournot duopoly with asymmetric firms. Describe it with a stochastic game setting like Prisoners' Dilemma. Let the set of states be $\Omega = \{\omega_1, \omega_2\}$, where $\omega_j = \langle N, A_1^{\omega_j}, A_2^{\omega_j}, K_1^{\omega_j}, K_2^{\omega_j} \rangle$, $j = 1, 2$, and $A_i^{\omega_j} = \{C_j, D_j\}$ is the set of actions of player $i = 1, 2$. Strategies C_j and D_j stands for "collude" and "deviate", respectively. For state ω_1, players' payoffs are:

$$\begin{array}{cc} & \begin{array}{cc} C_1 & D_1 \end{array} \\ \begin{array}{c} C_1 \\ D_1 \end{array} & \begin{pmatrix} (7,7) & (1,8) \\ (8,1) & (4,5) \end{pmatrix} \end{array}$$

whereas for state ω_2 players' payoffs are:

$$\begin{array}{cc} & \begin{array}{cc} C_2 & D_2 \end{array} \\ \begin{array}{c} C_2 \\ D_2 \end{array} & \begin{pmatrix} (9,9) & (1,10) \\ (16.5,1) & (6,5) \end{pmatrix} \end{array}$$

State ω_1 can be interpreted as a market with a low demand, and state ω_2 as a market with a high demand. Both one-shot games have the unique Nash equilibrium when both firms deviate with outcomes $(4, 5)$ and $(6, 5)$ in states ω_1 and ω_2 respectively. Conversely, the cooperative action profile that maximizes the sum of the payoffs are "to collude" with outcomes $(7, 7)$ and $(9, 9)$ respectively. When playing the cooperative action profile, players get equal payoffs, but in the Nash equilibrium outcome they obtain asymmetric payoffs. In particular, with a low demand Firm 1 has a lower payoff than Firm 2, and with a high demand Firm 2 has lower payoff than Firm 1. This scenario could be interpreted as the result of technical features of firms' production. For instance, Firm 2 can be endowed with a production technology being more efficient in producing low levels of output.

In state ω_2, players also differ in the profiles when one firm "colludes" and the competitor "deviates". In particular, Firm 1's deviation payoff is larger than Firm 2's one. Hence the asymmetry of the players influences the cooperative payoff imputation. Another feature of state ω_2 is that, when both firms collude, their summarized payoff is not much larger the one in action profile (D_2, C_2) (18 against 17.5). Therefore, if the probability of transiting from profile (D_2, C_2) to state ω_1 is larger than from profile (C_2, C_2), then players may agree on playing profile (D_2, C_2) to avoid transition from high to low demand state.

Let transition probabilities from states ω_1 and ω_2 be

$$\begin{pmatrix} (0.3, 0.7) & (0.9, 0.1) \\ (0.4, 0.6) & (0.3, 0.7) \end{pmatrix}, \quad \begin{pmatrix} (0.9, 0.1) & (0.4, 0.6) \\ (0.1, 0.9) & (0.3, 0.7) \end{pmatrix}$$

where the element (k, l) of the matrix consists of transition probabilities from state ω_j to states ω_1, ω_2, on condition that player 1 chooses actions k^{th} and player 2 chooses l^{th}. We may mention that the probability of transiting to state ω_1 in action profile (C_2, C_2) is much higher than the probability to transit to this state in action profile (D_2, C_2), that is 0.9 contrary to 0.1. Let the discount factor be $\delta = 0.99$ and the vector of the initial distribution over the set of states be $\pi_0 = (0.5, 0.5)$.

Cooperative game. Determine cooperative game G_c based on stochastic game G. For it, we compute cooperative solution $\eta^* = (\eta_1^*, \eta_2^*)$ in stationary strategies using (5) and (6). We obtain a unique stationary strategy $\eta_1^* = (C_1, D_2)$ for player 1, and $\eta_2^* = (C_1, C_2)$ for player 2 which give maximal total players' payoff $\bar{V}(\{1,2\}) = \pi_0 V(\{1,2\}) = 1704.61$. Following this profile, in state ω_1 the profile of cooperative strategies (when both players collude) gives payoff 7 for each firm. In state ω_2, with a cooperative strategy profile, Firm 1 deviates and Firm 2 colludes, and the payoff of firm 2 is less than its payoff in the Nash equilibrium. But this will be compensated by Firm 1 when they apply an imputation of their joint payoff. Therefore, the values of a characteristic function for a grand coalition are

$$V(\{1,2\}) = \begin{pmatrix} E_1^{\omega_1}(\eta^*) + E_2^{\omega_1}(\eta^*) \\ E_1^{\omega_2}(\eta^*) + E_2^{\omega_2}(\eta^*) \end{pmatrix} = \begin{pmatrix} 1702.43 \\ 1706.80 \end{pmatrix}.$$

By definition (51) the values of characteristic function for the empty set are zero:

$$V(\varnothing) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Calculate the values of characteristic function $V(S) = (V^{\omega_1}(S), V^{\omega_2}(S))$ for coalitions $S = \{1\}$ and $S = \{2\}$ using (7):

$$V(\{1\}) = \begin{pmatrix} 538.60 \\ 540.60 \end{pmatrix}, \ V(\{2\}) = \begin{pmatrix} 500.00 \\ 500.00 \end{pmatrix}.$$

These are Firms' payoffs in the Nash equilibrium when both firms deviate in all states, i.e., they adopt strategy profiles (D_1, D_1) and (D_2, D_2).

Using (10), we may calculate $\bar{V}(S)$ for the whole game and all coalitions:

$$\bar{V}(\varnothing) = 0.00, \ \bar{V}(\{1\}) = 539.60, \ \bar{V}(\{2\}) = 500.00, \ \bar{V}(\{1,2\}) = 1704.61.$$

Thus, we determine cooperative stochastic subgame $G_c^{\omega_j}$ as the set $\langle N, V^{\omega_j}(\cdot)\rangle$, $j = 1, 2$, and cooperative stochastic game G_c as the set $\langle N, \bar{V}(\cdot)\rangle$.

Cooperative solution: the Shapley value. We suppose that players choose the Shapley value as a cooperative solution of their total payoff in cooperative stochastic game G_c and in all subgames $G_c^{\omega_j}$, $j = 1, 2$. For two-player game the Shapley value is calculated by formula:

$$\sigma_i^{\omega_j} = V^{\omega_j}(\{i\}) + \frac{V^{\omega_j}(\{1,2\}) - V^{\omega_j}(\{1\}) - V^{\omega_j}(\{2\})}{2},$$

where $i = 1, 2$ and $j \in \{1, 2\}$, $j \neq i$. The Shapley values in subgames are

$$\sigma_1 = \begin{pmatrix} 870.516 \\ 873.698 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 831.916 \\ 833.098 \end{pmatrix}.$$

Then taking into account the vector of initial distribution π_0, we are able to determine the Shapley value $\bar{\sigma}$ in the whole game G_c by Definition 20:

$$\bar{\sigma} = (\bar{\sigma}_1, \bar{\sigma}_2) = (872.107, 832.507).$$

Subgame consistency. Now we verify if the Shapley value satisfies the principles of stable cooperation and begin with subgame consistency. If firms receive stage payoffs according to their initially defined payoffs, then their discounted payoffs in the whole game will be equal 1526.809 and 177.805 which are different from the components of the Shapley value which are 872.107 and 832.507. Define the imputation distribution procedure or transfer payments to the players such that they finally receive the components of the Shapley value and the imputation distribution procedure is subgame-consistent. Using that $\bar{\sigma}$ equals $\pi_0\sigma$ and equation (53), we obtain IDP:

$$\beta_1 = (\mathbb{I} - \delta \Pi(\eta^*))\sigma_1 = \begin{pmatrix} 6.500 \\ 9.052 \end{pmatrix}, \quad \beta_2 = (\mathbb{I} - \delta \Pi(\eta^*))\sigma_2 = \begin{pmatrix} 7.500 \\ 8.448 \end{pmatrix}.$$

Define σ-regularisation of initial stochastic game G using IDP and Definition 23. We redefine payoff functions of the players in the initial game in all states when players adopt cooperative strategy profile substituting the payoffs by corresponding components of the IDP. Therefore, players' payoffs in states ω_1 and ω_2 correspondingly equal:

$$\begin{pmatrix} (6.500, 7.500) & (1,8) \\ (8,1) & (4,5) \end{pmatrix},$$

$$\begin{pmatrix} (9,9) & (1,10) \\ (9.052, 8.448) & (6,5) \end{pmatrix}.$$

In a regularized game in the state with a low demand (state ω_1), both firms adopt action "to collude" and receive payoffs $(6.5, 7.5)$. Notice that their payoffs in the initial game are $(7,7)$. Therefore, Firm 1 gives 0.5 to Firm 2. In the state with a high demand (state ω_2) Firm 1 plays action "to deviate" while Firm 2 "colludes". This behavior is prescribed by the cooperative strategy profile. In state ω_2 players' payoffs are $(9.052, 8.448)$. Notice that the payoffs in the initial game are $(16.5, 1)$. Therefore, Firm 1 gets $16.5 - 9.052 = 7.448$ from Firm 2. If the regularization of the initial game is made by the above described method, the Shapley value and the corresponding IDP are subgame-consistent.

Strategic support. We now check for strategic support of the Shapley value, i. e., we check if Firms have benefits from individual deviations from the cooperative strategy profile. First, consider state ω_1. As the action profile played in cooperation is not the Nash equilibrium, then the players may have benefits from deviation. We verify if the inequality is true:

$$\sigma_i^{\omega_1} \geqslant F^{\omega_1}(\{i\}),$$

for each $i = 1, 2$, where

$$F^{\omega_1}(\{i\}) = \max_{\substack{a_i^{\omega_1} \in A_i^{\omega_1} \\ a_i^{\omega_1} \neq a_i^{\omega_1 *}}} \left\{ K_i^{\omega_1}(a_i^{\omega_1}, a_{N\setminus i}^{\omega_1 *}) + \delta \sum_{\omega' \in \Omega} p(\omega' | \omega_1, (a_i^{\omega_1}, a_{N\setminus i}^{\omega_1 *})) V^{\omega'}(\{i\}) \right\}.$$

Inequality (14) for Firm 1 is written in this way:

$$870.516 \geqslant 8 + 0.99\,(0.4\ 0.6) \begin{pmatrix} 538.60 \\ 540.60 \end{pmatrix} = 542.402,$$

and for Firm 2:

$$831.916 \geqslant 8 + 0.99\,(0.9\ 0.1)\begin{pmatrix}500.00\\500.00\end{pmatrix} = 503.$$

In state ω_2, cooperative action profile (D_2, C_2) is the Nash equilibrium. Therefore, players can't increase their payoffs by deviations. Therefore, we may conclude that inequality (14) holds for state ω_2. The condition of strategic support is satisfied.

Irrational-behavior-proof. To verify the condition of irrational behavior proof, we need to compare players' payoffs in two cases:

1) A firm plays individually during the whole game,
2) A firm cooperates with the other firm until some step, and after this it starts playing individually.

Notice that in the second case, when the firms cooperate, they receive payoffs in accordance with IDP, constructed on the basis of thee initially chosen cooperative solution.

If the player's payoff in case 1) is not greater than his payoff in case 2), then the principle of irrational behavior proof against irrational behavior is satisfied. This has been proved in Proposition 1, since

$$(\mathbb{I} - \delta\Pi(\eta^*))(\sigma_1 - V(\{1\})) = (\mathbb{I} - \delta\Pi(\eta^*))(\sigma_2 - V(\{2\})) = \begin{pmatrix}2.500\\3.448\end{pmatrix} \geqslant 0.$$

3.6. Strong transferable equilibrium

Theorem 1 can be generalized to the case when several players deviate, i. e., we may prove that if the condition similar to inequality (14) is satisfied in σ-regularization G_σ of stochastic game G, there exists a strong transferable equilibrium with payoffs $(\bar{\sigma}_1, \ldots, \bar{\sigma}_n)$. In this case, players can implement a specially constructed profile in trigger strategies, where as a punishment for deviated coalition, not deviated players will implement trigger strategies that allow a deviated coalition to obtain a minimax payoff in any subgame. Define a strong transferable equilibrium and prove a theorem similar to Theorem 1.

Definition 26. (Petrosyan and Kuzutin, 2000) We call profile $\tilde{\varphi} = (\tilde{\varphi}_1, \ldots, \tilde{\varphi}_n)$ strong transferable equilibrium in regularized game G_σ if for any coalition $S \subseteq N$, $S \neq \varnothing$, inequality

$$\sum_{i\in S} \bar{E}_i^\sigma(\tilde{\varphi}) \geqslant \sum_{i\in S} \bar{E}_i^\sigma(\tilde{\varphi} \parallel \varphi_S) \tag{83}$$

holds for any behaviour strategy of coalition S: $\varphi_S = (\varphi_i : i \in S) \in \prod_{i\in S} \Phi_i$. Here $\bar{E}_i^\sigma(\cdot)$ is a discounted payoff of player i in σ-regularisation of game G.

We will prove a theorem allowing us to obtain a condition on the game parameters for which in regularized game G_σ there exists a transferable equilibrium with players' payoffs equal to the corresponding components of the cooperative solution according to which the initial stochastic game is regularized.

Theorem 11. *If in regularized game G_σ such that cooperative solution satisfies condition $\bar{\sigma} = \pi_0\sigma$, the inequality holds:*

$$\sum_{i\in S} \beta_i \geqslant (\mathbb{I} - \delta\Pi(\eta^*))\tilde{F}(S) \tag{84}$$

for any coalition $S \subseteq N$, $S \neq \varnothing$, where $\tilde{F}(S) = (\tilde{F}^{\omega_1}(S), \ldots, \tilde{F}^{\bar{\omega}}(S))^T$,
$\tilde{F}^{\omega}(S) = \max\limits_{\substack{a_S^{\omega} \in \prod_{i \in S} \Delta(A_i^{\omega}) \\ a_S^{\omega} \neq a_S^{\omega *}}} \left\{ \sum_{i \in S} K_i^{\omega}(a^{\omega *} \parallel a_S^{\omega}) + \delta \sum_{\omega' \in \Omega} p(\omega'|\omega, a^{\omega *}) V^{\omega'}(S) \right\}$*, then in regularized game G_σ there exists a strong transferable equilibrium with players' payoffs $(\bar{\sigma}_1, \ldots, \bar{\sigma}_n)$.*

Proof. The proof of the theorem is close the proof of Theorem 1 but instead of strategy (63) we use the following behaviour strategy φ_i', $i \in N$:

$$\varphi_i'(h(k)) = \begin{cases} a_i^{\omega *}, & \text{if } \omega(k) = \omega, h(k) \subset h^*; \\ a_i^{\omega\prime}(S), & \text{if } \omega(k) = \omega \text{ and } \exists\, l \in [1, k-1], \\ & S \subset N,\ i \notin S\text{: } h(l) \subset h^* \text{ and} \\ & (\omega(l), a(l)) \notin h^*, \text{ but} \\ & (\omega(l), (a_S^*(l), a_{N \setminus S}(l)) \in h^*; \\ \text{any} & \text{otherwise,} \end{cases} \tag{85}$$

where $a_i^{\omega *}$ is an action of player i in a cooperative mode, while $a_i^{\omega\prime}(S) \in \Delta(A_i^{\omega})$ is an action of player i in a trigger mode which jointly with actions $a_{i'}^{\omega\prime}(S) \in \Delta(A_{i'}^{\omega})$ of players $i' \neq i$, $i' \in N \setminus S$ forms an action of coalition $N \setminus S$ against coalition S and allows coalition S to obtain minmax value $V^{\omega}(S)$ in subgame G^{ω}.

3.7. Strongly subgame consistency of the core

Now suppose that the solution of a cooperative stochastic game is the subset of the imputation set that contains more than one point. For definiteness, let such a solution be the core. We formulate the problem of strongly subgame consistency of the core and propose sufficient conditions for strongly subgame consistency of the core for stochastic games with infinite duration given by (1).

Suppose that the cores of stochastic game G_c and any subgame G_c^{ω}, $\omega \in \Omega$, are nonempty. 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 $\overline{CO}$. Reaching intermediate states $\omega \in \Omega$, player $i \in N$ chooses action $a_i^{\omega *}$ in accordance with cooperative strategy η_i^* and gets payoff $K_i^{\omega}(a^{\omega *})$. If the players recalculate the solution, i.e., they find a solution of cooperative subgame G_c^{ω}, then the current solution will be the core CO^{ω}. 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 $CO^{\omega'}$, $\omega' \in \Omega$, following state ω, would be an imputation from the core CO^{ω}. If this property holds for any state $\omega \in \Omega$, then the core of cooperative stochastic game G_c is strongly subgame-consistent.

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

$$EC(\omega) = \left\{ \sigma(\omega) = \sum_{\omega' \in \Omega} p(\omega'|\omega, a^{\omega *}) \sigma^{\omega'} \mid \sigma^{\omega'} \in CO^{\omega'} \right\}. \tag{86}$$

Set $EC(\omega)$ contains vectors $\sigma(\omega)$ which are mathematical expectations of all possible sets of the imputations from the cores of subgames starting in states which are realised after the current state with respect to probability distribution $\{p(\omega'|\omega, a^{\omega*}), \omega' \in \Omega\}$.

Remind Definition 21 of the imputation distribution procedure. The first condition (52) in a definition can be called the condition of "attainability of the imputation distribution procedure" because it allows to ensure that in any realized state the sum of payments to the players is equal to the sum of their payoffs when they implement cooperative strategies. The second condition guarantees players to receive the components of the initially chosen imputation from the core of cooperative game G_c in the sense of mathematical expectation, if payments to the players throughout the game will be made in accordance with distribution procedure $\{\beta^\omega : \omega \in \Omega\}$.

We now define the distribution procedure of imputation $\bar{\sigma} = (\bar{\sigma}_1, \dots, \bar{\sigma}_n)$, where $\bar{\sigma}_i = \pi_0 \sigma_i$, $(\sigma_1^\omega, \dots, \sigma_n^\omega) = \sigma^\omega \in CO^\omega$, such that the core is strongly subgame-consistent.

Definition 27. We call the core $\overline{CO}$ of cooperative stochastic game G_c strongly subgame-consistent if there exists a distribution procedure $\{\beta^\omega : \omega \in \Omega\}$ of the imputation from the core $\overline{CO}$ such that for any state $\omega \in \Omega$ the following inclusions hold:

$$\beta^\omega \oplus \delta EC(\omega) \subset CO^\omega, \tag{87}$$

$$B^\omega \in CO^\omega, \quad \omega \in \Omega \tag{88}$$

where

$$\beta^\omega \oplus \delta EC(\omega) = \Big\{ \beta^\omega + \delta\sigma(\omega) : \sigma(\omega) \in EC(\omega) \Big\}.$$

And distribution procedure $\{\beta^\omega : \omega \in \Omega\}$ is called strongly subgame-consistent.

Condition (87) means that the set of vectors equal to the sum of the imputation distribution procedure of the player in state ω and an imputation from the expected core for this state, is contained in the core of subgame starting from state ω. This condition imposes restrictions on payments to the players in the realized states, and very often is not satisfied for an arbitrary game, if payments to the players are made in accordance with the initially defined payoff functions.

We impose additional restrictions on the characteristic functions of subgames starting from the states of set Ω in order to obtain sufficient conditions for strongly subgame consistency of the core. Denote by $EV^\omega(S)$ the expected value of the characteristic function calculated for coalition $S \subseteq N$ for subgames following state ω:

$$EV^\omega(S) = \sum_{\omega' \in \Omega} p(\omega'|\omega, a^{\omega*}) V^{\omega'}(S).$$

Denote by

$$\Delta V^\omega(S) = V^\omega(S) - \delta EV^\omega(S)$$

the difference between the values of a characteristic function in state ω and the expected value of the characteristic function. We denote by ΔCO^ω an analog of the core constructed with function $\Delta V^\omega(S)$. We formulate a sufficient condition for strongly subgame consistency of IDP and the core $\overline{CO}$.

Theorem 12. *Let for each state $\omega \in \Omega$ the core CO^ω and the set ΔCO^ω be nonempty. If for every state $\omega \in \Omega$ distribution procedure $\{\beta^\omega : \omega \in \Omega\}$ of the imputation from the core $\overline{CO}$ satisfies conditions:*

$$\beta^\omega \in \Delta CO^\omega, \tag{89}$$

$$B^\omega \in CO^\omega, \quad \omega \in \Omega, \tag{90}$$

then the core $\overline{CO}$ and procedure $\{\beta^\omega : \omega \in \Omega\}$ are strongly subgame-consistent.

Proof. We prove that any vector $\beta^\omega \in \Delta CO^\omega$ satisfying conditions (89) and (90) is a strongly subgame-consistent distribution procedure of imputation $\bar{\sigma} \in \overline{CO}$, i. e., conditions (87) and (88) from Definition 27 hold. Condition (90) coincides with (88), so, we need to prove that inclusion (87) holds for each state $\omega \in \Omega$. In state ω consider any vector $\sigma(\omega) \in EC(\omega)$ and find sum $\beta^\omega + \delta\sigma(\omega)$. Now we verify if the latter vector belongs to the core $\overline{CO}$. First, calculate the sum of all components of the vector:

$$\begin{aligned}\sum_{i\in N}\beta_i^\omega + \delta\sum_{\omega'\in\Omega}p(\omega'|\omega,a^{\omega*})\sum_{i\in N}\sigma_i^{\omega'} = \\ = V^\omega(N) - \delta\sum_{\omega'\in\Omega}p(\omega'|\omega,a^{\omega*})V^{\omega'}(N) + \\ + \delta\sum_{\omega'\in\Omega}p(\omega'|\omega,a^{\omega*})\sum_{i\in N}\sigma_i^{\omega'} = V^\omega(N),\end{aligned}$$

which means that property of collective rationality holds.

Next, consider $S \subset N$, $S \neq N$:

$$\begin{aligned}\sum_{i\in S}\beta_i^\omega + \delta\sum_{\omega'\in\Omega}p(\omega'|\omega,a^{\omega*})\sum_{i\in S}\sigma_i^{\omega'} \geqslant \\ \geqslant V^\omega(S) + \delta\sum_{\omega'\in\Omega}p(\omega'|\omega,a^{\omega*})V^{\omega'}(S) - \\ - \delta\sum_{\omega'\in\Omega}p(\omega'|\omega,a^{\omega*})V^{\omega'}(S) = V^\omega(S).\end{aligned}$$

Since the choice of state $\omega \in \Omega$ is random, we conclude that the core of the game G_c and procedure $\{\beta^\omega : \omega \in \Omega\}$ are strongly subgame-consistent.

When analogs of the cores ΔCO^ω are nonempty for any state ω, Theorem 12 provides a method of construction of a strongly subgame-consistent distribution procedure of imputations from the core. Notice that generally not all imputations from the core can be realised with distribution procedure $\{\beta^\omega : \omega \in \Omega\}$ described above.

3.8. Stochastic game with one absorbing state

Noncooperative game. In this section we consider a two-player game with two states. The set of players is $N = \{1, 2\}$. Let state ω_1 be given by:

$$\omega_1: \quad \begin{array}{c} \\ C \\ D \end{array}\!\!\begin{array}{c} \begin{array}{cc} \quad C \quad & \quad D \quad \end{array} \\ \begin{pmatrix} (a, a+1) & (c, b) \\ (b, c) & (d+1, d) \end{pmatrix} \end{array} \tag{91}$$

Players have two pure actions, C (to cooperate) and D (to defect). The constants satisfy the inequalities:

$$b > a + 1, \qquad a > d + 1, \qquad d > c > 0.$$

We also assume

$$2a + 1 > b + c. \tag{92}$$

From inequality (92) it follows that players receive a larger total payoff by cooperating than defecting. The game represents Prisoners' Dilemma with asymmetric players: in action profile (C, C) the payoff of player 1 is less than the payoff of player 2, but in profile (D, D) the payoff of player 2 is less than the payoff of player 1. If action profiles (C, C) and (D, D) are chosen in state ω_1, a stochastic game remains in this state with probability 1. But if profiles (C, D) or (D, C) are chosen, the game transits to state ω_2 which is "absorbing", i.e. this state will be realised in all following stages of the game with probability 1. In state ω_2 both players have a unique action D and their payoffs will be equal to d:

$$\omega_2 : \qquad \begin{array}{c} D \\ D\,(d, d) \end{array} \tag{93}$$

The matrices of transition probabilities from states ω_1 and ω_2 are

$$\begin{pmatrix} (1,0) & (0,1) \\ (0,1) & (1,0) \end{pmatrix}, \quad (0,1)\,.$$

The discount factor is $\delta \in (0, 1)$ and the vector of the initial distribution on the set of states is $\pi_0 = (1, 0)$, i.e., a game starts with state ω_1.

Cooperative game. For this game we construct a cooperative game by determining the characteristic functions for all subgames and the whole game. We then show how we need to redistribute the stage payoffs adopting IDP to obtain the subgame consistency of the Shapley value. The condition of strategic stability gives the lower bound of the discount factor.

The first step is to determine cooperative form G_c of non-cooperative stochastic game G. In particular, we need to find a cooperative strategy profile and then calculate the values of characteristic functions for each subgame (starting from states ω_1 and ω_2) and for the whole game.

We compute cooperative strategy profile $\eta^* = (\eta_1^*, \eta_2^*)$ using (5) and (6). In a cooperative strategy profile both players choose C in state ω_1 and D in state ω_2. The total players' payoff with profile η^* is equal to the value of the characteristic function of coalition N:

$$\bar{V}(\{1,2\}) = \bar{E}_1(\eta^*) + \bar{E}_2(\eta^*) = 2a + 1 + \delta(2a + 1) + \ldots = \frac{2a+1}{1-\delta}. \tag{94}$$

In particular, the values of characteristic function $V^{\omega}(\{1, 2\})$ for both subgames are

$$V(\{1,2\}) = \begin{pmatrix} V^{\omega_1}(\{1,2\}) \\ V^{\omega_2}(\{1,2\}) \end{pmatrix} = \begin{pmatrix} \dfrac{2a+1}{1-\delta} \\ \dfrac{2d}{1-\delta} \end{pmatrix}. \tag{95}$$

We can now calculate the values of characteristic functions of coalitions $\{1\}$ and $\{2\}$ for both states using (9):

$$\begin{aligned}
V^{\omega_1}(\{1\}) &= \max_{\eta_1}\min_{\eta_2} E_1^{\omega_1}(\eta_1,\eta_2) = \min_{\eta_2}\max_{\eta_1} E_1^{\omega_1}(\eta_1,\eta_2) = \frac{d+1}{1-\delta},\\
V^{\omega_1}(\{2\}) &= \max_{\eta_2}\min_{\eta_1} E_2^{\omega_1}(\eta_1,\eta_2) = \min_{\eta_1}\max_{\eta_2} E_2^{\omega_1}(\eta_1,\eta_2) = \frac{d}{1-\delta},\\
V^{\omega_2}(\{1\}) &= V^{\omega_2}(\{2\}) = \frac{d}{1-\delta}.
\end{aligned}$$

By equation (51), the values of the characteristic functions for the empty set are zero:

$$V(\varnothing) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Using (10), we then calculate the values of the characteristic function $\bar{V}(\cdot)$ for all possible coalitions taking into account the initial distribution of states $\pi_0 = (1,0)$:

$$\bar{V}(\emptyset) = 0,\ \bar{V}(\{1\}) = \frac{d+1}{1-\delta},\ \bar{V}(\{2\}) = \frac{d}{1-\delta},\ \bar{V}(\{1,2\}) = \frac{2a+1}{1-\delta}.$$

In this way, we determine cooperative stochastic subgames $G_c^{\omega_j}$ as the set $\langle N, V^{\omega_j}(\cdot)\rangle$, $j=1,2$, and cooperative stochastic game G_c as the set $\langle N, \bar{V}(\cdot)\rangle$.

The Shapley value. We assume that players choose the Shapley value as an imputation of their total payoff in cooperative stochastic game G_c and in all subgames $G_c^{\omega_j}$, $j=1,2$. For a two-person game, this is given by:

$$\sigma_i^{\omega_j} = V^{\omega_j}(\{i\}) + \frac{V^{\omega_j}(\{1,2\}) - V^{\omega_j}(\{1\}) - V^{\omega_j}(\{2\})}{2},$$

where $i=1,2$ and $j \neq i$. The Shapley values for the subgames are:

$$\sigma_1 = \begin{pmatrix} \sigma_1^{\omega_1} \\ \sigma_1^{\omega_2} \end{pmatrix} = \begin{pmatrix} \dfrac{a+1}{1-\delta} \\ \dfrac{d}{1-\delta} \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} \sigma_2^{\omega_1} \\ \sigma_2^{\omega_2} \end{pmatrix} = \begin{pmatrix} \dfrac{a}{1-\delta} \\ \dfrac{d}{1-\delta} \end{pmatrix}.$$

Taking into account the vector of initial distribution π_0, we are able to determine the Shapley value $\bar{\sigma}$ in game G_c by Definition 20:

$$\bar{\sigma} = (\bar{\sigma}_1, \bar{\sigma}_2) = \left(\frac{a+1}{1-\delta}, \frac{a}{1-\delta}\right).$$

Subgame consistency of the Shapley value. We are now in a position to verify the principles of stable cooperation. Begin with subgame consistency. If players get payoffs according to the initially defined payoff functions, their total payoffs will be $\frac{a}{1-\delta}$ and $\frac{a+1}{1-\delta}$ in contrast to the components of the Shapley value $\frac{a+1}{1-\delta}$ and $\frac{a}{1-\delta}$. In order to obtain subgame consistency, we compute IDP by equating $\bar{\sigma}$ to $\pi_0\sigma$ by using (53):

$$\beta_1 = (\mathbb{I} - \delta \Pi(\eta^*))\sigma_1 = \begin{pmatrix} 1-\delta & 0 \\ 0 & 1-\delta \end{pmatrix} \begin{pmatrix} \dfrac{a+1}{1-\delta} \\ \dfrac{d}{1-\delta} \end{pmatrix} = \begin{pmatrix} a+1 \\ d \end{pmatrix},$$

$$\beta_2 = (\mathbb{I} - \delta \Pi(\eta^*))\sigma_2 = \begin{pmatrix} 1-\delta & 0 \\ 0 & 1-\delta \end{pmatrix} \begin{pmatrix} \dfrac{a}{1-\delta} \\ \dfrac{d}{1-\delta} \end{pmatrix} = \begin{pmatrix} a \\ d \end{pmatrix}.$$

We then determine σ-regularisation of the initial stochastic game G using the IDP and Definition 23. We re-establish the payoff functions of the initial game in state ω_1 when players adopt the cooperative action profiles. Therefore, the players' payoffs in state ω_1 are:

$$\begin{pmatrix} (a+1,a) & (c,b) \\ (b,c) & (d+1,d) \end{pmatrix}.$$

In state ω_1, when both players adopt the cooperative strategy profile (both players use action C in state ω_1), their payoffs are $(a+1,a)$. Since their payoffs in the initial game were $(a,a+1)$, player 2 transfers 1 to player 1. If the initial game is regularised by the method described above, the Shapley value and the corresponding PDP satisfy the principle of subgame consistency (see Theorem 9).

Strategic support of the Shapley value. We now evaluate the strategic support of the Shapley value by checking if players may deviate from the cooperative strategy profile. We consider the possible deviations of players in state ω_1 (in state ω_2 players have the unique action). In this state the cooperative action profile is not the Nash equilibrium, thus players may benefit from deviation. We should check if the following inequality

$$\sigma_i^{\omega_1} \geqslant F^{\omega_1}(\{i\}), \tag{96}$$

is true for any $i = 1,2$, where

$$F^{\omega_1}(\{i\}) = \max_{a_i^{\omega_1} \in \Delta(A_i^{\omega_1})} \left\{ K_i^{\omega_1}(a_i^{\omega_1}, a_{N\setminus i}^{\omega_1 *}) + \delta \sum_{\omega' \in \Omega} p(\omega' | \omega_1, (a_i^{\omega_1}, a_{N\setminus i}^{\omega_1 *})) V^{\omega'}(\{i\}) \right\}.$$

For player 1, inequality (96) yields:

$$\frac{a+1}{1-\delta} \geqslant b + \delta d + \delta^2 d + \ldots = b + \frac{\delta d}{1-\delta},$$

for player 2:

$$\frac{a}{1-\delta} \geqslant b + \delta d + \delta^2 d + \ldots = b + \frac{\delta d}{1-\delta}.$$

These two inequalities give the condition on δ when the principle of strategic support is satisfied:

$$\delta \geqslant \frac{b-a}{b-d}.$$

Principle of irrational-behaviour-proof. In order to verify irrational-behaviour proof, we need to compare the payoffs of each player when:

1) A player acts as an "individual player" during the whole game.
2) A player cooperates with a competitor until some stage and then plays individually.

If the payoff of 2) is not less than the payoff of 1), then this principle is satisfied. This is confirmed by Theorem 1, since:

$$(\mathbb{I} - \delta\Pi(\eta^*))(\sigma_1 - V(\{1\})) = \begin{pmatrix} 1-\delta & 0 \\ 0 & 1-\delta \end{pmatrix} \begin{pmatrix} \dfrac{a+1}{1-\delta} - \dfrac{d+1}{1-\delta} \\ \dfrac{d}{1-\delta} - \dfrac{d}{1-\delta} \end{pmatrix} = \begin{pmatrix} a-d \\ 0 \end{pmatrix} \geqslant 0,$$

$$(\mathbb{I} - \delta\Pi(\eta^*))(\sigma_2 - V(\{2\})) = \begin{pmatrix} 1-\delta & 0 \\ 0 & 1-\delta \end{pmatrix} \begin{pmatrix} \dfrac{a}{1-\delta} - \dfrac{d}{1-\delta} \\ \dfrac{d}{1-\delta} - \dfrac{d}{1-\delta} \end{pmatrix} = \begin{pmatrix} a-d \\ 0 \end{pmatrix} \geqslant 0.$$

Both players benefit from cooperation even if IDP is adopted initially at some stages and then the game is played as a non-cooperative one with initially defined payoff functions as compared with a game played individually by both players during the whole game.

Results. To sum up, we can formulate the conditions under which the Shapley value in the described stochastic game satisfies the three principles of stable cooperation (subgame consistency, strategic support, irrational-behavior-proof):

1. A discount factor is to be $\delta \geqslant \frac{b-a}{b-d}$.
2. A stochastic game is σ-regularised, i. e., the players' payoffs in state ω_1 are:

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{cc} C & D \\ \left(\begin{array}{c} (a+1,a) \\ (b,c) \end{array} \right. & \left. \begin{array}{c} (c,b) \\ (d+1,d) \end{array} \right) \end{array}$$

and in state ω_2 they must not be changed.

4. Conclusion

The paper summarizes the results on cooperative stochastic games with finite and infinite duration based on the author's and coauthors' publications. Section 2 is devoted to describing cooperative stochastic games with finite duration and considering some properties of cooperative solutions applying in dynamics. Section 3 contains a method of construction of a cooperative stochastic game with infinite duration. The principles of stable cooperation in these class of games are examined in this section. There are several numerical examples representing theoretical results. For the applications of theoretical results see the following publications (Bure and Parilina, 2017, Parilina, 2009, Parilina, 2008, Parilina and Sedakov, 2015).

Acknowlegments. The author thanks the Russian Science Foundation (project no. 17-11-01079) for financial support.

References

Avrachenkov, K., Cottatellucci, L. and Maggi, L. (2013). *Cooperative Markov decision processes: Time consistency, greedy players satisfaction, and cooperation maintenance.* International Journal of Game Theory, **42(1)**, 239–262.

Aumann, R. J. and Peleg, B. (1960). *Von Neumann-Morgenstern Solutions to Cooperative Games without Side Payments.* Bulletin of the American Mathematical Society, **66**, 173–179.

Baranova, E. M. (2006). *The Condition for Keeping Cooperation in Stochastic Cooperative Games.* Proceedings of the Russian-Finnish Graduate School Seminar "Dynamic Games and Multicriteria Optimization", edited by V.V. Mazalov, Petrozavodsk, 54–58.

Baranova, E. M. and Petrosjan, L. A. (2006). *Cooperative Stochastic Games in Stationary Strategies.* Game Theory and Applications, **11**, 1–7.

Bellman, R. (1957). *Dynamic programming.* Princeton: Princeton University Press, New Jersey.

Bure, V. M. and Parilina, E. M. (2017). *"Multiple access" game with incomplete information.* Mathematical Game Theory and Its Applications, **9(4)**, 3–17 (in Russian).

Chander, P., Tulkens, H. (1997). *The Core of an Economy with Multilateral Environmental Externalities.* International Journal of Game Theory, **23**, 379–401.

Chistyakov, S. and Petrosyan, L. (2011). *Strong Strategic Support of Cooperative Solutions in Differential Games.* Contributions to Game Theory and Management, **IV**, 105–111.

Driessen, T., Muto, S. and Nakayama, M. (1992). *A cooperative Game of Information Trading: the Core, the Nucleolus and the Kernel.* ZOR Methods and Models of Operations Research, **36(1)**, 55–72.

Filar, J. and Vrieze, K. (1997). *Competitive Markov Decision Processes.* N.Y.: Springer-Verlag New York.

Fink, A. M. (1964). *Equilibrium in a stochastic n-person game.* Journal of Science of the Hiroshima University, **A-I 28**, 89–93.

Gillies, D. B. (1959). *Solutions to general non-zero-sum games.* In: Tucker, A. W.; Luce, R. D. Contributions to the Theory of Games IV. (Annals of Mathematics Studies 40). Princeton: Princeton University Press, 47–85.

Gromova, E. V. and Petrosyan, L. A. (2015). *On a approach of construction of characteristic function in cooperative differential games.* Mathematical Game Theory and Its Applications, **7(4)**, 19–39 (in Russian).

Hörner, J., Rosenberg, D., Solan, E. and Vieille, N. (2010). *On a Markov game with one-sided information.* Operations Research, **58(4 PART 2)**, 1107–1115.

Jáskiewicz, A. and Nowak, A. S. (2016). *Stationary almost Markov perfect equilibria in discounted stochastic games.* Mathematics of Operations Research, **41(2)**, 430–441.

Kohlberg, E. (1971). *On the Nucleolus of a Characteristic Function Game.* SIAM Journal of Applied Mathematics, **20**, 62–66.

Kohlberg, E. (1972). *The Nucleolus as a Solution to a Minimization Problem.* SIAM Journal of Applied Mathematics, **23(1)**, 34–39.

Kohlberg, E. and Neyman, A. (2015). *The cooperative solution of stochastic games.* Harvard Business School. Working Paper, No. 15-071.

Kuhn, H. W. (1950). *Extensive Games.* Proceedings of National Academy of Sciences of the USA, **36**, 570–576.

Kuhn, H. W. (1953) *Extensive Games and the Problem of Information.* Annals of Mathematics Studies, **28**, 193–216.

Markovkin, M. V. (2006). *D. W. K. Yeung's Condition for Linear Quadratic Differential Games.* In: Dynamic Games and Their Applications, eds. L. A. Petrosyan, A. Y. Garnaev, St. Petersburg State University, St. Petersburg, 207–216.

Maschler, M. and Peleg, B. (1976). *Stable sets and stable points of set-valued dynamics system with applications to game theory.* SIAM Journal Control Optim., **14(2)**, 985–995.

WOLFRAM MATHEMATICA. https://www.wolfram.com/mathematica/

Mathworks (2017). https://se.mathworks.com

Meinhardt, H. I. *Graphical Extensions of the Mathematica Package TU Games.* http://library.wolfram.com/ infocenter/MathSource/5709/TuGamesView3D.pdf.

Mertens, J.-F. and Neyman, A. (1981a). *Minimax Theorems for Undiscounted Stochastic Games.* Game Theory and Mathematical Economics, pp. 83–87.

Mertens, J. F. and Neyman, A. (1981b). *Stochastic Games.* International Journal of Game Theory, **10**, 53–66.

Montero, M. (2005). *On the nucleolus as a Power Index.* Homo Oeconomicus, **22(4)**, 551–567.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior.* Princeton: Princeton University Press.

Neyman, A. (2008). *Existence of Optimal Strategies in Markov Games with Incomplete Information.* International Journal of Game Theory, **37(4)**, 581–596.

Neyman, A. (2013). *Stochastic Games with Short-Stage Duration.* Dynamic Games and Applications, **3**, 236–278.

Neyman, A. and Sorin, S. (2003). *Stochastic Games and Applications.* Dordrecht: Kluwer Academic Press.

Nowak, A. S. (1985). *Existence of equilibrium stationary strategies in discounted noncooperative stochastic games with uncountable state space.* Journal of Optimization Theory and Applications, **45(4)**, 591–602.

Nowak, A. S. (1999). *Sensitive equilibria for ergodic stochastic games with countable state spaces.* Mathematical Methods of Operations Research, **50(1)**, 65–76.

Nowak, A. S. and Radzik, T. (1994). *A solidarity value for n-person transferable utility games.* International Journal of Game Theory, **23(1)**, 43–48.

Parilina, E. M. (2014). *Strategic consistency of single-point optimality pronciples in cooperative stochastic games.* Mathematical Game Theory and Its Applications, **6(1)**, 56–72 (in Russian).

Parilina, E. M. (2015). *Stable cooperation in stochastic games.* Automation and Remote Control, **76**, 1111–1122.

Parilina, E. (2016). *Strategic Support of the Shapley Value in Stochastic Games.* Contributions to Game Theory and Management, **IX**, 246–265.

Parilina, E. M. (2009). *Cooperative stochastic game of data transmission in wireless network.* Mathematical Game Theory and Its Applications, **1(4)**, 93–110 (in Russian).

Parilina, E. M. (2008). *Subgame Consistency of Shapley Value in Cooperative Data Transmission Game in Wireless Network.* Contributions to Game Theory and Management, SPb, **1**, 381–994.

Parilina, E. and Sedakov, A. (2015). *Stochastic Approach for Determining Stable Coalition Structure.* International Game Theory Review, **17(4)**, pp. 155009-1—155009-22.

Parilina, E. M. and Petrosyan, L. A. (2017). *Strongly subgame consistent core in stochastic games.* Mathematical Game Theory and Its Applications, **9(2)**, 39–61 (in Russian).

Parilina, E. and Tampieri, A. (2018). *Stability and Cooperative Solution in Stochastic Games.* Theory and Decision, **84(4)**, 601–625.

Parilina, E. and Zaccour, G. (2015). *Node-Consistent Core for Games Played over Event Trees.* Automatica, **53**, 304–311.

Pecherski, S. L. and Yanovskaya, E. B. (2004). *Cooperative games: solutions and axioms.* Saint Petersburg: Publ. house of European university in SPb (in Russian).

Petrosjan, L. A. (2006). *Cooperative Stochastic Games.* In: Advances in Dynamic Games. Annals of the ISDG. Application to Economics, Engineering and Environmental Management, ed. by A. Haurie, S. Muto, L. A. Petrosjan, T.E.S. Raghavan, pp. 139–146.

Petrosjan, L. A. and Baranova, E. M. (2005). *Cooperative Stochastic Games in Stationary Strategies.* Proceedings of the Fifth International ISDG Workshop International Society of Dynamic Games, Segovia (Spain), pp. 225–234.

Petrosjan, L. and Zaccour, G. (2003). *Time-consistent Shapley value allocation of pollution cost reduction.* Journal of Economic Dynamics & Control, **27**, 381–398.

Petrosjan, L. A. and Zenkevich, N. A. (2015). *Conditions for Sustainable Cooperation.* Automation and Remote Control, **76(10)**, 1894–1904.

Petrosyan, L. A. (1977). *Time consistency of solutions in cooperative games with many participants.* Vestnik Leningradskogo universiteta. Seria 1, **19**, 46–52 (in Russian).

Petrosyan, L. A. (1992). *Construction of strongly time-consistent solutions in cooperative differential games.* Vestnik Leningradskogo universiteta. Seria 1, **2**, 33–38 (in Russian).

Petrosyan, L. A. (1993). *Strongly time-consistent optimality principles in multi-criteria problems of optimal control.* Tekhnicheskaya kibernetika, **1**, 169–174 (in Russian).

Petrosyan, L. A. (1998). *Semi-cooperative games.* Vestnik Leningradskogo universiteta. Seria 1, **2**, 57–63 (in Russian).

Petrosyan, L. A. and Baranova, E. M. (2003). *Stochastic games with random duration.* Trudi XXXIVth nauchnoi konferencii aspirantov i studentov "Procesi upravleniya i ustoichivost". SPb, 456–462 (in Russian).

Petrosyan, L. A. and Baranova, E. M. (2005a). *Cooperative stochastic games.* Tezisi dokladov Mezhdunarodnogo seminara "Teoriya upravleniya i teoriya obobjennih resheniy uravneniy Hamiltona-Yacobi". Ekaterinburg, pp. 33–35 (in Russian).

Petrosyan, L. A. and Baranova, E. M. (2005b). *Cooperative stochastic games in stationary strategies.* Sbornik trudov Mezhdunarodnoi conferencii "Ustoichivoct i procesi upravleniya". SPb, **1**, 495–503 (in Russian).

Petrosyan, L. A., Baranova, E. M. and Shevkoplyas, E. V. (2004). *Multistage cooperative games with random duration.* Trudi Instituta Matematiki i Mekhaniki, **10(2)**, 116–130 (in Russian).

Petrosyan, L. and Chistyakov, S. (2013). *Strategic support of Cooperative Solutions in 2-Person Differential Games with Dependent Motions.* Contributions to Game Theory and Management, **6**, 388–394.

Petrosyan, L. A. and Danilov, N. A. (1979). *Time consistent solutions of non-zero sum differential games with transferable payoffs.* Vestnik Leningradskogo universiteta. Seria 1, **1**, 46–54 (in Russian).

Petrosjan, L. A. and Grauer, L. V. (2002). *Strong Nash equilibrium in multistage games.* International Game Theory Review, **4(2)**, 255–264.

Gromova, E. V. and Petrosyan, L. A. (2015). *Strongly time-consistent cooperative solution for a differential game of pollution control.* Large-scale Systems Control, **55**, 140–159 (in Russian).

Petrosyan, L. A. and Kuzutin, D. V. (2000). *Games in extensive form: otimality and consistency.* Saint Petersburg: Izd-vo S.-Peterburgskogo universiteta (in Russian).

Petrosyan, L. A. and Shevkoplyas, E. V. (2000). *Cooperative differential games with random duration.* Vestnik Sankt-peterburgskogo universiteta. Seria 1, **4**, 14–18 (in Russian).

Petrosyan, L. and Sedakov, A. (2015). *Strategic support of cooperation in dynamic games on networks.* In: "Stability and Control Processes" in Memory of V.I. Zubov (SCP), 2015 International Conference, 256–260.

Reddy, P. V. and Zaccour, G. (2016). *A friendly computable characteristic function.* Mathematical Social Sciences, **82**, 18–25.

Schmeidler, D. (1969). *The Nucleolus of a Characteristic Function game.* SIAM Journal of Applied Mathematics, **17**, 1163–1170.

Schneider, R. (1993). *Convex bodies: the Brunn–Minkowski theory.* Cambridge Univ. Press.

Shapley, L. S. (1953a). *Stochastic Games.* Proceedings of National Academy of Sciences of the USA, **39**, 1095–1100.

Shapley, L. S. (1953b). *A Value for n-person Games.* In: H.W. Kuhn, A.W. Tucker, eds. Contributions to the Theory of Games II, Princeton: Princeton Univ. Press, 307–317.

Sedakov, A. A. (2015). *The strong time-consistent core.* Mathematical Game Theory and Its Applications, **7(2)**, 69–84.

Selten, R. (1975). *Reexamination of the perfectness concept for equilibrium points in extensive games.* International Journal of Game Theory, **4**, 25–55.

Shevkoplyas, E. V. (2010). *Stable cooperation in differential games with random duration.* Mathematical Game Theory and Its Applications, **2(3)**, 162–190 (in Russian).

Solan, E. (1998). *Discounted Stochastic Games.* Mathematics of Operations Research, **23**, 1010–1021.

Solan, E. (2009). *Stochastic Games.* In: Encyclopedia of Database Systems, Springer.

Solan, E. and Vieille, N. (2002). *Correlated Equilibrium in Stochastic Games.* Games and Economic Behavior, **38**, 362–399.

Solan, E. and Vieille, N. (2015). *Stochastic games.* Proceedings of the National Academy of Sciences of the United States of America, **112(45)**, 13743–13746.

Takahashi, M. (1964). *Stochastic games with infinitely many strategies.* Journal of Science of the Hiroshima University, **A-I 28**, 95–99.

http://mmiras.webs.uvigo.es/TUGlab/

Tur, A. V. (2014). *The irrational behavior proof condition for linear-quadratic discrete-time dynamic games with nontransferable payoffs.* Contributions to Game Theory and Management, **7**, 384–392.

Vieille, N. (2000). *Solvable states in n-player stochastic games.* SIAM Journal on Control and Optimization, **38(6)**, 1794–1804.

Vilkas, E. (1990). *Optimality in games and solutions.* Moscow: Nauka (in Russian).

Vorobiev, N. N. (1960). *Stable profiles in coalitional games.* Dokladi Akademii USSR, **131**, 493–495 (in Russian).

Vorobiev, N. N. (1967). *Coalitional games.* Teoriya Veroyatnostei i Ee Primenenie, **12(2)**, 289–306 (in Russian).

Vorobiev, N. N. (1985). *Game theory for ecomonists–cybernatists.* Moscow: Nauka (in Russian).

Yeung, D. W. K. (2006). *An irrational-behavior-proof condition in cooperative differential games.* International Game Theory Review, **8**, 739–744.

Yeung, D. W. K. and Petrosyan, L. A. (2011). *Subgame Consistent Cooperative Solution of Dynamic Games with Random Horizon.* Journal of Optimization Theory and Applications, **150(1)**, 78–97.