Article

Subgame Consistent Cooperative Behavior in an Extensive form Game with Chance Moves

Denis Kuzyutin [1,2] and **Nadezhda Smirnova** [1,2,*]

[1] Faculty of Applied Mathematics and Control Processes, Saint Petersburg State University, Universitetskaya nab. 7/9, 199034 St. Petersburg, Russia; d.kuzyutin@spbu.ru
[2] St. Petersburg School of Mathematics, Physics and Computer Science, National Research University Higher School of Economics (HSE), Soyuza Pechatnikov ul. 16, 190008 St. Petersburg, Russia
[*] Correspondence: nadezhda.v.smirnova@gmail.com

check for updates

Abstract: We design a mechanism of the players' sustainable cooperation in multistage n-person game in the extensive form with chance moves. When the players agreed to cooperate in a dynamic game they have to ensure time consistency of the long-term cooperative agreement. We provide the players' rank based (PRB) algorithm for choosing a unique cooperative strategy profile and prove that corresponding optimal bundle of cooperative strategies satisfies time consistency, that is, at every subgame along the optimal game evolution a part of each original cooperative trajectory belongs to the subgame optimal bundle. We propose a refinement of the backwards induction procedure based on the players' attitude vectors to find a unique subgame perfect equilibrium and use this algorithm to calculate a characteristic function. Finally, to ensure the sustainability of the cooperative agreement in a multistage game we employ the imputation distribution procedure (IDP) based approach, that is, we design an appropriate payment schedule to redistribute each player's optimal payoff along the optimal bundle of cooperative trajectories. We extend the subgame consistency notion to extensive-form games with chance moves and prove that incremental IDP satisfies subgame consistency, subgame efficiency and balance condition. An example of a 3-person multistage game is provided to illustrate the proposed cooperation mechanism.

Keywords: time consistency; multistage game; chance moves; subgame perfect equilibria; cooperative trajectory; imputation distribution procedure

1. Introduction

In a dynamic n-person game the players first choose their "optimal" strategies at the initial position x_0 (which form the optimal strategy profile for the whole game), and then have an option to change their strategies at any intermediate position x_t and switch to other strategies if these strategies constitute the locally optimal strategy profile for the subgame starting at x_t. The time consistency property (first introduced in References [1–3] for differential games) ensures that the players will not have an incentive to change their strategies at any subgame along the optimal game evolution, and hence plays an important role in the designing of the optimal players' behavior in non-cooperative and cooperative dynamic games (see, e.g., References [2–21], for details).

We consider an *n*-person finite multistage games in the extensive form (see, e.g., References [5,17,22,23]) with perfect information and with chance moves. Note that much research has been already done on time consistent solutions (or close concepts) in extensive-form games (see, e.g., References [4,6,13,17,21]). Time consistency concept was extended to dynamic games played over event trees in References [14,16,20] as well as to multicriteria extensive-form cooperative games (without chance moves) in References [7,8,10,11,15]. The property of "time consistency in the whole game" was extended to multicriteria extensive-form cooperative games with chance moves in Reference [9] (note

that in these games an optimal pure strategy profile does not generate the unique optimal trajectory in the game tree but rather the whole optimal bundle of the trajectories).

In the paper, we mainly focus on the dynamic aspects of cooperation in a dynamic extensive-form game with chance moves, and propose to design a mechanism of the players' sustainable cooperation which satisfies three properties. First, a fragment of each cooperative trajectory from the optimal bundle for the original game Γ^{x_0} should "remain optimal" at each subgame Γ^{x_t} along the cooperative game evolution, that is, it should belong to the subgame optimal bundle of cooperative trajectories. Secondarily, a cooperative payoff-to-go at the subgame Γ^{x_t} is no less than the non-cooperative payoff-to-go for all players. Finally, when the players re-evaluate their expected cooperative payoff after each passed chance move, they have no incentive to change original cooperative agreement.

To this aim, we first need to provide a rule for choosing a unique cooperative strategy profile as well as the unique optimal bundle of cooperative trajectories. We introduce the *Players' Rank Based (PRB) algorithm* and prove that this algorithm generates the unique optimal bundle of cooperative trajectories which satisfies time consistency. Note that a rather close approach—the so-called Refined Leximin (RL) algorithm—was introduced recently in Reference [8]. Let us notice the main differences of these two algorithms. The RL algorithm is applicable for multicriteria game without chance moves and is based on the ranking of the criteria, while the PRB algorithm is designed for single-criterion extensive-form game with chance moves and employed the players' ranks. Further, the RL algorithm allows to choose a unique cooperative trajectory while the PRB algorithm generates the unique optimal bundle of the cooperative trajectories in the game tree. To the best of the authors' knowledge, other approaches to choosing an optimal bundle of the cooperative trajectories in extensive-form game with chance moves have not been considered yet.

Then, to construct a characteristic function (which describes the worth of each coalition in cooperative game) we use an equilibrium-based approach, namely the γ-characteristic function introduced in Reference [24]. Hence, the players have to accept a specific method for choosing a unique Subgame Perfect Equilibria (SPE) [25] in an extensive-form game with chance moves. To solve this problem we provide the novel refinement of the backwards induction procedure (see, e.g., References [5,17,23])—the so-called *Attitude SPE algorithm*. A similar approach to construct a unique SPE in extensive-form game with perfect information was explored in References [17,26,27] and was called the Type Equilibrium (TE) algorithm. Both algorithms are the refinements of the general backwards induction procedure that take into account the attitudes of each player towards other players. Let us point out the main differences of these algorithms. The TE algorithm is applicable for the game without chance moves and for the case when the payoffs are only determined in terminal nodes. In addition, the TE algorithm allows to construct SPE that is "unique" in the sense of payoffs (i.e., there may exist several optimal trajectories which generate the same equilibrium payoffs) while the Attitude SPE algorithm allows to choose unique SPE strategy profile as well as unique bundle of trajectories. Another rather close approach to find a unique SPE—the so-called Indifferent Equilibrium (IE) algorithm—was introduced in Reference [28]. Again, the IE algorithm is applicable only for the game without chance moves and for the partial case when the payoffs are determined in terminal nodes. Moreover, IE algorithm in general allows to construct a SPE in behavior strategies while the proposed Attitude SPE algorithm always generates a SPE in pure strategies.

It is worth noting, that other approaches to analyze an extensive-form game, except for the backwards induction procedure and its refinements mentioned above, imply that the researcher first needs to obtain a strategic representation of the original extensive game and then analyzes this strategic (or normal-form) game (see, e.g., References [29–31]). For instance, the software tool "Game Theory Explorer" [29] is based on the strategic-form representation and then applying the modified Lemke-Howson algorithm [32] to find all Nash equilibria. The majority of existing algorithms are developed to find Nash equilibria in mixed strategies for 2-person games and do not allow to construct SPE in pure strategies. Moreover, as it was noted in Reference [31], in general the strategic-form representation is exponential in the size of the original game tree. In contrast, the proposed Attitude

SPE algorithm is a rather simple recursive algorithm which deals with n-person extensive-form game (with perfect information) itself and allows to compute a unique SPE in pure strategies.

After computing the γ-characteristic function we suppose that the players adopt some single-valued cooperative solution φ (for instance, the Shapley value [33], the nucleolus [34], etc.) which satisfies the individual and collective rationality property. Finally, to guarantee the sustainability of the achieved long-term cooperative agreement we employ the *Imputation Distribution Procedure (IDP)* based approach (see, e.g., References [3,12,14,16–18,20,35]), that is, a payment schedule to redistribute the ith player's expected cooperative payoff along the optimal bundle of cooperative trajectories. In this paper, we mainly focus on the following good properties an IDP may satisfy: subgame efficiency, strict balance condition [10,15,17] and an appropriate refinement of the time consistency property, called *subgame consistency*. The point is that the "time consistency in the whole game" property [9,14,16,20] is based on an a priori assessment of the ith player's expected optimal payoff (before the game Γ^{x_0} starts). However, when the players make a decision in the subgame Γ^{x_t} after the chance move occurs, they need to re-estimate their expected optimal payoffs-to-go since the original optimal bundle of cooperative trajectories shrinks after each chance node. To deal with this interesting feature of the game with chance moves we adopt the notion of subgame consistency that was firstly proposed in Reference [36] for cooperative stochastic differential games and then extend it to stochastic dynamic games in References [37,38].

Since we derive a suitable definition of subgame consistency for other class of games, the proposed Definition 6 differs from ones provided in References [37,38] but captures the same idea. Let us point out the main differences with References [37,38]. While D. Yeung and L. Petrosyan do not consider the issue of multiple equilibria and study the stochastic games in which there exists a unique Nash equilibrium in each subgame, we focus on the problem of how to select a unique (subgame perfect) Nash equilibrium in extensive-form game with chance moves and derive the corresponding algorithm. Secondarily, the characteristic function has not been constructed in References [37,38] and, hence, the players are restricted to using the simplest cooperative solutions (for instance, they may share equally the excess of the total expected cooperative payoff over the expected sum of individual non-cooperative payoffs), whereas we provide a method for calculating the γ-characteristic function. Hence, the players may use different solution concepts based on the characteristic function approach. Finally, it turns out that the incremental IDP specified for extensive-form games with chance moves in Reference [9] satisfies not only the subgame consistency but also subgame efficiency and strict balance condition.

Therefore, the suggested PRB algorithm, the Attitude SPE algorithm combined with the γ-characteristic function, and the incremental payment schedule for any single-valued cooperative solution (meeting individual and collective rationality) together constitute a required mechanism of the players' sustainable cooperation that satisfies exactly three properties mentioned above for any extensive-form game with chance moves.

It is worth noting that the extensive-form games, as well as dynamic games played over event trees, differential games and multistage games with discrete dynamics are used to model various real-world situations where several decision makers (or players) with different objectives may cooperate (see, e.g., References [5,12,14,16,17,20,39–44]). Hence, a proposed approach to implement a long-term cooperative agreement may have a number of possible applications.

The rest of the paper is organized as follows: Section 2 recalls the main ingredients of the class of games of interest. In Section 3, we specify the attitude SPE algorithm that allows constructing a unique SPE in a extensive-form game with chance moves. In Section 4, we provide the PRB algorithm and prove that the optimal bundle of cooperative trajectories generated by this algorithm satisfy time consistency. Section 5 reveals a drawback of the IDP "time consistency in the whole game" property and presents a subgame consistency definition that is applicable for extensive-form games with chance moves. We prove that incremental IDP satisfies a number of good properties and consider an example of a 3-person multistage game with chance moves to illustrate the incremental IDP implementation. Section 6 provides a brief review of the results and discussion.

2. Extensive-Form Game with Chance Moves

We consider a finite multistage game in extensive form following References [6,13,17,22,23]. First we need to define the basic notations and briefly remind some properties of extensive-form game that will be used in the sequel:

- $N = \{1, \ldots, n\}$ is the set of all players.
- K is the game tree with the root x_0 and the set of all nodes P.
- $S(x)$ is the set of all direct successors (descendants) of the node x, and $S^{-1}(y)$ is the unique predecessor (parent) of the node $y \neq x_0$ such that $y \in S(S^{-1}(y))$.
- P_i is the set of all decision nodes of the ith player (at these nodes the player i chooses the following node), $P_i \cap P_j = \varnothing$, for all $i, j \in N, i \neq j$.
- $P_{n+1} = \{z^j\}_{j=1}^m$ denotes the set of all terminal nodes (final positions), $S(z^j) = \varnothing \; \forall z^j \in P_{n+1}$.
- P_0 is the set of all nodes at which a chance moves, where $\pi(y|x) > 0$ denotes the probability of transition from node $x \in P_0$ to node $y \in S(x)$. We suppose that for each $x \in P_0$ it holds that $S(x) \cap P_0 = \varnothing$. Lastly, $\bigcup\limits_{i=0}^{n+1} P_i = P$.
- $\omega = (x_0, \ldots, x_{t-1}, x_t, \ldots, x_T)$ is the trajectory (or the path) in the game tree, $x_{t-1} = S^{-1}(x_t)$, $1 \leq t \leq T$, $x_T = z^j \in P_{n+1}$; where index t in x_t denotes the ordinal number of this node within the trajectory ω and can be interpreted as the "time index".
- $h_i(x) = (h_{i/1}(x), \ldots, h_{i/r}(x))$ is the payoff of the ith player at the node $x \in P$. We assume that for all $i \in N$, $k = 1, \ldots, r$, and $x \in P$ the payoffs are non-negative, that is, $h_{i/k}(x) \geqslant 0$.

In the following, we will use $G^{cm}(n)$ to denote the class of all finite multistage n-person games with chance moves in extensive form defined above, where $\Gamma^{x_0} \in G^{cm}(n)$ denotes a game with root x_0. Note that Γ^{x_0} is an extensive-form game with perfect information (see, e.g., References [17,22,23] for details).

Since all the solutions we are interested in throughout the paper are attainable when the players restrict themselves to the class of pure strategies we will focus on this class of strategies. The pure strategy $u_i(\cdot)$ of the ith player is a function with domain P_i that specifies for each node $x \in P_i$ the next node $u_i(x) \in S(x)$ which the player i has to choose at x. Let U_i denote the (finite) set of all ith player's pure strategies, $U = \prod_{i \in N} U_i$.

Denote by $p(y|x, u)$ the conditional probability that node $y \in S(x)$ is reached if node x has been already reached (the probability of transition from x to y) while the players use the strategies u_i, $i \in N$. Note that for all $x \in P_i$, $i = 1, \ldots, n$, and for all $y \in S(x)$ $p(y|x, u) = 1$ if $u_i(x) = y$ and $p(y|x, u) = 0$ if $u_i(x) \neq y$. For chance moves, that is, if $x \in P_0$ $p(y|x, u) = \pi(y|x)$ for all $y \in S(x)$ for each $u \in U$.

Then one can calculate the probability $p(\omega, u)$ of realization of the trajectory $\omega = (x_0, \ldots, x_\tau, x_{\tau+1}, \ldots, x_T)$, $x_T \in P_{n+1}$, $x_{\tau+1} \in S(x_\tau)$, $\tau = 0, \ldots, T-1$, when the players use the strategies u_i from the strategy profile $u = (u_1, \ldots, u_n)$.

$$p(\omega, u) = p(x_1|x_0, u) \cdot p(x_2|x_1, u) \cdot \ldots \cdot p(x_T|x_{T-1}, u) = \prod_{\tau=0}^{T-1} p(x_{\tau+1}|x_\tau, u). \tag{1}$$

Denote by $\Omega(u) = \{\omega_k(u) | p(\omega_k, u) > 0\}$ the finite set (or the bundle) of the trajectories ω_k which are generated by strategy profile $u \in U$. Note that for all $\omega_k(u) \in \Omega(u)$, $u_j(x_\tau) = x_{\tau+1}$ for all $x_\tau \in \omega_k(u) \cap P_j$, $j \in N$, $0 \leq \tau \leq T-1$.

Let $\tilde{h}_i(\omega) = \sum\limits_{\tau=0}^{T} h_i(x_\tau)$ denote the ith player's vector payoff corresponding to the trajectory $\omega = (x_0, \ldots, x_t, x_{t+1}, \ldots, x_T)$.

Denote by

$$H_i(u) = \sum_{\omega_k \in \Omega(u)} p(\omega_k, u) \cdot \tilde{h}_i(\omega_k) = \sum_{\omega_k \in \Omega(u)} p(\omega_k, u) \cdot \sum_{\tau=0}^{T(k)} h_i(x_\tau) \tag{2}$$

the (expected) value of the ith player's payoff function which corresponds to the strategy profile $u = (u_1, \ldots, u_n)$. Let $\Omega_{n+1}(u) = \{\Omega(u) \cap P_{n+1}\}$ denote the set of all terminal nodes of the trajectories $\omega_k(u) \in \Omega(u)$.

Remark 1 ([9]). *If the pure strategy profiles u and v generate different bundles $\Omega(u)$ and $\Omega(v)$ of the trajectories, that is, $\Omega(u) \neq \Omega(v)$, then $\Omega_{n+1}(u) \cap \Omega_{n+1}(v) = \varnothing$.*

According to References [17,22,23] each intermediate node $x_t \in P \setminus P_{n+1}$ generates a subgame Γ^{x_t} with the subgame tree K^{x_t} and the subgame root x_t as well as a factor-game Γ^D with the factor-game tree $K^D = (K \setminus K^{x_t}) \cup \{x_t\}$. Decomposition of the original extensive game Γ^{x_t} at node x_t onto the subgame Γ^{x_t} and the factor-game Γ^D generates the corresponding decomposition of the pure (and mixed) strategies (see References [17,22] for details).

Let $P_i^{x_t}(P_i^D)$, $i \in N$, denote the restriction of P_i on the subgame tree $K^{x_t}(K^D)$, and $u_i^{x_t}(u_i^D)$, $i \in N$, denote the restriction of the ith player's pure strategy $u_i(\cdot)$ in Γ^{x_0} on $P_i^{x_t}(P_i^D)$. The pure strategy profile $u^{x_t} = (u_1^{x_t}, \ldots, u_n^{x_t})$ generates the bundle of the subgame trajectories $\Omega^{x_t}(u^{x_t}) = \{\omega_k^{x_t}(u^{x_t}) | p(\omega_k^{x_t}, u^{x_t}) > 0\}$. Similarly to (2), let us denote by

$$H_i^{x_t}(u^{x_t}) = \sum_{\omega_k^{x_t} \in \Omega^{x_t}(u^{x_t})} p(\omega_k^{x_t}, u^{x_t}) \cdot \sum_{\tau=t}^{T(k)} h_i(x_\tau) = \sum_{\omega_k^{x_t} \in \Omega^{x_t}(u^{x_t})} p(\omega_k^{x_t}, u^{x_t}) \cdot \tilde{h}_i(\omega_k^{x_t}) \tag{3}$$

the expected value of the ith player's payoff in Γ^{x_t}, and by $U_i^{x_t}$ the set of all possible ith player's pure strategies in the subgame Γ^{x_t}, $U^{x_t} = \prod_{i \in N} U_i^{x_t}$. Note that for each trajectory $\omega = (x_0, \ldots, x_t, x_{t+1}, \ldots, x_T)$, $1 \leqslant t \leqslant T - 1$, $x_T \in P_{n+1}$,

$$p(\omega, u) = \prod_{\tau=0}^{t-1} p(x_{\tau+1}|x_\tau, u) \cdot \prod_{\tau=t}^{T-1} p(x_{\tau+1}|x_\tau, u) =$$
$$= p(\underline{\omega}^{x_t}, u) \cdot p(\omega^{x_t}, u) = p(\underline{\omega}^{x_t}, u^D) \cdot p(\omega^{x_t}, u^{x_t}), \tag{4}$$

where $\underline{\omega}^{x_t} = (x_0, x_1, \ldots, x_{t-1}, x_t)$ denotes a fragment of trajectory ω implemented before the subgame Γ^{x_t} starts, and $p(\underline{\omega}^{x_t}, u) = p(x_t, u)$ denotes the probability that node x_t is reached when the players employ the strategies u_i, $i \in N$. It is worth noting that factor-game $\Gamma^D = \Gamma^D(u^{x_t})$ is usually defined for given strategy profile u^{x_t} in the subgame Γ^{x_t} since we assume that

$$h_i^D(x_0, x_1, \ldots, x_{t-1}, x_t) = \sum_{\tau=0}^{t-1} h_i(x_\tau) + H_i^{x_t}(u^{x_t}) = \tilde{h}_i(\underline{\omega}^{x_t} \setminus \{x_t\}) + H_i^{x_t}(u^{x_t}) \tag{5}$$

(see, e.g., References [17,22] for details). Moreover, given intermediate node x_t, the bundle $\Omega(u) = \{\omega_k(u) | p(\omega_k, u) > 0\}$ can be divided in two subsets, that is, $\Omega(u) = \{\Psi_m\} \cup \{\chi_l\}$, where $x_t \in \Psi_m$, and $x_t \notin \chi_l$, $\{\Psi_m\} \cap \{\chi_l\} = \varnothing$. Then, taking (1), (3), (4) and (5) into account, we get

$$H_i(u) = \sum_m p(\Psi_m, u) \cdot \tilde{h}_i(\Psi_m) + \sum_l p(\chi_l, u) \cdot \tilde{h}_i(\chi_l) =$$

$$= \sum_m p(x_t, u) \cdot p(\Psi_m^{x_t}, u^{x_t}) \cdot [\tilde{h}_i(\underline{\Psi}_m^{x_t} \setminus \{x_t\}) + \tilde{h}_i(\Psi_m^{x_t})] +$$

$$+ \sum_l p(\chi_l, u) \cdot \tilde{h}_i(\chi_l) = p(x_t, u^D) \cdot \tilde{h}_i(x_0, \dots, x_{t-1}) \cdot \sum_m p(\Psi_m^{x_t}, u^{x_t}) +$$

$$+ p(x_t, u^D) \cdot \sum_m p(\Psi_m^{x_t}, u^{x_t}) \cdot \tilde{h}_i(\Psi_m^{x_t}) + \sum_l p(\chi_l, u) \cdot \tilde{h}_i(\chi_l) =$$

$$= p(x_t, u^D) \cdot \tilde{h}_i(x_0, \dots, x_{t-1}) + p(x_t, u^D) \cdot H_i^{x_t}(u^{x_t}) +$$

$$+ \sum_l p(\chi_l, u) \cdot \tilde{h}_i(\chi_l) = p(x_t, u^D) \cdot h_i^D(x_0, \dots, x_t) + \sum_l p(\chi_l, u) \cdot \tilde{h}_i(\chi_l). \tag{6}$$

Note that, since $P_i = P_i^{x_t} \cup P_i^D$, one can compose the ith player's pure strategy $W_i = (u_i^D, v_i^{x_t}) \in U_i$ in the original game Γ^{x_0} from her strategies $v_i^{x_t} \in U_i^{x_t}$ in the subgame Γ^{x_t} and $u_i^D \in U_i^D$ in the factor-game Γ^D [17,22].

3. Refined Backwards Induction Procedure to Construct a Unique SPE

Definition 1 ([45]). *A strategy profile $u = (u_1, u_2, \dots, u_n)$ is a Nash Equilibrium (NE) in $\Gamma^{x_0} \in G^{cm}(n)$, if*

$$H_i(v_i, u_{-i}) \leqslant H_i(u_i, u_{-i}), \ \forall v_i \in U_i, \ \forall i \in N.$$

Let $NE(\Gamma^{x_0})$ denote the set of all pure strategy Nash equilibria in Γ^{x_0}.

Definition 2 ([25]). *A strategy profile u is a subgame perfect (Nash) equilibrium (SPE) in $\Gamma^{x_0} \in G^{cm}(n)$, if $\forall x \in P \setminus P_{n+1}$ it holds that $u^x \in NE(\Gamma^x)$, i. e. the restriction of u on each subgame Γ^x forms a NE in this subgame.*

To construct SPE in an extensive-form game with perfect information one may employ a so-called backwards induction procedure (see, e.g., References [12,17,22,23,46,47]).

However, the backwards induction procedure may generate multiple subgame perfect equilibriums in an extensive form game with different payoffs to the players (see, e.g., References [5,12,17,23]). To choose a unique SPE and unique corresponding bundle of trajectories we use an approach based on the players' attitude vectors. Namely, let the ith player's attitude vector $F_i = \{f_i(1), \dots, f_i(n)\}$ be a permutation of numbers $\{1, \dots, n\}$ meeting the condition $f_i(i) = 1$. If $f_i(j) = k$ one may interpret the player j as an "ith player's associate of level k".

In the paper we will use these attitude vectors when constructing SPE via backwards induction procedure in the following way. Let $x \in P_i$, $H_i^y(\underline{u}^y)$ denote the ith player's expected payoff in the subgame Γ^y, $y \in S(x)$ while $\underline{u}^y$ be a SPE in this subgame. Assume that there exist multiple nodes $y_1, \dots, y_q$ such that $h_i(y_1) + H_i^{y_1}(\underline{u}^{y_1}) = \dots = h_i(y_q) + H_i^{y_q}(\underline{u}^{y_q})$, that is, player i is indifferent to the choice of particular node $\bar{y}$ from $\{y_1, \dots, y_q\}$ while the ith player's choice may affect the other players' payoffs. If $f_i(j) = 2$, suppose that the ith player aims to maximize firstly the jth player's expected payoff $H_j^y(\underline{u}^y)$ when choosing a unique node y from $y_1, \dots, y_q$. If again there are several nodes y with the same value $H_j^y(\underline{u}^y)$ the ith player purposes to maximize secondarily the expected payoff $H_l^y(\underline{u}^y)$ of such player l that $f_i(l) = 3$, and so on. Note that similar approach to construct a unique SPE in extensive-form game with perfect information but without chance moves was explored in References [17,26,27] for the case when the payoffs are only determined in terminal nodes.

Now let us provide a rigorous specification of this backwards induction procedure refinement which we will refer to as the *Attitude SPE or A-SPE algorithm*.

Attitude SPE algorithm. Suppose that the players attitude vectors $F_1, F_2, \dots, F_n$ are of common knowledge, i. e. each player knows these vectors, and all the players are aware of it. Let the length

of the trajectory $\omega = (x_0, \ldots, x_t, x_{t+1}, \ldots, x_T)$ equals to $T - 1$, and the multistage game Γ^{x_0} length equals to the maximal length of the trajectory ω in Γ^{x_0}. We'll construct the unique subgame perfect equilibrium $\underline{u} = (\underline{u}_1, \ldots, \underline{u}_n)$ in Γ^{x_0} by induction in the length L of the subgame Γ^x.

Step $L = 1$: Consider a subgame Γ^x of the length $L = 1$. If $x \in P_i$, $i = 1, \ldots, n$, we have two cases.

Case 1: there exists a unique $z^k \in S(x) = P^x_{n+1}$ such that $h_i(z^k) = \max\limits_{z \in S(x)} h_i(z)$. Then suppose that
$$\underline{u}_i(x) = z^k, \; p(z^k | x, \underline{u}) = 1, \; p(z | x, \underline{u}) = 0 \; \forall z \in S(x) \setminus \{z^k\}.$$
Case 2: there exist $q > 1$ nodes $z^{k_q} \in S(x) = P^x_{n+1}$ such that $h_i(z^{k_1}) = h_i(z^{k_2}) = \ldots = h_i(z^{k_q}) = \max\limits_{z \in S(x)} h_i(z)$. Then suppose that the ith player chooses the terminal position $z^k \in \{z^{k_1}, \ldots, z^{k_q}\} = S^{i,1}(x)$ such that

$$h_j(z^k) = \max\limits_{z \in S^{i,1}(x)} h_j(z), \text{ where } f_i(j) = 2. \tag{7}$$

Let $S^{i,2}(x)$ denote the set of all nodes $z^k \in S^{i,1}(x)$ meeting (7). If $S^{i,2}(x)$ consists of unique node z^k then $\underline{u}_i(x) = z^k$, $p(z^k | x, \underline{u}) = 1$, $p(z | x, \underline{u}) = 0 \; \forall z \in S(x) \setminus \{z^k\}$. Otherwise, suppose that the ith player chooses terminal node $z^k \in S^{i,2}(x)$ such that

$$h_l(z^k) = \max\limits_{z \in S^{i,2}(x)} h_l(z), \text{ where } f_i(l) = 3. \tag{8}$$

Let $S^{i,3}(x)$ denote the set of all final nodes $z^k \in S^{i,2}(x)$ satisfying (8), and so on.
$\vdots$

Finally, if $S^{i,n}(x)$ contains unique node z^k, then $\underline{u}_i(x) = z^k$, $p(z^k | x, \underline{u}) = 1$, $p(z | x, \underline{u}) = 0$ $\forall z \in S(x) \setminus \{z^k\}$. Otherwise, suppose that player i chooses the final node z^k from $S^{i,n}(x)$ with minimal ordinal number k.

Note that for all cases $H_j(\underline{u}) = h_j(z^k)$, $j \in N$.

If $x \in P_0$ then $S(x) = P^x_{n+1}$ and we do not need to define a strategy of any player at x, while $H_j(\underline{u}) = \sum\limits_{z^k \in S(x)} \pi(z^k | x) \cdot h_j(z^k)$. Hence, the players' behavior $\underline{u}^x = (\underline{u}^x_1, \ldots, \underline{u}^x_n) \in NE(\Gamma^x)$ and the expected payoffs $H^x_j(\underline{u}^x)$, $j \in N$ are defined for all subgames Γ^x of the length 1. In addition, for games Γ^y, $y \in P_{n+1}$ of length $L = 0$ we assume that $H^y_i(\underline{u}^y) = h_i(y)$, $i \in N$.

Step $2, \ldots, L - 1$: Suppose that at each subgame Γ^y of the length $(L - 1)$ or less the unique SPE $\underline{u}^y = (\underline{u}^y_1, \ldots, \underline{u}^y_n)$ has been already constructed ("inductive assumption"), and $H^y_i(\underline{u}^y)$, $i \in N$, is the corresponding vector of all the players' payoffs.

Step L: Consider the game Γ^{x_0} of the length $L \geqslant 1$. Note that for all $y \in S(x_0)$ the length of the subgame Γ^y is less than L. If $x_0 \in P_0$ then

$$H_j(\underline{u}) = \sum\limits_{y \in S(x_0)} \pi(y|x_0) \cdot (h_j(y) + H^y_j(\underline{u}^y)) \geqslant \sum\limits_{y \in S(x_0)} \pi(y|x_0) \cdot (h_j(y) + H^y_j(u^y_j, \underline{u}^y_{-j})) = H_j(u^y_j, \underline{u}^y_{-j}) \tag{9}$$

for all $u_j = u^y_j \in U^y_j = U_j$ since $\underline{u}^y \in NE(\Gamma^y)$ due to induction assumption, and each player $j \in N$ can deviate from $\underline{u}_j$ only in the subgames Γ^y, $y \in S(x_0)$.

If $x_0 \in P_i$ for some $i \in N$, we have two cases.

Case 1: there exists a unique $\bar{y} \in S(x_0)$ such that

$$h_i(\bar{y}) + H^{\bar{y}}_i(\underline{u}^{\bar{y}}) = \max\limits_{y \in S(x_0)} \left(h_i(y) + H^y_i(\underline{u}^y) \right). \tag{10}$$

Then we suppose that $\underline{u}_i(x_0) = \bar{y}$; $\underline{u}_j(x) = \underline{u}^y_j(x)$ if $x \in P_j \cap K^y$, $y \in S(x_0)$, $j = 1, \ldots, n$.

Case 2: there exist $q > 1$ nodes $\overline{y}_1, \ldots, \overline{y}_q \in S(x_0)$ such that

$$h_i(\overline{y}_1) + H_i^{\overline{y}_1}(\underline{u}^{\overline{y}_1}) = \ldots = h_i(\overline{y}_q) + H_i^{\overline{y}_q}(\underline{u}^{\overline{y}_q}) = \max_{y \in S(x_0)} \left(h_i(y) + H_i^y(\underline{u}^y) \right). \tag{11}$$

Then we suppose that the ith player chooses $\overline{y} \in \{\overline{y}_1, \ldots, \overline{y}_q\} = S^{i,1}(x_0)$ such that

$$h_j(\overline{y}) + H_j^{\overline{y}}(\overline{u}^{\overline{y}}) = \max_{y \in S^{i,1}(x_0)} \left(h_j(y) + H_j^y(\underline{u}^y) \right), \text{ where } f_i(j) = 2. \tag{12}$$

Let $S^{i,2}(x_0)$ denote the set of all nodes $\overline{y} \in S^{i,1}(x_0)$ satisfying (12). If $S^{i,2}(x_0)$ consists of unique node $\overline{y}$ then we suppose that $\underline{u}_i(x_0) = \overline{y}$; $\underline{u}_j(x) = u_j^y(x)$ if $x \in P_j \cap K^y$, $y \in S(x_0)$, $j = 1, \ldots, n$. Otherwise, suppose that the ith player chooses node $\overline{y} \in S^{i,2}(x_0)$ such that

$$h_l(\overline{y}) + H_l^{\overline{y}}(\overline{u}^{\overline{y}}) = \max_{y \in S^{i,2}(x_0)} \left(h_l(y) + H_l^y(\underline{u}^y) \right), \text{ where } f_i(l) = 3. \tag{13}$$

Let $S^{i,3}(x_0)$ denote the set of all nodes $\overline{y} \in S^{i,2}(x_0)$ meeting (13), and so on.

$\vdots$

Finally, if $S^{i,n}(x_0)$ contains several nodes $\overline{y}_m$, denote by $\underline{l} = \min_{\overline{y}_m \in S^{i,n}(x_0)} \{l \mid z^l \in P_{n+1}^{\overline{y}_m} \cap \Omega(\underline{u}^{\overline{y}_m})\}$ the minimal number of terminal nodes of the trajectories generated by subgame perfect equilibriums $\underline{u}^{\overline{y}_m}$ in the subgames $\Gamma^{\overline{y}_m}$, $\overline{y}_m \in S^{i,n}(x_0)$ (see Remark 1). Note that there exists unique trajectory $\omega = (x_0, \ldots, z^{\underline{l}})$ from x_0 to $z^{\underline{l}}$ in the game Γ^{x_0}, and let $\overline{y} = \omega \cap S^{i,n}(x_0)$. Again, we suppose that $\underline{u}_i(x_0) = \overline{y}$; $\underline{u}_j(x) = u_j^y(x)$ if $x \in P_j \cap K^y$, $y \in S(x_0)$, $j = 1, \ldots, n$.

Now we prove that for both cases no player has profitable deviation in Γ^{x_0} from the strategy profile $\underline{u} = (\underline{u}_1, \ldots, \underline{u}_n)$ constructed above.

$$H_i(\underline{u}) = h_i(\overline{y}) + H_i^{\overline{y}}(\underline{u}^{\overline{y}}) \geqslant h_i(y) + H_i^y(\underline{u}^y) \geqslant h_i(y) + H_i^y(u_i^y, \underline{u}^y_{-i}) \tag{14}$$

for all $y \in S(x_0)$, $u_i^y \in U_i^y$ due to (10), (11) and the induction assumption that $\underline{u}^y \in NE(\Gamma^y)$, $y \in S(x_0)$.
For other players $j \in N$, $j \neq i$, we have

$$H_j(\underline{u}) = h_j(\overline{y}) + H_j^{\overline{y}}(\underline{u}^{\overline{y}}) \geqslant h_j(\overline{y}) + H_j^{\overline{y}}(u_j^{\overline{y}}, \underline{u}^{\overline{y}}_{-j}) = H_j(u_j, \underline{u}_{-j}) \tag{15}$$

for all $u_j \in U_j$ since $\underline{u}^{\overline{y}} \in NE(\Gamma^{\overline{y}})$, and the only deviation of player $j \in N$, $j \neq i$ from $\underline{u}_j$ in the subgame $\Gamma^{\overline{y}}$ may affect the players' payoffs.

Hence, taking (9), (14) and (15) into account we obtain by induction that the strategy profile $\underline{u} = (\underline{u}_1, \ldots, \underline{u}_n)$ constructed above forms unique subgame perfect equilibria in Γ^{x_0}.

Proposition 1. *If the players attitude vectors $F_1, F_2, \ldots, F_n$ are of common knowledge, the Attitude SPE algorithm allows to construct a unique subgame perfect equilibrium $\underline{u} = (\underline{u}_1, \ldots, \underline{u}_n)$ in pure strategies for any extensive-form game $\Gamma^{x_0} \in G^{cm}(n)$ with chance moves as well as a unique bundle of trajectories $\Omega(\underline{u})$.*

It is worth noting than the existence of (subgame perfect) pure strategy equilibrium in extensive form game with perfect information and chance moves was first proved in References [46,47] for the partial case when the payoffs are only defined in terminal nodes. Hence, Proposition 1 could be considered as a corollary of these results. However, we provide a rigorous algorithm how to construct a unique *SPE* in extensive-form game with chance moves as well as a (unique) corresponding bundle of trajectories. We will use this algorithm, in particular, to calculate the characteristic function of the cooperative extensive-form game in Section 4.

Let us use the following example to demonstrate how the Attitude SPE algorithm works.

Example 1. *(A 3-player multistage game with chance moves).*

Let $P_0 = \{\overline{x}_1, \overline{x}_3\}$, $P_1 = \{\overline{x}_0, \overline{x}_4^2\}$, $P_2 = \{\overline{x}_2^1, \overline{x}_5\}$, $P_3 = \{x_2^2, x_4^3\}$, $P_{n+1} = \{z_1, \ldots, z_{10}\}$. *The players'*
payoffs and probabilities $\pi(y|x)$, $x \in P_0$ are written in the game tree.

Suppose that the players' attitude vectors are $F_1 = (f_1(1), f_1(2), f_1(3)) = (1, 3, 2)$, $F_2 = (2, 3, 1)$ and
$F_3 = (3, 1, 2)$.

When using the Attitude SPE algorithm, at each node $x \in P_i$, $i = 1, 2, 3$, the ith player has to choose the
alternative marked in bold violet in Figure 1. Note that

$$H^{x_3}(\underline{u}^{x_3}) = \begin{pmatrix} 12 \\ 0 \\ 0 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 0 \\ 24 \\ 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 24 \\ 0 \\ 24 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 0 \\ 18 \\ 12 \end{pmatrix} = \begin{pmatrix} 24 \\ 10 \\ 16 \end{pmatrix} \ and \ H^{z_2} = h(z_2) = \begin{pmatrix} 0 \\ 10 \\ 20 \end{pmatrix}.$$

Hence, $S^{2,1}(x_2^1) = \{z_2, x_3\}$, and $\underline{u}_2(x_2^1) = z_2$ due to the player's 2 attitude vector F_2.

The A-SPE algorithm generates unique SPE $\underline{u} = (\underline{u}_1, \underline{u}_2, \underline{u}_3)$, where $\underline{u}_1(x_0) = x_1$, $\underline{u}_1(x_4^2) = z_8$;
$\underline{u}_2(x_2^1) = z_2$, $\underline{u}_2(x_5) = z_9$; $\underline{u}_3(x_2^2) = z_4$, $\underline{u}_3(x_4^3) = z_6$, while $H(\underline{u}) = (11, 22, 18)$. We will use this SPE later
in Section 4 when calculating the γ-characteristic function.
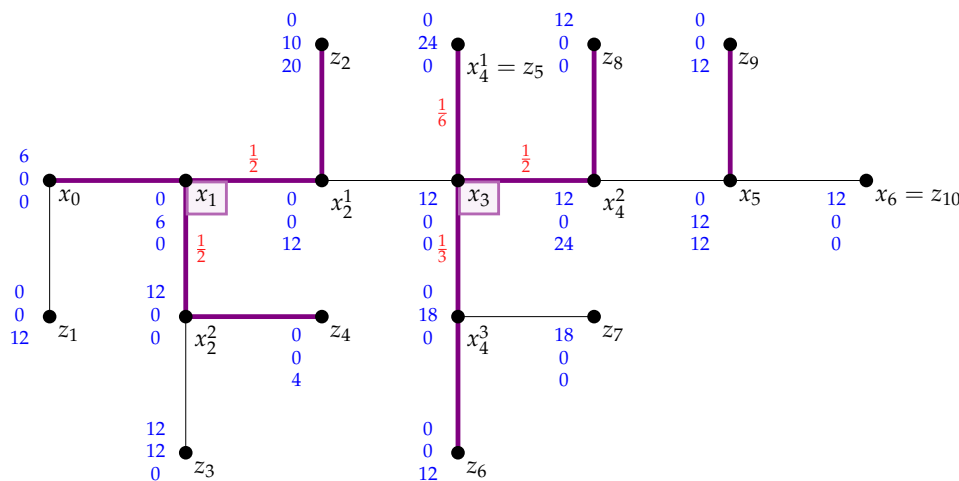

Figure 1. 3-person extensive-form game: A-Subgame Perfect Equilibria (SPE) algorithm implementation.

4. Cooperative Strategies and Trajectories

If the players agree to cooperate in multicriteria game Γ^{x_0}, first they are expected to maximize the
total payoff $\sum\limits_{i=1}^{n} H_i(u)$ of the grand coalition. Let $\overline{U}(\Gamma^{x_0})$ denote the set of all pure strategy profiles u,
such that

$$\sum_{i \in N} H_i(u) = \max_{v \in U} \sum_{i \in N} H_i(v) = \overline{H}. \tag{16}$$

The set $\overline{U}(\Gamma^{x_0})$ is known to be nonempty and it may contain multiple strategy profiles (see,
e.g., Reference [17]). Hence, the players need to agree on a specific approach they are going to use to
choose a unique optimal cooperative strategy profile $\overline{u} \in \overline{U}(\Gamma^{x_0})$ as well as the corresponding optimal
bundle of cooperative trajectories in the game tree. To this aim we introduce the so-calle *Players' Rank
Based (PRB) algorithm.* Note that rather close approach—using the ranking of the criteria to choose a
unique cooperative trajectory—was proposed recently in Reference [8] for multicriteria extensive-form
games without chance moves. Namely, suppose that the players have agreed on the so-called "rank"
of each player within the grand coalition N, and $r(k) = i$ means that the rank of player i equals k,
$k = 1, ..., n$.

　　Players' rank based (PRB) algorithm.

Step 0. Consider the set $\overline{U}(\Gamma^{x_0})$. If all strategy profiles $u \in \overline{U}(\Gamma^{x_0})$ generate the same bundle of trajectories $\Omega(u)$ (see, e.g., References [17,22,23] for discussion on a certain redundancy of the pure strategy definition in extensive game), let the players choose any strategy profile $\overline{u} \in \overline{U}(\Gamma^{x_0})$ as the cooperative strategy profile and $\Omega(\overline{u})$ denote the corresponding bundle of cooperative trajectories.

Step $k = 1$. Otherwise, that is, if the strategy profiles from $\overline{U}(\Gamma^{x_0})$ generate different (and hence, disjoint—see Remark 1) bundles of the trajectories, calculate

$$\max_{v \in \overline{U}(\Gamma^{x_0})} H_{r(1)}(v) = \overline{H}_{r(1)}.$$

Let $\overline{U}_{r(1)}(\Gamma^{x_0})$ denote the set of all strategy profiles u such that $H_{r(1)}(u) = \overline{H}_{r(1)}$. If all strategy profiles $u \in \overline{U}_{r(1)}(\Gamma^{x_0})$ generate the same bundle of trajectories $\Omega(u)$, the players may choose any strategy profile $\overline{u} \in \overline{U}_{r(1)}(\Gamma^{x_0})$ as the cooperative strategy profile. Otherwise proceed to the next step.

Step $k = 2$. Consider the set $\overline{U}_{r(1)}(\Gamma^{x_0})$. If all strategy profiles $u \in \overline{U}_{r(1)}(\Gamma^{x_0})$ generate the same bundle of trajectories $\Omega(u)$, the players may choose any strategy profile $\overline{u} \in \overline{U}_{r(1)}(\Gamma^{x_0})$ as the cooperative strategy profile. Otherwise, proceed to the next step.

Step k ($k = 2, \dots, n$).

$\vdots$

Step $k + 1$. Finally, if the strategy profiles from $u \in \overline{U}_{r(n)(\Gamma^{x_0})}$ generate different bundles of the trajectories, we suppose that the players choose such $\overline{u} \in \overline{U}_{r(n)(\Gamma^{x_0})}$ that $\Omega(\overline{u}) = \{\omega_m(\overline{u}) = (x_0, \dots, x_{T(m)} = z_l) \mid p(\omega_m, \overline{u}) > 0\}$ contains the trajectory $\omega(\overline{u})$ with minimal number l of the terminal node z_l (see Remark 1).

Henceforth, we will refer to the strategy profile $\overline{u} \in \overline{U}(\Gamma^{x_0})$ and the bundle of the trajectories $\Omega(\overline{u})$ as the *optimal cooperative strategy profile* and the *optimal bundle of cooperative trajectories* respectively.

In the dynamic setting it is significant that a specific method which the players agreed to accept in order to choose a unique optimal cooperative strategy profile $\overline{u} \in \overline{U}(\Gamma^{x_0})$ as well as the corresponding optimal bundle of cooperative trajectories satisfies time consistency (see, e.g., References [1,2,6,13,17]), that is, a fragment of the optimal bundle of the cooperative trajectories in the subgame should remain optimal in this subgame. Suppose that at each subgame $\Gamma^{\overline{x}_t}$ along the cooperative trajectories, that is $\overline{x}_t \in \omega(\overline{u})$, $\omega(\overline{u}) \in \Omega(\overline{u})$, the players choose the strategy profile $u^{\overline{x}_t} \in U^{\overline{x}_t}$ such that

$$u^{\overline{x}_t} \in \arg\max_{v^{\overline{x}_t} \in U^{\overline{x}_t}} \sum_{i \in N} H_i^{\overline{x}_t}(v^{\overline{x}_t}). \tag{17}$$

Let $\overline{U}(\Gamma^{\overline{x}_t})$ denote the set of all pure strategy profiles $u^{\overline{x}_t} \in U^{\overline{x}_t}$ which satisfy (17) and the players use the same approach to choose a unique optimal cooperative strategy profile $\overline{u}^{\overline{x}_t} \in \overline{U}(\Gamma^{\overline{x}_t})$ in the subgame as for the original game Γ^{x_0} (namely, the PRB algorithm).

Proposition 2. *A cooperative strategy profile for $\Gamma^{x_0} \in G^{cm}(n)$ based on the PRB algorithm satisfies time consistency. Namely, let $\overline{u} \in U$ satisfies (16), and $\Omega(\overline{u})$ be the optimal bundle of cooperative trajectories. Then for each subgame $\Gamma^{\overline{x}_t}$, $\overline{x}_t \in \omega(\overline{u}) = (\overline{x}_0, \dots, \overline{x}_t, \overline{x}_{t+1}, \dots, \overline{x}_T), 1 \leqslant t < T$, with $\overline{x}_0 = x_0$, $\omega(\overline{u}) \in \Omega(\overline{u})$, it holds that*

$$\sum_{i \in N} H_i^{\overline{x}_t}(\overline{u}^{\overline{x}_t}) = \max_{v^{\overline{x}_t} \in U^{\overline{x}_t}} \sum_{i \in N} H_i^{\overline{x}_t}(v^{\overline{x}_t}), \tag{18}$$

while $\omega^{\overline{x}_t} = (\overline{x}_t, \overline{x}_{t+1}, \dots, \overline{x}_T) \in \Omega(\overline{u}^{\overline{x}_t})$, that is, $\omega^{\overline{x}_t}$ belongs to the optimal bundle of cooperative trajectories in the subgame $\Gamma^{\overline{x}_t}$.

Proof. The optimal bundle of cooperative trajectories $\Omega(\overline{u})$ generated by $\overline{u} \in \overline{PO}(\Gamma^{x_0})$ can be divided onto two subsets $\{\Psi_m\} = \{\omega \in \Omega(\overline{u}) \mid \overline{x}_t \in \omega\}$ and $\{\chi_l\} = \{\omega \in \Omega(\overline{u}) \mid \overline{x}_t \notin \omega\}$ while $\{\Psi_m\} \cap \{\chi_l\} = \varnothing$, $\{\Psi_m\} \cup \{\chi_l\} = \Omega(\overline{u})$. Then, taking (5) and (6) into account we get

$$
\begin{aligned}
H_i(\overline{u}) &= \sum_m p(\Psi_m, \overline{u}) \cdot \tilde{h}_i(\Psi_m) + \sum_l p(\chi_l, \overline{u}) \cdot \tilde{h}_i(\chi_l) = \\
&= p(\overline{x}_t, \overline{u}) \cdot \left[\tilde{h}_i(\overline{x}_0, \overline{x}_1, \ldots, \overline{x}_{t-1}) + H_i^{\overline{x}_t}(\overline{u}^{\overline{x}_t}) \right] + \sum_l p(\chi_l, \overline{u}^D) \cdot \tilde{h}_i(\chi_l),
\end{aligned}
\tag{19}
$$

and (16) for $\overline{u}$ takes the form

$$
\begin{aligned}
\sum_{i \in N} p(\overline{x}_t, \overline{u}) \cdot \left(\tilde{h}_i(\overline{x}_0, \overline{x}_1, \ldots, \overline{x}_{t-1}) + H_i^{\overline{x}_t}(\overline{u}^{\overline{x}_t}) \right) + \\
+ \sum_{i \in N} \sum_l p(\chi_l, \overline{u}^D) \cdot \tilde{h}_i(\chi_l) = \max_{v \in U} H_i(v).
\end{aligned}
\tag{20}
$$

Suppose that $\overline{u}^{\overline{x}_t}$ does not satisfy (18), that is, there exists $v^{\overline{x}_t} \in U^{\overline{x}_t}$ such that

$$
\sum_{i \in N} H_i^{\overline{x}_t}(\overline{u}^{\overline{x}_t}) < \sum_{i \in N} H_i^{\overline{x}_t}(v^{\overline{x}_t}).
\tag{21}
$$

Denote by $\Omega(v^{\overline{x}_t}) = \{\lambda_m^{\overline{x}_t} = (\overline{x}_t, \ldots, \overline{x}_{T(m)}) \mid p(\lambda_m^{\overline{x}_t}, v^{\overline{x}_t}) > 0\}$ the bundle of all trajectories in the subgame $\Gamma^{\overline{x}_t}$ generated by $v^{\overline{x}_t}$. Then (21) takes the form

$$
\sum_{i \in N} \sum_m p(\Psi_m^{\overline{x}_t}, \overline{u}^{\overline{x}_t}) \cdot \tilde{h}_i^{\overline{x}_t}(\Psi_m^{\overline{x}_t}) < \sum_{i \in N} \sum_m p(\lambda_m^{\overline{x}_t}, v^{\overline{x}_t}) \cdot \tilde{h}_i^{\overline{x}_t}(\lambda_m^{\overline{x}_t}).
\tag{22}
$$

Denote by $W_i = (\overline{u}_i^D, v_i^{\overline{x}_t})$, $i \in N$, the ith player's compound pure strategy in Γ^{x_0}. The strategy profile $W = (W_1, \ldots, W_n)$ generates the strategy bundle $\Omega(W)$ that can be divided onto two disjoint subsets $\{\lambda_m\} = \{\omega \in \Omega(W) \mid \overline{x}_t \in \omega\}$ and $\{\chi_l\} = \{\omega \in \Omega(W) \mid \overline{x}_t \notin \omega\}$, where the second subset for $\Omega(W)$ coincides with the second subset for $\Omega(\overline{u})$ since $W^D = u^D$, and $\lambda_m = (\overline{x}_0, \ldots, \overline{x}_t) \cup (\overline{x}_t, \ldots, \overline{x}_{T(m)}) = (\overline{x}_0, \ldots, \overline{x}_t) \cup \lambda_m^{\overline{x}_t}$.

Adding $\sum_{i \in N} \tilde{h}_i(\overline{x}_0, \overline{x}_1 \ldots, \overline{x}_{t-1})$ to both sides of (22) we get

$$
\sum_{i \in N} \left(\tilde{h}_i(\overline{x}_0, \ldots, \overline{x}_{t-1}) + H_i^{\overline{x}_t}(\overline{u}^{\overline{x}_t}) \right) < \sum_{i \in N} \left(\tilde{h}_i(\overline{x}_0, \ldots, \overline{x}_{t-1}) + H_i^{\overline{x}_t}(v^{\overline{x}_t}) \right).
\tag{23}
$$

Then we can multiply both sides of (23) on $p(\overline{x}_t, \overline{u}) = p(\overline{x}_t, \overline{u}^D) = p(\overline{x}_t, W^D) = p(\overline{x}_t, W) > 0$ and then add $\sum_{i \in N} \sum_l p(\chi_l, \overline{u}^D) \cdot \tilde{h}_i(\chi_l)$ to both sides of the last inequality. Taking into account (4)–(6) and (20) we obtain

$$
\sum_{i \in N} H_i(\overline{u}) < \sum_{i \in N} H_i(W)
$$

for the constructed strategy profile $W \in U$. The last inequality contradicts the fact that $\overline{u} \in \overline{U}(\Gamma^{x_0})$, hence (18) is valid.

Arguing in a similar way (for the case when different strategy profiles from $\overline{U}(\Gamma^{\overline{x}_t})$ generate different bundles of the trajectories) we can verify that $\omega^{\overline{x}_t} = (\overline{x}_t, \ldots, \overline{x}_T)$ — a fragment of the cooperative trajectory $\omega \in \Omega(\overline{u})$, starting at $\overline{x}_t$ — belongs to the optimal bundle of cooperative trajectories in the subgame $\Gamma^{\overline{x}_t}$, that is, $\omega^{\overline{x}_t} \in \Omega(\overline{u}^{\overline{x}_t})$. □

We will assume in this paper that all the players have agreed to apply the PRB algorithm in order to choose the cooperative strategy profile $\overline{u} = (\overline{u}_1, \ldots, \overline{u}_n)$ that generates the optimal bundle $\Omega(\overline{u})$ of cooperative trajectories in $\Gamma^{x_0} \in G^{cm}(n)$. The next step of cooperation is to define a characteristic function $V^{x_0}(S)$. There are different notions of characteristic functions (see, e.g., References [23,24,48]),

in this paper we adopt the so-called γ-characteristic function introduced in Reference [24]. Namely, we assume that $V^{x_0}(S)$ is given by the SPE (based on the Attitude SPE algorithm) outcome of S in the noncooperative game between members of S maximizing their joint payoff, and non members playing individually.

The γ-characteristic function $V^{\overline{x}_t}$ for the subgame $\Gamma^{\overline{x}_t}$, $\overline{x}_t \in \omega_m(\overline{u}) = (\overline{x}_0, \ldots, \overline{x}_t, \ldots, \overline{x}_{T(m)})$, $\omega_m(\overline{u}) \in \Omega(\overline{u})$ along the optimal bundle of cooperative trajectories can be constructed using the same approach. Note that

$$V^{\overline{x}_t}(N) = \sum_{\omega_m^{\overline{x}_t} \in \Omega(\overline{u}^{\overline{x}_t})} p(\omega_m^{\overline{x}_t}, \overline{u}^{\overline{x}_t}) \cdot \sum_{\tau=t}^{T(m)} \sum_{i \in N} h_i(\overline{x}_\tau), \ t = 0, 1, \ldots, T(m). \tag{24}$$

Let $\Gamma^{x_0}(N, V^{x_0})$ denote extensive-form cooperative game $\Gamma^{x_0} \in G^{cm}(n)$ with γ-characteristic function, and $\Gamma^{\overline{x}_t}(N, V^{\overline{x}_t})$ denote the corresponding subgame.

We assume that the players adopt a single-valued cooperative solution φ^{x_0} (for instance, the Shapley value [33], the nucleolus [34], etc.) for the cooperative game $\Gamma^{x_0}(N, V^{x_0})$ which satisfies the collective rationality property

$$\sum_{i=1}^{n} \varphi_i^{x_0} = V^{x_0}(N) = \sum_{\omega_m \in \Omega(\overline{u})} p(\omega_m, \overline{u}) \cdot \sum_{\tau=0}^{T(m)} \sum_{i \in N} h_i(\overline{x}_\tau) \tag{25}$$

and the individual rationality property

$$\varphi_i^{x_0} \geqq V^{x_0}(\{i\}), \ i = 1, \ldots, n. \tag{26}$$

In addition, we assume that the same properties (25) and (26) are valid for the cooperative solutions $\varphi^{\overline{x}_t}$ at each subgame $\Gamma^{\overline{x}_t}(N, V^{\overline{x}_t})$, $t = 0, \ldots, T - 1$.

It is worth noting that the last assumption as well as the choice of γ-characteristic function ensure that every player has an incentive to cooperate at each subgame along the optimal game evolution since the ith player's cooperative payoff-to-go at $\Gamma^{\overline{x}_t}(N, V^{\overline{x}_t})$, $t = 0, \ldots, T - 1$, is at least equal to her non-cooperative counterpart: $\varphi_i^{\overline{x}_t} \geqslant H_i^{\overline{x}_t}(\underline{u}^{\overline{x}_t})$.

5. Subgame Consistency and Incremental IDP

Let $\beta = \{\beta_i(\overline{x}_\tau)\}$, $i = 1, \ldots, n$; $\tau = 1, \ldots, T(l)$, $\overline{x}(\tau) \in \omega_l(\overline{u})$, $\omega_l(\overline{u}) \in \Omega(\overline{u})$ denote the Imputation Distribution Procedure (IDP) for the cooperative solution $\left(\varphi_i^{\overline{x}_0}\right)_{i \in N}$ or the payment schedule (see, e.g., References [3,8–12,14–18,20] for details). The IDP approach means that all the players have agreed to allocate the total cooperative payoff $V^{x_0}(N)$ between the players along the optimal bundle $\Omega(\overline{u})$ of cooperative trajectories $\omega_l(\overline{u})$ according to some specific rule which is called IDP. Namely, $\beta_i(\overline{x}_\tau)$ denotes the actual current payment which the player i receives at position $\overline{x}_\tau$ (instead of $h_i(\overline{x}_\tau)$) if the players employ the IDP β. Moreover, one can design such an IDP β that all the players will be interested in cooperation in any subgame $\Gamma^{\overline{x}_\tau}$, $\overline{x}(\tau) \in \omega_l(\overline{u})$, $\omega_l(\overline{u}) \in \Omega(\overline{u})$, that is, at any intermediate time instant.

Definition 3. *The IDP $\beta = \{\beta_i(\overline{x}_\tau)\}$ satisfies subgame efficiency, if at any intermediate node $\overline{x}_t \in \omega(\overline{u})$, $\omega(\overline{u}) \in \Omega(\overline{u})$, $0 \leqslant t < T$, it holds that:*

$$\sum_{\omega_m^{\overline{x}_t} \in \Omega(\overline{u}^{\overline{x}_t})} p(\omega_m^{\overline{x}_t}, \overline{u}^{\overline{x}_t}) \cdot \sum_{\tau=t}^{T(m)} \beta_i(\overline{x}_\tau) = \varphi_i^{\overline{x}_t}, \ i \in N. \tag{27}$$

Equation (27) means that the expected sum of the payments to player i along the optimal subgame $\Gamma^{\overline{x}_t}$ evolution equals to what she is entitled to in this subgame. Then the IDP for each player can

be reasonably implemented as a rule for step-by-step allocation of the ith player's current expected optimal payoff. Note that for $t = 0$ the subgame efficiency definition coincides with the efficiency at initial node x_0 or the efficiency in the whole game Γ^{x_0} condition (see References [9,14,16,20]).

Definition 4 ([10]). *The IDP $\beta = \{\beta_i(\bar{x}_\tau)\}$ satisfies the strict balance condition if for each node $\bar{x}_\tau \in \omega_m(\bar{u})$, $\omega_m(\bar{u}) \in \Omega(\bar{u})\ \forall t = 0, \ldots, T(m)$*

$$\sum_{i \in N} \beta_i(\bar{x}_\tau) = \sum_{i \in N} h_i(\bar{x}_\tau). \tag{28}$$

Equation (28) ensures the "admissibility" of the IDP, that is, the sum of payments to the players in any node $\bar{x}_\tau$ is equal to the sum of payoffs that they can collect in this node.

The next advantageous dynamic property of an IDP—the time consistency, introduced in Reference [3]—was extended to dynamic games played over event trees in References [14,16,20] as well as to multicriteria extensive-form cooperative games (with chance moves) in Reference [9].

To write down properly the time consistency condition for some intermediate node $\bar{x}_t \in \omega(\bar{u}) = (\bar{x}_0, \bar{x}_1, \ldots, \bar{x}_{t-1}, \bar{x}_t, \bar{x}_{t+1}, \ldots, \bar{x}_T), \omega(\bar{u}) \in \Omega(\bar{u}), 1 \leqslant t < T$, in multistage game Γ^{x_0} with chance moves we need to pay attention to all chance nodes on the path $(\bar{x}_0, \ldots, \bar{x}_{t-1}) = \underline{\omega}^{x_t} \setminus \{\bar{x}_t\}$.
Namely, let us numerate the chance nodes from $P_0 \cap (\underline{\omega}^{x_t} \setminus \{\bar{x}_t\})$ in order of their occurrence on the path $(\bar{x}_0, \ldots, \bar{x}_{t-1})$, that is, $y_1 = \bar{x}_{t(1)}, y_2 = \bar{x}_{t(2)}, \ldots, y_\theta = \bar{x}_{t(\theta)}, 0 \leqslant t(1) < t(2) < \ldots < t(\theta) < t$.

Definition 5 ([9]). *The IDP $\beta = \{\beta_{i/k}(\bar{x}_\tau)\}$ for the cooperative solution φ^{x_0} is called time consistent in the whole game $\Gamma^{x_0}(N, V^{x_0}) \in G^{cm}(n)$ if at any intermediate node $\bar{x}_t \in \omega(\bar{u}), \omega(\bar{u}) \in \Omega(\bar{u}), 1 \leqslant t < T$, for all $i \in N$, it holds that*

<u>*case $\theta = 0$*</u> *(no chance nodes on the path $(\bar{x}_0, \ldots, \bar{x}_{t-1})$):*

$$\sum_{\tau=0}^{t-1} \beta_i(\bar{x}_\tau) + \varphi_i^{\bar{x}_t} = \varphi_i^{x_0}, \tag{29}$$

<u>*case $\theta = 1$*</u> *(only one chance node $y_1 = \bar{x}_{t(1)}$ before $\bar{x}_t$):*

$$\sum_{\tau=0}^{t(1)} \beta_i(\bar{x}_\tau) + p(\bar{x}_{t(1)+1}, \bar{u}) \cdot \left\{ \sum_{\tau=t(1)+1}^{t-1} \beta_i(\bar{x}_\tau) + \varphi_i^{\bar{x}_t} \right\} + \\
+ \sum_{x^k \in S(\bar{x}_{t(1)}) \setminus \{\bar{x}_{t(1)+1}\}} p(x^k, \bar{u}) \cdot \varphi_i^{x^k} = \varphi_i^{x_0}, \tag{30}$$

<u>*case $\theta = 2$*</u> *(two chance nodes $y_1 = \bar{x}_{t(1)}, y_2 = \bar{x}_{t(2)}$ before $\bar{x}_t$):*

$$\sum_{\tau=0}^{t(1)} \beta_i(\bar{x}_\tau) + p(\bar{x}_{t(1)+1}, \bar{u}) \cdot \left\{ \sum_{\tau=t(1)+1}^{t(2)} \beta_i(\bar{x}_\tau) + p(\bar{x}_{t(2)+1} \mid \bar{x}_{t(2)}, \bar{u}) \times \right. \\
\times \left[\sum_{\tau=t(2)+1}^{t-1} \beta_i(\bar{x}_\tau) + \varphi_i^{\bar{x}_t} \right] + \sum_{x^m \in S(\bar{x}_{t(2)}) \setminus \{\bar{x}_{t(2)+1}\}} p(x^m \mid \bar{x}_{t(2)}, \bar{u}) \cdot \varphi_i^{x^m} \right\} + \\
+ \sum_{x^k \in S(\bar{x}_{t(1)}) \setminus \{\bar{x}_{t(1)+1}\}} p(x^k, \bar{u}) \cdot \varphi_i^{x^k} = \varphi_i^{x_0}, \tag{31}$$

...

Note that for partial case when $\overline{x}_t \in S(\overline{x}_{t(1)})$, that is, if $\overline{x}_t$ follows the chance node $\overline{x}_{t(1)}$ Equation (30) takes the simpler form

$$\sum_{\tau=0}^{t(1)} \beta_i(\overline{x}_\tau) + \sum_{x^k \in S(\overline{x}_{t(1)})} p(x^k, \overline{u}) \cdot \varphi_i^{x^k} = \varphi_i^{x_0}.$$

A similar note is valid for equation (31), and so forth.

Roughly speaking, Definition 5 implies that the payments collected by the ith player (according to the payment schedule β) before reaching some intermediate node $\overline{x}_t$ plus the expected ith player's component of the Shapley value in the subgame $\Gamma^{\overline{x}_t}$ starting at $\overline{x}_t$ plus this player's expected Shapley value components in other subgames along the cooperative trajectories which do not contain $\overline{x}_t$ corresponds to what the player i is entitled to in the original game $\Gamma^{x_0}(N, V^{x_0})$.

It is worth noting that Definition 5 indeed provides a reasonable consistency requirements which a good payment schedule β should satisfy when the player evaluates IDP β at the initial node x_0, that is, before the game $\Gamma^{x_0}(N, V^{x_0})$ starts (and the words "in the whole game" in Definition 5 properly reflect this feature). However, when the player purposes to evaluate IDP β in the subgame $\Gamma^{\overline{x}_t}$, that is, after reaching some intermediate node $\overline{x}_t$ (in case when $\theta \geqslant 1$) this player will unlikely take into account the expected future payoffs in all the subgames which are unattainable if the node $\overline{x}_t$ has been already reached, that is, the last addends in the LHS of (30) and (31). To overcome this problem we suggest the players to use a notion of subgame consistency—a refinement of time consistency that was firstly proposed in Reference [36] for cooperative stochastic differential games and then extend it to stochastic dynamic games in References [37,38]. Let us provide a rigorous definition of the IDP subgame consistency for extensive-form games with chance moves that is applicable in all the subgames along the optimal bundle of cooperative trajectories.

Definition 6. *The IDP* $\beta = \{\beta_i(\overline{x}_\tau)\}$ *is called subgame consistent if at any intermediate node* $\overline{x}_t \in \omega(\overline{u})$, $\omega(\overline{u}) \in \Omega(\overline{u})$, $1 \leqslant t \leqslant T$, *for all* $i \in N$, *it holds that*

case $1 \leqslant t \leqslant t(1)$ *(no chance nodes before the subgame* $\Gamma^{\overline{x}_t}$ *root* $\overline{x}_t$):

$$\sum_{\tau=0}^{t-1} \beta_i(\overline{x}_\tau) + \varphi_i^{\overline{x}_t} = \varphi_i^{x_0}, \tag{32}$$

case $t(1)+1 < t \leqslant t(2)$ *(only one chance node* $y_1 = \overline{x}_{t(1)}$ *before* $\overline{x}_t$):

$$\sum_{\tau=t(1)+1}^{t-1} \beta_i(\overline{x}_\tau) + \varphi_i^{\overline{x}_t} = \varphi_i^{\overline{x}_{t(1)+1}}, \tag{33}$$

case $t(2)+1 < t \leqslant t(3)$ *(two chance nodes before* $\overline{x}_t$):

$$\sum_{\tau=t(2)+1}^{t-1} \beta_i(\overline{x}_\tau) + \varphi_i^{\overline{x}_t} = \varphi_i^{\overline{x}_{t(2)+1}}, \tag{34}$$

$\vdots$

case $t(\theta)+1 < t \leqslant T$ *(no chance nodes after* $\overline{x}_t$):

$$\sum_{\tau=t(\theta)+1}^{t-1} \beta_i(\overline{x}_\tau) + \varphi_i^{\overline{x}_t} = \varphi_i^{\overline{x}_{t(\theta)+1}}. \tag{35}$$

The subgame consistency definition differs from the "time consistency in the whole game" property (see References [9,14,16,20]) which is based on an a priori assessment of the ith player's

expected optimal payoff (before the game starts). However, when the players make a decision in the subgame after the chance move occurs they need to recalculate the expected optimal payoff since the original optimal bundle of cooperative trajectories shrinks after each chance node. Note that we can not write out the subgame consistency condition for $t = t(1) + 1, t(2) + 1, ..., t(\theta) + 1$, that is, for the nodes $\bar{x}_t$ that immediately follow the chance nodes.

One can suggest different imputation distribution procedures that may or may not satisfy the useful properties listed above. The review of different IDP for multistage games (without chance moves) as well as the analysis of their properties can be found in References [10,12,15,17]. Below we consider the refinement of the so-called incremental IDP (see, e.g., References [10,14,16,17,20,21]) that was recently introduced for multistage games with chance moves [9].

Definition 7 ([9]). *The incremental IDP for the cooperative solution φ^{x_0} in multistage game with chance moves Γ^{x_0} is defined as follows:*

$$\beta_i(x_t) = \varphi_i^{x_t} - \sum_{x_{t+1}^k \in S(x_t)} p(x_{t+1}^k | x_t, \bar{u}) \cdot \varphi_i^{x_{t+1}^k} \tag{36}$$

for $x_t \in \omega_l(\bar{u}) = (x_0, \ldots, x_t, \ldots, x_{T(l)})$, $\omega_l(\bar{u}) \in \Omega(\bar{u})$, $t = 0, \ldots, T(l) - 1$;

$$\beta_i(x_{T(l)}) = \varphi_i^{x_{T(l)}} \tag{37}$$

for $x_{T(l)} \in \Omega(\bar{u}) \cap P_{n+1}$.

Remark 2. *Formulas (36), (37) are similar to the imputation distribution procedures suggested in References [14,16,20] for (single-criterion) stochastic discrete-time dynamic games played over event trees. If $x_t \in P_i$, $i = 1, \ldots, n$ Equation (36) takes the simpler form $\beta_i(x_t) = \varphi_i^{x_t} - \varphi_i^{x_{t+1}}$, where $\bar{u}_i(x_t) = x_{t+1}$, that coincides with the "classical" incremental IDP.*

Let us use again 3-person extensive-form game from Example 1 to demonstrate a proposed scheme of cooperation.

Example 2. *(Cooperative behavior in 3-player game from Ex. 1).*
Suppose that the players have agreed on the following ranks: $r(1) = 1$, $r(2) = 2$ and $r(3) = 3$. When implementing the PRB algorithm we get the optimal bundle $\Omega(\bar{u})$ which contains four cooperative trajectories (marked in bold, deep blue in Figure 2): $\omega_1 = (\bar{x}_0, \bar{x}_1, \bar{x}_2^1, \bar{x}_3, \bar{x}_4^2, \bar{x}_5, \bar{x}_6)$, $\omega_2 = (\bar{x}_0, \bar{x}_1, \bar{x}_2^1, z_3)$, $\omega_3 = (\bar{x}_0, \bar{x}_1, \bar{x}_2^1, \bar{x}_3, x_4^1)$ and $\omega_4 = (\bar{x}_0, \bar{x}_1, \bar{x}_2^1, \bar{x}_3, x_4^3, z_7)$. Note that players use the ranks when making decision at node $\bar{x}_5$.
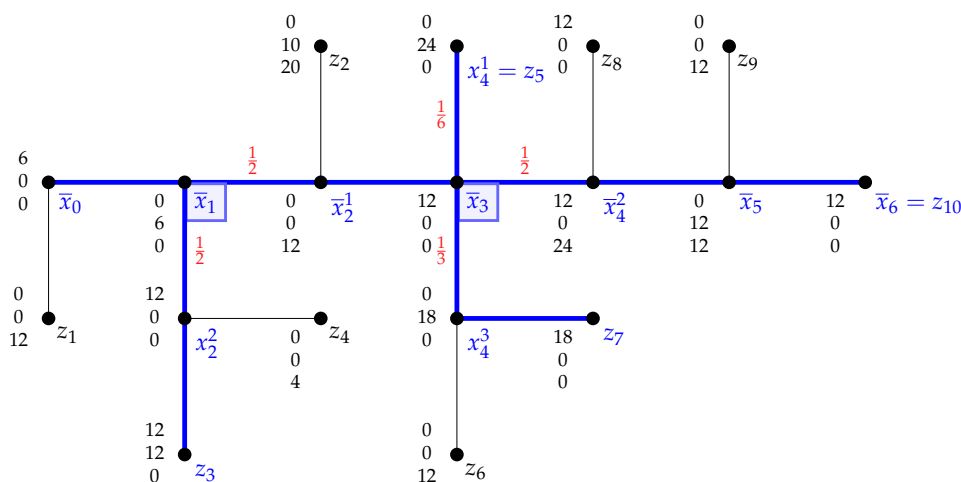

Figure 2. 3-player extensive-form game: cooperative behavior.

To demonstrate the implementation of the incremental IDP and its properties we will adopt the Shapley value as a single valued cooperative solution. The values of the γ-characteristic function V^{x_0} for the original game $\Gamma^{x_0}(N, V^{x_0})$ and the Shapley value φ^{x_0} are

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{x_0}(S)$	12	11	18	38	48	33	68

$$\varphi^{x_0} = \begin{pmatrix} 25\frac{1}{6} \\ 17\frac{1}{6} \\ 25\frac{2}{3} \end{pmatrix}.$$

Consider, for instance, the incremental IDP along the longest cooperative trajectory $\omega_2 = (\bar{x}_0, \dots, \bar{x}_6)$ from $\Omega(\bar{u})$. If we calculate γ-characteristic functions using Attitude SPE algorithm for the subgames, we get the following results.

Subgame $\Gamma^{\bar{x}_1}(N, V^{\bar{x}_1})$:

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{\bar{x}_1}(S)$	6	11	18	32	42	33	62

$$\varphi^{\bar{x}_1} = \begin{pmatrix} 19\frac{1}{6} \\ 17\frac{1}{6} \\ 25\frac{2}{3} \end{pmatrix}.$$

Subgame $\Gamma^{\bar{x}_2^1}(N, V^{\bar{x}_2^1})$:

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{\bar{x}_2^1}(S)$	0	10	32	40	60	42	76

$$\varphi^{\bar{x}_2^1} = \begin{pmatrix} 21 \\ 17 \\ 38 \end{pmatrix}.$$

Subgame $\Gamma^{x_2^2}(N, V^{x_2^2})$:

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{x_2^2}(S)$	12	0	4	12	24	12	36

$$\varphi^{x_2^2} = \begin{pmatrix} 17\frac{1}{3} \\ 5\frac{1}{3} \\ 13\frac{1}{3} \end{pmatrix}.$$

Subgame $\Gamma^{\bar{x}_3}(N, V^{\bar{x}_3})$:

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{\bar{x}_3}(S)$	24	10	16	40	48	26	64

$$\varphi^{\bar{x}_3} = \begin{pmatrix} 31 \\ 13 \\ 20 \end{pmatrix}.$$

Subgame $\Gamma^{\bar{x}_4^2}(N, V^{\bar{x}_4^2})$:

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{\bar{x}_4^2}(S)$	24	0	24	36	60	24	72

$$\varphi^{\bar{x}_4^2} = \begin{pmatrix} 36 \\ 6 \\ 30 \end{pmatrix}.$$

Subgame $\Gamma^{x_4^3}(N, V^{x_4^3})$:

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{x_4^3}(S)$	0	18	12	18	18	30	36

$$\varphi^{x_4^3} = \begin{pmatrix} 3 \\ 18 \\ 15 \end{pmatrix}.$$

Subgame $\Gamma^{\bar{x}_5}(N, V^{\bar{x}_5})$:

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{\bar{x}_5}(S)$	0	12	24	12	24	36	36

$$\varphi^{\bar{x}_5} = \begin{pmatrix} 0 \\ 12 \\ 24 \end{pmatrix}.$$

Finally, $\varphi^{\bar{x}_6} = h^{\bar{x}_6} = (12, 0, 0)$.

One can calculate the incremental IDP $\{\beta_i(\bar{x}_\tau), \ \bar{x}_\tau \in \omega_2\}$ using (36) and (37):

	$\bar{x}_0$	$\bar{x}_1$	$\bar{x}_2^1$	$\bar{x}_3$	$\bar{x}_4^2$	$\bar{x}_5$	$\bar{x}_6$
$\beta_1(\bar{x}_\tau)$	6	0	-10	12	36	-12	12
$\beta_2(\bar{x}_\tau)$	0	6	4	0	-6	12	0
$\beta_3(\bar{x}_\tau)$	0	0	18	0	6	24	0

Note that the subgame consistency conditions at nodes $\bar{x}_1$, $\bar{x}_3$ and $\bar{x}_5$ according to (32)–(34) respectively take the form:

$$\beta_i(\bar{x}_0) + \varphi_i^{\bar{x}_1} = \varphi_i^{x_0}, i \in N, \text{ or } \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 19\frac{1}{6} \\ 17\frac{1}{6} \\ 25\frac{2}{3} \end{pmatrix} = \begin{pmatrix} 25\frac{1}{6} \\ 17\frac{1}{6} \\ 25\frac{2}{3} \end{pmatrix},$$

$$\beta_i(\overline{x}_2^1) + \varphi_i^{\overline{x}_3} = \varphi_i^{x_2^1}, i \in N, \text{ or } \begin{pmatrix} -10 \\ 4 \\ 18 \end{pmatrix} + \begin{pmatrix} 31 \\ 13 \\ 20 \end{pmatrix} = \begin{pmatrix} 21 \\ 17 \\ 38 \end{pmatrix},$$

$$\beta_i(\overline{x}_4^2) + \varphi_i^{\overline{x}_5} = \varphi_i^{x_4^2}, i \in N, \text{ or } \begin{pmatrix} 36 \\ -6 \\ 6 \end{pmatrix} + \begin{pmatrix} 0 \\ 12 \\ 24 \end{pmatrix} = \begin{pmatrix} 36 \\ 6 \\ 30 \end{pmatrix}.$$

It is known that the classical incremental IDP for multistage (and differential) games may imply negative current payments to some players at some positions (see References [4,10,17,38] for details). As one can observe in Ex. 2, this drawback of the incremental IDP may appear in the extensive-form game with chance moves as well. Two approaches how to overcome this possible disadvantage were suggested in References [4,10]. Unfortunately, as it was firstly proved in Reference [10], in general it is impossible to design a time consistent IDP which satisfies both the balance condition and non-negativity constraint.

Proposition 3. *The incremental IDP (36), (37) satisfies strict balance condition (28), the subgame efficiency condition (27), and the subgame consistency conditions (32)–(35).*

Proof. Incremental IDP β was proved to satisfiy strict balance condition (28) in Reference [9]. The proof of subgame consistency can be carried out by direct verification. For instance, consider the case when $t(1) + 1 < t \leqslant t(2)$. Then, using Remark 2 we get

$$\sum_{\tau=t(1)+1}^{t-1} \beta_i(\overline{x}_\tau) = \left(\varphi_i^{\overline{x}_{t(1)+1}} - \varphi_i^{\overline{x}_{t(1)+2}} \right) + \ldots + \left(\varphi_i^{\overline{x}_{t-1}} - \varphi_i^{\overline{x}_t} \right) = \varphi_i^{\overline{x}_{t(1)+1}} - \varphi_i^{\overline{x}_t}.$$

Obviously, (33) is satisfied.

The proof that IDP (36), (37) satisfies subgame efficiency (27) is based on direct calculations but rather cumbersome in general case (i.e., for arbitrary game Γ^{x_0}). Let us demonstrate how it works for the game in Example 2. For instance we verify that the incremental IDP meets the subgame efficiency condition at node $\overline{x}_3$.

Note that $\Omega(\overline{u}^{\overline{x}_3}) = \left\{ \omega_1^{\overline{x}_3} = (\overline{x}_3, x_4^1); \omega_2^{\overline{x}_3} = (\overline{x}_3, \overline{x}_4^2, \overline{x}_5, \overline{x}_6); \omega_3^{\overline{x}_3} = (\overline{x}_3, x_4^3, z_7) \right\}$ while $p(\omega_1^{\overline{x}_3}, \overline{u}^{\overline{x}_3}) = \pi(x_4^1|\overline{x}_3)$, $p(\omega_2^{\overline{x}_3}, \overline{u}^{\overline{x}_3}) = \pi(\overline{x}_4^2|\overline{x}_3)$ and $p(\omega_3^{\overline{x}_3}, \overline{u}^{\overline{x}_3}) = \pi(x_4^3|\overline{x}_3)$. Then, using (32), (33), Remark 2, equality $\sum_{x_4^k \in S(\overline{x}_3)} \pi(x_4^k|\overline{x}_3) = 1$ and the notation

$$\Phi_i^4 = \sum_{k=1}^{3} \pi(x_4^k|\overline{x}_3) \cdot \varphi_i^{x_4^k},$$

we obtain

$$\sum_{\omega_k^{\overline{x}_3} \in \Omega(\overline{u}^{\overline{x}_3})} p(\omega_k^{\overline{x}_3}, \overline{u}^{\overline{x}_3}) \cdot \sum_{\tau=3}^{T(k)} \beta_i(\overline{x}_\tau) = \pi(x_4^1|\overline{x}_3) \cdot \left[\left(\varphi_i^{\overline{x}_3} - \Phi_i^4 \right) + \varphi_i^{x_4^1} \right] +$$

$$+ \pi(\overline{x}_4^2|\overline{x}_3) \cdot \left[\left(\varphi_i^{\overline{x}_3} - \Phi_i^4 \right) + \left(\varphi_i^{\overline{x}_4^2} - \varphi_i^{\overline{x}_5} \right) + \left(\varphi_i^{\overline{x}_5} - \varphi_i^{\overline{x}_6} \right) + \varphi_i^{\overline{x}_6} \right] +$$

$$+ \pi(\overline{x}_4^3|\overline{x}_3) \cdot \left[\left(\varphi_i^{\overline{x}_3} - \Phi_i^4 \right) + \left(\varphi_i^{x_4^3} - \varphi_i^{z_7} \right) + \varphi_i^{z_7} \right] = \varphi_i^{\overline{x}_3} \cdot \sum_{k=1}^{3} \pi(x_4^k|\overline{x}_3) +$$

$$+ \pi(x_4^1|\overline{x}_3) \cdot \left[-\Phi_i^4 + \varphi_i^{x_4^1} \right] + \pi(\overline{x}_4^2|\overline{x}_3) \cdot \left[-\Phi_i^4 + \varphi_i^{\overline{x}_4^2} \right] + \pi(x_4^3|\overline{x}_3) \cdot \left[-\Phi_i^4 + \varphi_i^{x_4^3} \right] =$$

$$= \varphi_i^{\overline{x}_3} - \Phi_i^4 \cdot \sum_{k=1}^{3} \pi(x_4^k|\overline{x}_3) + \sum_{k=1}^{3} \pi(x_4^k|\overline{x}_3) \cdot \varphi_i^{x_4^k} = \varphi_i^{\overline{x}_3}.$$

□

According to Proposition 3, the incremental payment schedule (36), (37) can be used to implement a long-term cooperative agreement in an extensive-form game with chance moves.

6. Conclusions

In the paper we purposes to design a mechanism of the players' sustainable long-term cooperation that satisfies a number of good properties. To this aim we formalised the players' rank based algorithm for selecting a unique optimal bundle of cooperative trajectories, and proved that corresponding cooperative strategy profile satisfies time consistency. To calculate γ-characteristic function one need to have a specific method for constructing a unique (subgame perfect) equilibrium at any extensive-form game with chance moves. Hence, we formalised a backwards induction procedure refinement based on the players' attitude vectors—the so-called attitude SPE algorithm.

As a result of reexamination of the "IDP time consistency in the whole game" concept, we suggest to adopt the concept of subgame consistency, introduced in Reference [36] for differential stochastic games and then extend it to dynamic stochastic games in References [37,38]. The definition of subgame consistency for extensive-form game with chance moves is provided. This property takes into account such an interesting feature of the games under consideration that when the players make a decision in the subgame Γ^{x_t} after the chance move occurs, they need to recalculate their expected optimal payoffs-to-go since the original optimal bundle of cooperative trajectories shrinks after each chance node. It is worth noting that a similar approach, based on the IDP subgame consistency notion could be applied to dynamic games played over event trees ([14,16,20]). We proved that the incremental IDP specified for multistage games with chance moves in Reference [9] satisfy subgame consistency and subgame efficiency as well as the strict balance condition.

It follows from Propositions 1–3 that two specified algorithms combined with the γ-characteristic function, and the incremental payment schedule together constitute a mechanism of the players' sustainable cooperation that satisfies a number of good properties and could be used in extensive-form games with chance moves. Note that the main result of the paper—Proposition 3—does not depend on the specific method which the players employ to calculate the characteristic function as well as on the specific single-valued cooperative solution meeting (25) and (26).

Since this is the first time that subgame consistent solutions are examined for extensive-form games with chance moves, further research along this line is expected. It is surely of interest to develop appropriate software application to implement proposed algorithms in arbitrary extensive-form game with chance moves. Possibly, one can use the so-called Game Theory Explorer [30] when developing such software tools for 2-person extensive games. Further, it might be interesting to run experiments with large-scale datasets, after the software application that allows to construct unique SPE, the optimal bundle of cooperative trajectories, γ-characteristic function, and so forth, will be developed.

Let us notice some preliminary suggestions on how one can use such software application to run simulations. First, one can vary the main parameter—the length of the game tree, and the additional parameters such as the game structure, the players' payoffs, probabilities of transitions, and so forth, to obtain practical estimations of the proposed algorithms complexity and scalability. Secondarily, one can generate external disturbances of the stage payoffs and probabilities and vary the players' attitude vectors to carry out the sensitivity analysis of the proposed non-cooperative and cooperative solutions. Further, it is of interest to get experimental estimations of the price of anarchy and the price of stability for the class of games under consideration. Finally, one can use such software application to check whether the additional properties (non-negativity, irrational-behavior-proof conditions, etc.) of the proposed incremental IDP and other payment schedules (see, e.g., Reference [15]) are satisfied for given extensive-form game with chance moves.

Author Contributions: Conceptualization, D.K.; methodology, D.K.; formal analysis, D.K.; investigation, D.K. and N.S.; writing—original draft preparation, D.K. and N.S.; writing—review and editing, D.K. and N.S.; visualization, D.K. and N.S.; supervision, D.K. All authors have read and agreed to the published version of the manuscript.

References

1. Haurie, A. A note on nonzero-sum diferential games with bargaining solution. *J. Optim. Theory Appl.* **1976**, *18*, 31–39. [CrossRef]
2. Petrosyan, L. Time-consistency of solutions in multi-player differential games. *Astronomy* **1977**, *4*, 46–52.
3. Petrosyan, L.A.; Danilov, N.N. Stability of solutions in non-zero sum differential games with transferable payoffs. *Astronomy* **1979**, *1*, 52–59.
4. Gromova, E.V.; Plekhanova, T.M. On the regularization of a cooperative solution in a multistage game with random time horizon. *Discret. Appl. Math.* **2019**, *255*, 40–55. [CrossRef]
5. Haurie, A.; Krawczyk, J. B.; Zaccour, G. *Games and Dynamic Games*; Scientific World: Singapore, 2012.
6. Kuzyutin, D. On the problem of the stability of solutions in extensive games. *Vestnik St. Petersburg Univ. Math.* **1995**, *4*, 18–23.
7. Kuzyutin, D. On the consistency of weak equilibria in multicriteria extensive games. In *Contributions to Game Theory and Management*; Petrosyan, L.A., Zenkevich, N.A., Eds.; St. Petersburg State Univ. Press: St. Petersburg, Russia, 2012; Volume V, pp. 168–177.
8. Kuzyutin, D.; Gromova, E.; Pankratova, Y. Sustainable cooperation in multicriteria multistage games. *Oper. Res. Lett.* **2018**, *46*, 557–562. [CrossRef]
9. Kuzyutin, D.; Gromova, E.; Smirnova, N. On the cooperative behavior in multistage multicriteria game with chance moves. In *Mathematical Optimization Theory and Operations Research (MOTOR 2020)*; Intern. Conference Proceedings; LNCS Series; Springer: Berlin, Germany, 2020; forthcoming.
10. Kuzyutin, D.; Nikitina, M. Time consistent cooperative solutions for multistage games with vector payoffs. *Oper. Res. Lett.* **2017**, *45*, 269–274. [CrossRef]
11. Kuzyutin, D.; Nikitina, M. An irrational behavior proof condition for multistage multicriteria games. In *Consrtuctive Nonsmooth Analysis and Related Topics (Dedic. to the Memory of V.F.Demyanov)*; CNSA 2017; IEEE: New York, NY, USA, 2017; pp. 178–181.
12. Petrosyan, L.A.; Danilov, N.N. *Cooperative Differential Games and Their Applications*; Publishing House of Tomsk University: Tomsk, Russia, 1985.
13. Petrosyan, L.A.; Kuzyutin, D.V. On the stability of E-equilibrium in the class of mixed strategies. *Vestn. St. Petersburg Univ. Math.* **1995**, *3*, 54–58.
14. Parilina, E.; Zaccour, G. Node-consistent Shapley value for games played over event trees with random terminal time. *J. Optim. Theory Appl.* **2017**, *175*, 236–254. [CrossRef]
15. Kuzyutin, D.; Smirnova, N.; Gromova, E. Long-term implementation of the cooperative solution in multistage multicriteria game. *Oper. Res. Perspect.* **2019**, *6*, 100107. [CrossRef]
16. Parilina, E.; Zaccour, G. Node-consistent core for games played over event trees. *Automatica* **2015**, *55*, 304–311. [CrossRef]
17. Petrosyan, L.; Kuzyutin, D. *Games in Extensive Form: Optimality and Stability*; Saint Petersburg University Press: St. Petersburg, Russia, 2000.
18. Petrosyan, L.; Zaccour, G. Time-consistent Shapley value allocation of pollution cost reduction. *J. Econ. Dyn. Control.* **2003**, *27*, 381–398. [CrossRef]
19. Sedakov, A. On the Strong Time Consistency of the Core. *Autom. Remote Control* **2018**, *79*, 757–767. [CrossRef]
20. Reddy, P.; Shevkoplyas, E.; Zaccour, G. Time-consistent Shapley value for games played over event trees. *Automatica* **2013**, *49*, 1521–1527. [CrossRef]
21. Zakharov, V.; Dementieva, M. Multistage Cooperative Games and Problem of Time Consistency. *Int. Game Theory Rev.* **2004**, *6*, 157–170. [CrossRef]
22. Kuhn, H. Extensive games and the problem of information. *Ann. Math.* **1953**, *28*, 193–216.
23. Myerson, R. *Game Theory. Analysis of Conflict*; Harvard University Press: Cambridge, MA, USA, 1997.
24. Chandler, P.; Tulkens, H. The core of an economy with multilateral environmental externalities. *Int. J. Game Theory* **1997**, *26*, 379–401. [CrossRef]

25. Selten, R. Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games. *Int. J. Game Theory* **1975**, *4*, 25–55. [CrossRef]

26. Petrosyan, L.; Zenkevich, N. *Game Theory*; World Scientific Publisher: Singapure, London, 1996.

27. Kuzyutin, D.; Osokina, O.; Romanenko, I. On the consistency of optimal behavior in extensive games. In *Game Theory and Applications*; Petrosyan, L., Mazalov V., Eds.; Nova Science Publ.: New York, NY, USA, 1997; pp. 107–116.

28. Petrosyan, L.A.; Mamkina, S.I. Games with changing coalitional partition . *Vestn. St. Petersburg Univ. Math.* **2004**, *3*, 60–69.

29. McKelvey, R.; McLennan, A.; aTurocy, T. Gambit: Software Tools for Game Theory. Version 16.0.1. 2016. Available online: http://www.gambit-project.org (accessed on 25 June 2020).

30. Savani, R.; von Stengel, B. Game theory explorer—software for the applied game theorist. *Comput. Manag. Sci.* **2015**, *12*, 5–33. [CrossRef]

31. Von Stengel, B. Computing equilibria for two-person games. In *Handbook of Game Theory*; Aumann, R., Hart, S., Eds.; North-Holland: Amsterdam, Netherlands, 2002; Volume 3, pp. 1723–1759.

32. Lemke, C. Bimatrix equilibrium points and mathematical programming. *Manag. Sci.* **1965**, *11*, 681–689. [CrossRef]

33. Shapley, L. A value for n-person games. In *Contributions to the Theory of Games, II*; Kuhn, H., Tucker, A.W., Eds.; Princeton University Press: Princeton, NJ, USA, 1953; pp. 307–317.

34. Schmeidler, D. The nucleolus of a characteristic function game. *SIAM J. Appl. Math.* **1969**, *17*, 1163–1170. [CrossRef]

35. Petrosian, O.; Zakharov, V. IDP-Core: Novel Cooperative Solution for Differential Games. *Mathematics* **2020**, *8*, 721. [CrossRef]

36. Yeung, D.; Petrosyan, L. Subgame consistent cooperative solutions in stochastic differential games. *J. Optim. Theory Appl.* **2004**, *120*, 651–666. [CrossRef]

37. Yeung, D.; Petrosyan, L. Subgame consistent solutions for cooperative stochastic dynamic games. *J. Optim. Theory Appl.* **2010**, *145*, 579–596. [CrossRef]

38. Yeung, D.; Petrosyan, L. Subgame-consistent cooperative solutions in randomly furcating stochastic dynamic games. *Math. Comput. Model.* **2013**, *57*, 976–991. [CrossRef]

39. Breton, M.; Dahmouni, I.; Zaccour, G. Equilibria in a two-species fishery. *Math. Biosci.* **2019**, *309*, 78–91. [CrossRef] [PubMed]

40. Crettez, B.; Hayek, N.; Zaccour, G. Do charities spend more on their social programs when they cooperate than when they compete? *Eur. J. Oper. Res.* **2020**, *283*, 1055–1063. [CrossRef]

41. Finus, M. *Game Theory and International Environmental Cooperation*; Edward, E., Ed.; Edward Elgar Publ.: Northampton, MA, USA, 2001.

42. Mazalov, V.V.; Rettiyeva, A.N. The discrete-time bioresource sharing model. *J. Appl. Math. Mech.* **2011**, *75*, 180–188. [CrossRef]

43. Ougolnitsky, G.; Usov, A. Spatially distributed differential game theoretic model of fisheries. *Mathematics*. **2019**, *7*, 732. [CrossRef]

44. Yeung, D.; Petrosyan, L. *Subgame Consistent Economic Optimization: An Advanced Cooperative Dynamic Game Analysis*; Springer: New York, NY, USA, 2012.

45. Nash, J.F. Equilibrium points in n-person games. *Proc. Nat. Acad. Sci. USA* **1950**, *36*, 48–49. [CrossRef]

46. Birch, B.J. On games with almost complete information. *Proc. Camb. Philos. Soc.* **1955**, *51*, 275–287. [CrossRef]

47. Dalkey, N. Equivalence of information patterns and essentially determinate games. *Contrib. Theory Games*. **1953**, *21*, 217–244.

48. Gromova, E.V.; Petrosyan, L.A. On an approach to constructing a characteristic function in cooperative differential games. *Autom. Remote. Control.* **2017**, *78*, 1680–1692. [CrossRef]