

**Approximated cooperative equilibria  
for games played over event trees**

E. Parilina  
G. Zaccour

G-2015-55

May 2015

---

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2015.

The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*.

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2015.



# Approximated cooperative equilibria for games played over event trees

Elena Parilina<sup>a</sup>

Georges Zaccour<sup>b</sup>

<sup>a</sup> *Saint Petersburg State University, Faculty of Applied Mathematics and Control Processes, Saint Petersburg 198504, Russia*

<sup>b</sup> *GERAD & Chair in Game Theory and Management, HEC Montréal, Montréal (Québec) Canada, H3T 2A7*

e.parilina@spbu.ru

georges.zaccour@gerad.ca

May 2015

Les Cahiers du GERAD

G–2015–55

Copyright © 2015 GERAD

**Abstract:** We consider the class of stochastic games played over finite event trees, that is, games where the random process is an act of nature and is not influenced by the players' actions. We suppose that the players agree to form the grand coalition and maximize their joint payoff. If the cooperative solution is not an equilibrium, then players may cheat on the agreement, unless a mechanism is designed to ensure that all players implement their cooperative controls over time (and nodes). To sustain cooperation over the event tree, we use behavior strategies known as grim trigger strategies. As we are dealing with a finite horizon, it is well known that deviation from cooperation in the last stage cannot be deterred, as there is no possibility for punishing the deviator(s). Consequently, we focus on epsilon (or approximated) equilibria. More specifically, we prove the existence of an epsilon-perfect equilibrium, where the value of epsilon is calculated using the game's parameters. We illustrate our findings with a numerical example.

**Key Words:** Stochastic games, S-adapted strategies, cooperative solution, perfect  $\varepsilon$ -equilibrium, trigger strategies.

**Résumé:** Nous considérons un jeu stochastique joué sur un arbre d'événements, et supposons que les joueurs sont d'accord pour former la grande coalition et maximiser leur profit joint. Si la solution coopérative n'est pas un équilibre, alors les joueurs peuvent tricher sur l'accord, à moins d'élaborer un mécanisme qui assure que les joueurs implémentent leurs commandes coopératives à travers le temps (et les nœuds de l'arbre de l'événement). Pour soutenir la coopération, nous utilisons les stratégies comportementales connues sous le nom de stratégies d'enclenchement. Comme le jeu est à horizon fini, il est bien connu qu'une déviation de la coopération à la dernière période ne peut pas être dissuadée, car il n'y a aucune possibilité pour punir le ou les joueur(s) qui dévient. Par conséquent, nous nous concentrons sur des équilibres approximés. Plus précisément, nous prouvons l'existence d'un équilibre epsilon-parfait, où la valeur d'epsilon est calculée en utilisant les paramètres du jeu. Nous illustrons nos résultats avec des exemples numériques.

**Mots clés:** Jeux stochastiques, stratégies S-adaptées, solution coopérative, équilibre epsilon-parfait, stratégies d'enclenchement.

---

**Acknowledgments:** This research was partially conducted during the research stay of the first author at GERAD. The work of the first author was partially supported by Saint Petersburg State University (research project 9.38.245.2014). The work of the second author is supported by SSHRC and NSERC Canada.

# 1 Introduction

An important issue in cooperative dynamic games is the sustainability of the agreement over time, that is, how to ensure that all players stick to their cooperative controls as time goes by. The literature in (state-space) dynamic games has dealt with this issue along essentially two lines, namely, the design of time-consistent mechanisms and cooperative equilibria. In a nutshell, the determination of a time-consistent solution involves a two-step procedure. The first step is the computation of the cooperative solution of the dynamic game and selection of an imputation, e.g., the Shapley value or an imputation from the core. Second, a payment schedule must be defined over time such that: (i) the (possibly discounted) total stream of payments to a player corresponds to her share in the overall cooperative game, that is, her imputation; (ii) at any intermediate instant of time, the cooperative payoff-to-go dominates its noncooperative counterpart. Observe that these payoffs-to-go are compared along the cooperative state trajectory, implying that the players have implemented their cooperative controls so far.<sup>1</sup> Further, a time-consistent solution is not an equilibrium, nor is it based on unilateral-deviation thinking, that is, either there is an agreement where all parties are on board, or there is no agreement at all. For a review of time consistency in differential games, see Yeung and Petrosjan (2005) and Zaccour (2008). In a cooperative equilibrium approach, as the name suggests, the idea is to make sure that the cooperative solution is an equilibrium, and, hence self-supported. This is achieved by letting the players use non-Markovian (or history-based) strategies that effectively deter any cheating on the cooperative agreement.

The objective of this paper is to design a cooperative equilibrium solution for stochastic dynamic games played over event trees, that is, games where the random process is an act of nature and is not influenced by the players' actions.<sup>2</sup> This class of games, which involves flow (control) and stock (state) variables, is useful to model competition and cooperation between players interacting repeatedly over time in the presence of an accumulation process. As an example, the set of players could be firms belonging to the same industry, where each firm makes an investment (control variable) to increase its production capacity (state variable), and with the price of the product being dependent on all firms' outputs and on some random event (weather, state of the economy, etc.). This class of games was initially introduced in Zaccour (1987) and Haurie et al. (1990) to study noncooperative equilibria in the European natural gas market, involving four suppliers competing over a long-term planning horizon in nine markets described by stochastic demand laws. The solution concept was termed  $S$ -adapted equilibrium, where the  $S$  stands for *sample* of realizations of the random process (see Haurie et al. (2012) for details, and Genc et al. (2007), Genc and Sen (2008) and Pineau et al. (2011) for applications of this class of games to energy markets). Recently, interest has shifted to cooperative games played over an event tree, with a focus on the sustainability of cooperation over time. Reddy et al. (2013) proposed a node-consistent decomposition of the Shapley value, and Parilina and Zaccour (2015) constructed a node-consistent core for these games. In this paper, our concern is not the allocation of the imputation over nodes, but the construction in a finite-horizon setting of an approximated cooperative equilibrium solution.

We use a grim trigger strategy, which is a behavior strategy based on the following simple rule: if cooperation has prevailed till now, then choose the cooperative control in the current stage; and if a deviation has been observed, then implement a noncooperative (or punishing) control for the rest of the game. The so-called folk theorem about the existence of a subgame perfect equilibrium in trigger strategies for infinitely repeated games was proved long ago (see, e.g., Aumann and Shapley (1994)). Dutta (1995) proved a similar theorem for stochastic games. Tolwinski et al. (1986) considered nonzero-sum differential games in strategies with memory. These strategies were called cooperative, as they were built as behavior strategies incorporating cooperative open-loop controls and feedback strategies used as threats in order to enforce the cooperative agreement. Recently, Chistyakov and Petrosyan (2013) examined the problem of strong strategic support for cooperative solutions in differential games, and Parilina (2014) stated some conditions for strategic support of cooperative solutions in stochastic games.

---

<sup>1</sup>Kaitala and Pohjola (1990) proposed the concept of an agreeable solution, which requires that cooperative payoff-to-go dominates noncooperative payoff-to-go along any state trajectory.

<sup>2</sup>The main difference between this class of games and classical stochastic games is that here players cannot influence the transition between one decision node (or state) and another.

Folk theorems are for infinite-horizon dynamic games. It is well known that enforcing cooperation in finite-horizon games is more difficult, not to say generally elusive. The reason is that, at the last stage, defection from the agreement is individually rational and this deviation cannot be punished. Using a backward-induction argument, it is easy to show that the unique subgame perfect equilibrium in repeated and multistage games is to implement Nash equilibrium controls at each stage of the finite game. This clear-cut theoretical result has not always received empirical support, and in fact, experiments show that cooperation may be realized, at least partially, in finite-horizon games (see, e.g. Angelova et al. (2013)).

The literature has come out with different ways to cope with the difficulties in enforcing cooperation in finite-horizon dynamic games. For instance, Eswaran and Lewis (1986) proposed to support collusive behavior in finite repeated games by having the players post bonds, which can be forfeited if they detect from cooperative behavior. Radner (1980) proposed the idea of  $\varepsilon$ -equilibrium and proved its existence for finitely repeated games.<sup>3</sup> A third alternative can be used in the class of finitely repeated games when there exist more than one Nash equilibria in a one-shot game (Benoit and Krishna (1985)).

Recently, the problem of the existence of a subgame-perfect  $\varepsilon$ -equilibrium in pure strategies has been investigated. Solan and Vieille (2003) considered a two-player game with perfect information that has no subgame-perfect  $\varepsilon$ -equilibrium in pure strategies, for a small  $\varepsilon$ , but has a subgame-perfect  $\varepsilon$ -equilibrium in behavior strategies for any  $\varepsilon > 0$ . Flesch et al. (2014) also examined the existence of a subgame-perfect  $\varepsilon$ -equilibrium in perfect information games with infinite horizon and Borel measurable payoffs. Flesch and Predtetchinski (2015) proposed the concept of  $\varphi$ -tolerance equilibrium perfect-information games of infinite duration where  $\varphi$  is a function of history. A strategy profile is said to be a  $\varphi$ -tolerance equilibrium, if, for any history  $h$ , this strategy profile is a  $\varphi(h)$ -equilibrium in the subgame starting at  $h$ . This concept is close to the one we investigate in this paper and to the contemporaneous perfect  $\varepsilon$ -equilibrium proposed in Mailath et al. (2005). Contrary to Solan and Vieille (2003), Flesch et al. (2014) and Flesch and Predtetchinski (2015), we examine finite horizon games with perfect information. Further, different authors retained different measures for payoffs. For instance, Radner (1980) and Benoit and Krishna (1985) assumed that the payoff is the average payoff for one stage of the game, whereas in Mailath et al. (2005) and here, the players' payoffs are given by a stream of discounted payments.

The rest of the paper is organized as follows: Section 2 recalls the main ingredients of the class of games played over an event tree. Section 3 states the problem of strategic support and the main results. We provide an illustrative example in Section 4, and briefly conclude in Section 5.

## 2 Game over event tree<sup>4</sup>

Let  $\mathcal{T} = \{0, 1, \dots, T\}$  be the set of periods. The stochastic process is represented by an event tree, which has a root node  $n^0$  in period 0 and a set of nodes  $\mathcal{N}^t = \{n_1^t, \dots, n_{N_t}^t\}$  in period  $t = 1, \dots, T$ . Denote by  $a(n_i^t) \in \mathcal{N}^{t-1}$  the unique predecessor of node  $n_i^t \in \mathcal{N}^t$  on the event-tree graph for  $t = 1, \dots, T$ , and by  $\mathcal{S}(n_i^t) \subset \mathcal{N}^{t+1}$  the set of all possible direct successors of node  $n_i^t \in \mathcal{N}^t$  for  $t = 0, \dots, T-1$ . A path from the root node  $n^0$  to a terminal node  $n_i^T$  is called a *scenario*. Each scenario has a probability and the probabilities of all scenarios sum up to 1. We denote by  $\pi(n_i^t)$  the probability of passing through node  $n_i^t$ , which corresponds to the sum of the probabilities of all scenarios that contain this node. In particular,  $\pi(n^0) = 1$  and  $\pi(n_i^T)$  is equal to the probability of the single scenario that terminates in (leaf) node  $n_i^T$ . Observe that each node  $n_i^t \in \mathcal{N}^t$  represents a possible sample value of the history of the stochastic process up to time  $t$ . The tree graph structure represents the nesting of information as one time period succeeds the other.

Denote by  $M = \{1, \dots, m\}$  the set of players. For each player  $j \in M$ , we define a set of decision variables indexed over the set of nodes. Denote by  $u_j(n_i^t) \in \mathbb{R}^{m_j}$  the decision variables of player  $j$  at node  $n_i^t$ , and let  $u(n_i^t) = (u_1(n_i^t), \dots, u_m(n_i^t))$ . Let  $X \subset \mathbb{R}^p$ , with  $p \in \mathbb{N}_+$ , be a state set. For each node  $n_i^t \in \mathcal{N}^t$ ,  $t = 0, 1, \dots, T$ , let  $U_j^{n_i^t} \subset \mathbb{R}^{\mu_j^{n_i^t}}$ , with  $\mu_j^{n_i^t} \in \mathbb{N}_+$ , be the control set of player  $j$ . Denote by

<sup>3</sup>Nash-equilibrium,  $\varepsilon$ -equilibrium and subgame perfect  $\varepsilon$ -equilibrium for repeated games are described in Kalai (1987).

<sup>4</sup>This section draws heavily on Haurie et al. (2012) and Parilina and Zaccour (2015).

$U^{n_i^t} = U_1^{n_i^t} \times \cdots \times U_j^{n_i^t} \times \cdots \times U_m^{n_i^t}$  the product control sets. A transition function  $f^{n_i^t}(\cdot, \cdot) : X \times U^{n_i^t} \mapsto X$  is associated with each node  $n_i^t$ . The state equations are given by

$$x(n_i^t) = f^{a(n_i^t)}(x(a(n_i^t)), u(a(n_i^t))), \quad (1)$$

$$u(a(n_i^t)) \in U^{a(n_i^t)}, \quad n_i^t \in \mathcal{N}^t, t = 1, \dots, T. \quad (2)$$

At each node  $n_i^t$ ,  $t = 0, \dots, T-1$ , the reward to player  $j$  is a function of the state and the controls of all players, given by  $\phi_j^{n_i^t}(x(n_i^t), u(n_i^t))$ . At a terminal node  $n_i^T$ , the reward to player  $j$  is given by the function  $\Phi_j^{n_i^T}(x(n_i^T))$ .

We assume that player  $j \in M$  maximizes her expected stream of payoffs discounted at rate  $\lambda_j$  ( $0 < \lambda_j < 1$ ). The state equations and the reward functions define the following multistage game, where we let

$$\mathbf{x} = \{x(n_i^t) : n_i^t \in \mathcal{N}^t, t = 0, \dots, T\},$$

$$\mathbf{u} = \{u(n_i^t) : n_i^t \in \mathcal{N}^t, t = 0, \dots, T-1\},$$

and  $J_j(\mathbf{x}, \mathbf{u})$  be the payoff to player  $j$ , that is,

$$J_j(\mathbf{x}, \mathbf{u}) = \sum_{t=0}^{T-1} \lambda_j^t \sum_{n_i^t \in \mathcal{N}^t} \pi(n_i^t) \phi_j^{n_i^t}(x(n_i^t), u(n_i^t)) + \lambda_j^T \sum_{n_i^T \in \mathcal{N}^T} \pi(n_i^T) \Phi_j^{n_i^T}(x(n_i^T)), \quad j \in M, \quad (3)$$

s.t.

$$x(n_i^t) = f^{a(n_i^t)}(x(a(n_i^t)), u(a(n_i^t))), \quad (4)$$

$$u(a(n_i^t)) \in U^{a(n_i^t)}, \quad n_i^t \in \mathcal{N}^t, t = 1, \dots, T,$$

$$x(n^0) = x^0 \text{ given.} \quad (5)$$

**Definition 1** An admissible  $S$ -adapted strategy of player  $j$  is a vector  $\mathbf{u}_j = \{u_j(n_i^t) : n_i^t \in \mathcal{N}^t, t = 0, \dots, T-1\}$ , that is, a plan of actions adapted to the history of the random process represented by the event tree.

The  $S$ -adapted strategy vector of the  $m$  players is  $\mathbf{u} = (\mathbf{u}_j : j \in M)$ . We can thus define a game in normal form, with payoffs  $W_j(\mathbf{u}, x^0) = J_j(\mathbf{x}, \mathbf{u})$ ,  $j \in M$ , where  $\mathbf{x}$  is obtained from  $\mathbf{u}$  as the unique solution of the state equations that emanate from the initial state  $x^0$ .

We point out that there is a subtle, but important, difference between open-loop and  $S$ -adapted information structures (and equilibria). Indeed, whereas in an open-loop information structure, the controls and the state equations are defined over time, they are defined (indexed) over the set of nodes of the event tree in an  $S$ -adapted information structure.

If the players agree to cooperate, then they will maximize the sum of their discounted payoffs throughout the entire horizon, that is,

$$\max_{\mathbf{u}_j : j \in M} \sum_{j \in M} W_j(\mathbf{u}, x^0).$$

Denote the resulting vector of cooperative controls by  $\mathbf{u}^*$ , which is found for the game starting from node  $n^0$  and state  $x^0$ :

$$\mathbf{u}^* = \arg \max_{\mathbf{u}_j : j \in M} \sum_{j \in M} W_j(\mathbf{u}, x^0). \quad (6)$$

Further, denote by  $\mathbf{x}^* = \{x^*(n_i^t) : n_i^t \in \mathcal{N}^t, t = 0, 1, \dots, T\}$  the cooperative state trajectory generated by the cooperative controls  $\mathbf{u}^*$ .

For later use, we also need to determine the subgame starting from state  $x^*(n_i^t)$  at node  $n_i^t \in \mathcal{N}^t$ ,  $t = 1, \dots, T-1$ . This subgame takes place on a tree subgraph  $\Gamma(n_i^t)$  of the initial graph. The payoff of player  $j \in M$  in this subgame is given as follows:

$$W_j(\mathbf{u}(n_i^t), x^*(n_i^t)) = \sum_{\theta=t}^{T-1} \lambda_j^{\theta-t} \sum_{n_i^\theta \in \mathcal{N}^\theta} \pi(n_i^\theta | n_i^t) \phi_j^{n_i^\theta}(x^*(n_i^\theta), u(n_i^\theta)) + \lambda_j^{T-t} \sum_{n_i^T \in \mathcal{N}^T} \pi(n_i^T | n_i^t) \Phi_j^{n_i^T}(x^*(n_i^T)), \quad (7)$$

where  $\mathcal{N}_\Gamma^\theta = \mathcal{N}^\theta \cap \Gamma(n_i^t)$ ,  $\mathbf{u}(n_i^t) = (\mathbf{u}_j(n_i^t) : j \in M)$  is an  $S$ -adapted strategy profile and  $\mathbf{u}_j(n_i^t) = \{u_j(n_i^\theta) : n_i^\theta \in \Gamma(n_i^t)\}$  is an admissible  $S$ -adapted strategy of player  $j$  in the subgame starting from node  $n_i^t$ , with initial state  $x^*(n_i^t)$ . The term  $\pi(n_i^\theta | n_i^t)$  is the conditional probability<sup>5</sup> that node  $n_i^\theta$  will be realized if the subgame starts from node  $n_i^t$ .

If the players cooperate in the subgame starting from state  $x^*(n_i^t)$ , then they maximize the sum of their total discounted payoffs, i.e.,

$$\max_{\mathbf{u}_j(n_i^t): j \in M} \sum_{j \in M} W_j(\mathbf{u}(n_i^t), x^*(n_i^t)),$$

and the cooperative controls in the subgame are given as follows:

$$\mathbf{u}^*(n_i^t) = \arg \max_{\mathbf{u}_j(n_i^t): j \in M} \sum_{j \in M} W_j(\mathbf{u}(n_i^t), x^*(n_i^t)). \quad (8)$$

Therefore, the payoff of player  $j$  in the cooperative subgame starting from node  $n_i^t$ , with initial state  $x^*(n_i^t)$ ,  $n_i^t \in \mathcal{N}^t$ , is equal to  $W_j(\mathbf{u}^*(n_i^t), x^*(n_i^t))$ ,  $t = 1, \dots, T$ . Denote the trajectory of  $\mathbf{u}^*(n_i^t)$  on the path emanating from node  $n_v^\tau \in \mathcal{N}^\tau$ ,  $\tau > t$ , and terminating at node  $n_w^T \in \mathcal{N}^T$  as  $\mathbf{u}^*([n_v^\tau, n_w^T] / n_i^t)$ .

### 3 Approximated cooperative equilibrium

Now, we consider the game played over an event tree as a game in extensive form with closed-loop information structure. This means that each player knows not only the current node  $n_i^t \in \mathcal{N}^t$ ,  $t = 0, \dots, T$  and what she played on the (unique) path leading from the initial node  $n^0$  to  $a(n_i^t)$ , but also what the other players did in all previous periods. Let this path be  $(n^0, n_{i_1}^1, \dots, n_{i_{t-1}}^{t-1}, n_i^t)$  and denote it by  $P(n_i^t)$ . The collection of nodes and corresponding strategy profiles realized on path  $P(n_i^t)$ , except node  $n_i^t$ , is called the *History of Node*  $n_i^t$  and is denoted by  $H(n_i^t) = ((n^0, u(n^0)), (n_{i_1}^1, u(n_{i_1}^1)), \dots, (n_{i_{t-1}}^{t-1}, u(n_{i_{t-1}}^{t-1})))$ .

**Definition 2** A behavior strategy  $\sigma_j = \{\sigma_j(n_i^t) : n_i^t \in \mathcal{N}^t, t = 0, \dots, T-1\}$  of player  $j \in M$  in the game played over an event tree is a mapping that associates to each node  $n_i^t$  an action  $u_j(n_i^t) \in U_j^{n_i^t}$  with each history  $H(n_i^t)$ , that is,

$$\sigma_j(n_i^t) : H(n_i^t) \longrightarrow U_j^{n_i^t}.$$

In other words, a behavior strategy tells a player what to do at each node  $n_i^t$  of the event tree  $\Gamma(n^0)$ . Denote by  $\Sigma_j$  the set of behavior strategies of player  $j$ , by  $\sigma = (\sigma_1, \dots, \sigma_m)$  a behavior strategy profile, and by  $\Sigma = \Sigma_1 \times \dots \times \Sigma_m$  the set of possible strategy profiles. For a given behavior strategy profile, we can compute the expected payoff in all subgames, including the whole game, for any given initial state. To avoid adding new notations, we denote the payoff of player  $j$  in the subgame starting at node  $n_i^t$  and in state  $x(n_i^t)$  as a function of the behavior strategy profile, by

$$W_j(\sigma, x(n_i^t)) = W_j(\mathbf{u}(n_i^t), x(n_i^t)),$$

where  $\mathbf{u}(n_i^t)$  is a trajectory of controls in the subgame starting at node  $n_i^t$  and determined by profile  $\sigma$ .

**Definition 3** A behavior strategy profile  $\hat{\sigma}$  is an  $\varepsilon$ -equilibrium if, for each player  $j \in M$  and each strategy  $\sigma_j \in \Sigma_j$ , the following inequality holds:

$$W_j(\hat{\sigma}, x^0) \geq W_j((\hat{\sigma}_{-j}, \sigma_j), x^0) - \varepsilon. \quad (9)$$

**Definition 4** A behavior strategy profile  $\hat{\sigma}$  is a perfect  $\varepsilon$ -equilibrium if, for each player  $j \in M$ , each node  $n_i^t \in \mathcal{N}^t$ , each strategy  $\sigma_j \in \Sigma_j$ , and if, for each history  $H(n_i^t)$ , the following inequality holds:

$$W_j(\hat{\sigma}, x(n_i^t) | H(n_i^t)) \geq W_j((\hat{\sigma}_{-j}, \sigma_j), x(n_i^t) | H(n_i^t)) - \varepsilon,$$

where  $W_j(\hat{\sigma}, x(n_i^t) | H(n_i^t))$  is player  $j$ 's payoff in the subgame starting at node  $n_i^t$  in state  $x(n_i^t)$  given by history  $H(n_i^t)$  when the players use strategy profile  $\hat{\sigma}$ .

<sup>5</sup>The conditional probability  $\pi(n_i^\theta | n_i^t)$  can be calculated with the formula:  $\pi(n_i^\theta | n_i^t) = \pi(n_i^\theta) / \pi(n_i^t)$  if  $\pi(n_i^t) \neq 0$ ; otherwise, the subgame starting from node  $n_i^t$  cannot materialize.



To strategically support cooperation in the finite-horizon dynamic game played over an event tree, we shall construct an approximated equilibrium in behavior strategies with a closed-loop information structure. The development is in line with what has been done in repeated games, modulo the fact that here we additionally have a vector of state variables evolving over time. Before formally defining the grim trigger strategy that will be used in this construction, let us suppose (i) that the players want to realize the cooperative trajectory  $\mathbf{u}^*$  from (6), and (ii) that if a player  $j$  deviates from cooperation at node  $a$  ( $n_i^t$ ), then the other players will minmax her payoff in the subgame starting at  $n_i^t$  in state  $x(n_i^t)$ . This minmax value is defined by

$$\bar{W}_j(x(n_i^t)) = \min_{\sigma_{-j} \in \prod_{p \neq j} \Sigma_p} \max_{\sigma_j \in \Sigma_j} W_j((\sigma_{-j}, \sigma_j), x(n_i^t)),$$

that is, the gain that players in  $M \setminus \{j\}$  cannot prevent player  $j$  from achieving in the subgame starting in node  $n_i^t$  and state  $x(n_i^t)$ . Any vector  $\sigma_{-j}$  satisfying

$$\hat{\sigma}_{-j} = \arg \min_{\sigma_{-j} \in \prod_{p \neq j} \Sigma_p} \max_{\sigma_j \in \Sigma_j} W_j((\sigma_{-j}, \sigma_j), x(n_i^t)), \quad (10)$$

is a punishing minmax strategy against player  $j$  deviating, and is part of the grim trigger strategy defined below. Denote by  $\hat{\mathbf{u}}^j(n_i^t) = \{\hat{u}^j(n_i^t) : n_i^t \in \Gamma(n_i^t)\}$  the controls corresponding to minmax strategy profile  $\hat{\sigma} = (\hat{\sigma}_{-j}, \hat{\sigma}_j)$ , which punishes player  $j$  in the subgame starting in node  $n_i^t$  and state  $x(n_i^t)$ .

The grim trigger behavior strategy of a player consists of two behavior types or two modes:

**The nominal mode.** If the history of node  $n_i^t$  coincides with

$$H^*(n_i^t) = ((n^0, u^*(n^0)), (n_{i_1}^1, u^*(n_{i_1}^1)), \dots, (n_{i_{t-1}}^{t-1}, u^*(n_{i_{t-1}}^{t-1}))), \quad (11)$$

i.e., all players used their cooperative controls on the path  $P(n_{i_{t-1}}^{t-1})$ , that is, from  $n^0$  until  $n_{i_{t-1}}^{t-1}$ , then player  $j, j \in M$  implements  $u_j^*(n_i^t)$  in node  $n_i^t$ .

**The trigger mode.** If the history of node  $n_i^t$  is such that there exists a node  $n$  on the path  $P(n_{i_{t-1}}^{t-1})$  and a deviating player  $j \in M, j \neq p$ , that is, the history  $H(n)$  of node  $n$  is part of  $H^*(n_i^t)$ , and  $(n, u(n))$  is not part of  $H^*(n_i^t)$ , but if we replace the control  $u_j(n)$  of player  $j$  in node  $n$  by the cooperative control  $u_j^*(n)$ , then the pair  $(n, (u_{-j}(n), u_j^*(n)))$  will be  $(n, u^*(n))$  and part of history  $H^*(n_i^t)$ . Then, player  $p$ 's strategy is part of punishing strategy  $\sigma_{-j}$  of  $M \setminus \{j\}$  from (10), with punishment starting from the successor nodes of the node at which player  $j$  defected from cooperation.

Formally speaking, the trigger behavior strategy of player  $p \in M$  is defined as follows:

$$\hat{\sigma}_p(H(n_i^t)) = \begin{cases} u_p^*(n_i^t), & \text{if } H(n_i^t) = H^*(n_i^t), \\ \hat{u}_p^j(n_i^t), & \text{if there exists a node } n \text{ on path } P(n_i^t) \\ & \text{and player } j \in M, j \neq p \text{ such that } H(n) \subset H^*(n_i^t), \\ & \text{and } (n, u(n)) \notin H^*(n_i^t), \text{ but } (n, (u_{-j}(n), u_j^*(n))) \in H^*(n_i^t), \end{cases} \quad (12)$$

where  $\hat{u}_p^j(n_i^t)$  is player  $p$ 's control in node  $n_i^t$ . The control  $\hat{u}_p^j(n_i^t)$  implements the punishing strategy given in (10) against player  $j$  (hence the superscript  $j$  in  $\hat{u}_p^j(n_i^t)$ ) in the subgame starting in the unique node that belongs to the set  $\mathcal{S}(n) \cap P(n_i^t)$ .

To avoid further complicating the notation, we omitted the state argument in the punishing control and the trigger strategy, but we stress that they depend on the state value. Let node  $n_1$  be a direct successor of node  $n$  in which player  $j$  deviates. The collection of controls  $(u_{-j}^*(n), u_j(n))$  is then realized, and the state value in node  $n_1$  can be calculated using the state dynamics  $x(n_1) = f^n(x^*(n), (u_{-j}^*(n), u_j(n)))$ . The control  $\hat{u}_p^j(n_1)$  is part of the control profile

$$\hat{\mathbf{u}}^j(n_1) = \arg \min_{\substack{\mathbf{u}_{-j}(n_1) = \\ (\mathbf{u}_p(n_1): p \in M \setminus \{j\})}} \max_{\mathbf{u}_j(n_1)} W_j((\mathbf{u}_{-j}(n_1), \mathbf{u}_j(n_1)), x(n_1)).$$

Now, in the subgame starting from node  $n_l^t \in \mathcal{N}^t$  in state  $x(n_l^t)$ , the collection of punishing controls corresponding to players in  $M \setminus \{j\}$  minmaxing strategies while player  $j$  maximizing her payoff is given by

$$\hat{\mathbf{u}}^j(n_l^t) = (\hat{\mathbf{u}}_p^j(n_l^t) : p \in M),$$

where  $\hat{\mathbf{u}}_p^j(n_l^t) = \{\hat{u}_p^j(n_l^\theta) : n_l^\theta \in \Gamma(n_l^t)\}$ . This collection of controls generates a minmax trajectory of states in player  $j$ 's punishment in this subgame, that is,

$$\hat{\mathbf{x}}^j(n_l^t) = \{\hat{x}^j(n_l^\theta) : n_l^\theta \in \Gamma(n_l^t)\}.$$

To construct the trigger strategies, we need to find  $m$  punishing strategy profiles for each subgame. Our main result follows.

**Theorem 1** *In the game played over an event tree there exists a perfect  $\varepsilon$ -equilibrium in trigger strategies with players' payoffs  $(W_1(\mathbf{u}^*, x^0), \dots, W_m(\mathbf{u}^*, x^0))$  where*

$$\varepsilon = \max_{j \in M} \max_{n_l^t \in \mathcal{N}^t} \max_{t=1, \dots, T-1} \varepsilon_j(n_l^t), \quad (13)$$

where

$$\begin{aligned} \varepsilon_j(n_l^t) = & \max_{u_j(n_l^t) \in U_j^{n_l^t}} \left\{ \phi_j^{n_l^t}(x^*(n_l^t), (u_{-j}^*(n_l^t), u_j(n_l^t))) - \phi_j^{n_l^t}(x^*(n_l^t), u^*(n_l^t)) \right. \\ & + \sum_{\theta=t+1}^{T-1} \lambda_j^{\theta-t} \sum_{n_l^\theta \in \mathcal{N}_\Gamma^\theta} \pi(n_l^\theta | n_l^t) \left( \phi_j^{n_l^\theta}(\hat{x}^j(n_l^\theta), \hat{u}^j(n_l^\theta)) - \phi_j^{n_l^\theta}(x^*(n_l^\theta), u^*(n_l^\theta)) \right) \\ & \left. + \lambda_j^{T-t} \sum_{n_l^T \in \mathcal{N}_\Gamma^T} \pi(n_l^T | n_l^t) \left( \Phi_j^{n_l^T}(\hat{x}(n_l^T)) - \Phi_j^{n_l^T}(x^*(n_l^T)) \right) \right\}, \quad (14) \end{aligned}$$

and  $\hat{u}^j(n_l^\theta)$  is a control profile in node  $n_l^\theta$  corresponding to a minmax strategy profile  $\hat{\sigma}^j$  punishing player  $j$  in the subgame, starting at the node belonging to the set  $\mathcal{S}(n_l^t)$  and in state  $f^{n_l^t}(x(n_l^t), (u_{-j}^*(n_l^t), u_j(n_l^t)))$ . Therefore, the differences in the second and third lines also depend on the control  $u_j(n_l^t)$ . The state  $\hat{x}^j(n_l^\theta)$ ,  $n_l^\theta \in \Gamma(n_l^t)$  is a state trajectory corresponding to  $\hat{\mathbf{u}}^j$ .

**Proof.** Consider the trigger behavior strategy  $\hat{\sigma} = (\hat{\sigma}_p : p \in M)$  defined in (12), and the subgame starting from any node  $n_l^t \in \mathcal{N}^t$ ,  $t = 0, \dots, T-1$ . If player  $j$  does not deviate from the cooperative trajectory in node  $n_l^t$ , then her payoff in this subgame will be given by

$$\begin{aligned} W_j(\mathbf{u}^*(n_l^t), x^*(n_l^t)) = & \phi_j^{n_l^t}(x^*(n_l^t), u^*(n_l^t)) + \sum_{\theta=t+1}^{T-1} \lambda_j^{\theta-t} \sum_{n_l^\theta \in \mathcal{N}_\Gamma^\theta} \pi(n_l^\theta | n_l^t) \phi_j^{n_l^\theta}(x^*(n_l^\theta), u^*(n_l^\theta)) \\ & + \lambda_j^{T-t} \sum_{n_l^T \in \mathcal{N}_\Gamma^T} \pi(n_l^T | n_l^t) \Phi_j^{n_l^T}(x^*(n_l^T)), \quad (15) \end{aligned}$$

where  $\mathcal{N}_\Gamma^\theta = \mathcal{N}^\theta \cap \Gamma(n_l^t)$ ,  $\mathbf{u}^*(n_l^t) = (\mathbf{u}_j^*(n_l^t) : j \in M)$  is an  $S$ -adapted cooperative strategy profile.

Suppose player  $j$  deviates in node  $n_l^t$  from the cooperative trajectory. In this case, she may secure the following payoff in the subgame starting at node  $n_l^t$ , given the information that the behavior strategy profile  $\hat{\sigma} = (\hat{\sigma}_p(\cdot) : p \in M)$  determined by (12) will materialize:

$$\begin{aligned} & \max_{u_j(n_l^t) \in U_j^{n_l^t}} \left\{ \phi_j^{n_l^t}(x^*(n_l^t), (u_{-j}^*(n_l^t), u_j(n_l^t))) \right. \\ & \left. + \sum_{\theta=t+1}^{T-1} \lambda_j^{\theta-t} \sum_{n_l^\theta \in \mathcal{N}_\Gamma^\theta} \pi(n_l^\theta | n_l^t) \phi_j^{n_l^\theta}(\hat{x}^j(n_l^\theta), \hat{u}^j(n_l^\theta)) + \lambda_j^{T-t} \sum_{n_l^T \in \mathcal{N}_\Gamma^T} \pi(n_l^T | n_l^t) \Phi_j^{n_l^T}(\hat{x}^j(n_l^T)) \right\}, \quad (16) \end{aligned}$$

where punishing minmax strategy starts to be implemented in nodes from  $\mathcal{S}(n_i^t)$ . Then, we may compute the benefit from deviation of player  $j$  in node  $n_i^t$  as a difference between (16) and (15), namely:

$$\begin{aligned} \varepsilon_j(n_i^t) = & \max_{u_j(n_i^t) \in U_j^{n_i^t}} \left\{ \phi_j^{n_i^t}(x^*(n_i^t), (u_{-j}^*(n_i^t), u_j(n_i^t))) - \phi_j^{n_i^t}(x^*(n_i^t), u^*(n_i^t)) \right. \\ & + \sum_{\theta=t+1}^{T-1} \lambda_j^{\theta-t} \sum_{n_i^\theta \in \mathcal{N}_i^\theta} \pi(n_i^\theta | n_i^t) \left( \phi_j^{n_i^\theta}(\hat{x}^j(n_i^\theta), \hat{u}^j(n_i^\theta)) - \phi_j^{n_i^\theta}(x^*(n_i^\theta), u^*(n_i^\theta)) \right) \\ & \left. + \lambda_j^{T-t} \sum_{n_i^T \in \mathcal{N}_i^T} \pi(n_i^T | n_i^t) \left( \Phi_j^{n_i^T}(\hat{x}(n_i^T)) - \Phi_j^{n_i^T}(x^*(n_i^T)) \right) \right\}, \end{aligned}$$

Calculating the maximum benefit from deviation for any subgame and any player, we obtain the value of  $\varepsilon$  in the theorem statement, that is,

$$\varepsilon = \max_{j \in M} \max_{\substack{n_i^t \in \mathcal{N}^t \\ t=1, \dots, T-1}} \varepsilon_j(n_i^t).$$

And for  $\varepsilon$  equal to this value, the behavior strategy profile determined by (12) is a perfect  $\varepsilon$ -equilibrium by construction.  $\square$

## 4 Numerical illustrations

To illustrate the results of the previous section, we consider a three-player stochastic version of the deterministic model of pollution control in Germain et al. (2003). Denote by  $M = \{1, 2, 3\}$  the set of players, and by  $\mathcal{T} = \{0, 1, 2, 3\}$  the set of periods. Let  $u(n_i^t) = (u_1(n_i^t), u_2(n_i^t), u_3(n_i^t))$  be the vector of countries' emissions of some pollutant and denote by  $x(n_i^t)$  the stock of pollution at node  $n_i^t$  in time period  $t$ . The evolution of this stock is governed by the following difference equation:

$$x(n_i^t) = (1 - \delta(n_i^t))x(n_i^{t-1}) + \sum_{j \in M} u_j(n_i^t), \quad (17)$$

$$x(n^0) = x^0, \quad (18)$$

with the initial stock  $x^0$  at root node  $n^0$  being given, and  $\delta(n_i^t)$  ( $0 < \delta(n_i^t) < 1$ ) is the stochastic rate of pollution absorption by nature at node  $n_i^t$ . We suppose that  $\delta(n_i^t)$  can take two possible values, that is,  $\delta(n_i^t) \in \{\underline{\delta}, \bar{\delta}\}$ , with  $\underline{\delta} < \bar{\delta}$ . The event tree is depicted in Figure 1. Let nodes  $n_2^1, n_2^2, n_4^2$  correspond to the low level of pollution reduction  $\underline{\delta}$ , and nodes  $n^0, n_1^1, n_1^2, n_3^2$  correspond to the high level of pollution reduction  $\bar{\delta}$ .

The damage cost is an increasing convex function in the pollution stock having the quadratic form  $D_j(x(n_i^t)) = \alpha_j x^2(n_i^t)$ ,  $j \in M$ , where  $\alpha_j$  is a strictly positive parameter. The cost of emissions is also given by a quadratic function  $C_j(u_j(n_i^t)) = \frac{\gamma_j}{2} (u_j(n_i^t) - e)^2$ , where  $e$  and  $\gamma_j$  are strictly positive constants.

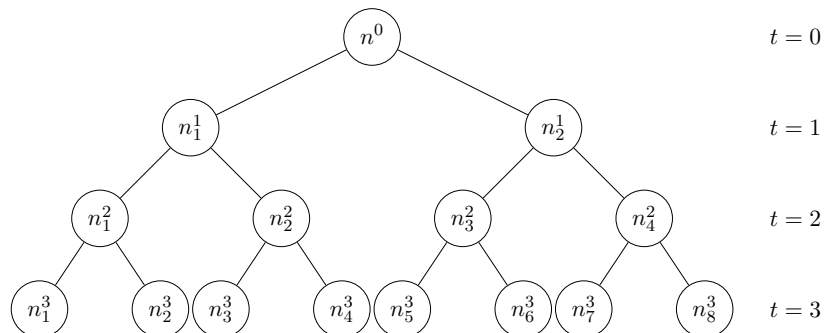


Figure 1: Event tree graph for  $T = 3$ .

The total discounted cost  $J_j(\mathbf{x}, \mathbf{u})$  to be minimized by player  $j \in M$  is given by

$$\sum_{t=0}^2 \lambda_j^t \sum_{n_i^t \in \mathcal{N}^t} \pi(n_i^t) (C_j(u_j(n_i^t)) + D_j(x(n_i^t))) + \lambda_j^3 \sum_{n_i^3 \in \mathcal{N}^3} \pi(n_i^3) D_j(x(n_i^3)),$$

where  $\mathbf{x} = \{x(n_i^t)\}$  and  $\mathbf{u} = \{u(n_i^t)\}$ ,  $n_i^t \in \mathcal{N}^t$ ,  $\lambda_j \in (0, 1)$  is a discount rate of player  $j$ , subject to (17), given initial stock  $x_0 = 0$  before the game starts and constraints:  $u_j(n_i^t) \in [0, e]$  for any player  $j \in M$  and any node  $n_i^t \in \mathcal{N}^t$ ,  $t = 0, 1, 2$ .

We use the following parameters for the numerical simulation:

$$\begin{aligned} \alpha_1 &= 0.1, \alpha_2 = 0.2, \alpha_3 = 0.3, \\ \gamma_1 &= 0.9, \gamma_2 = 0.8, \gamma_3 = 0.7, \\ \underline{\delta} &= 0.45, \bar{\delta} = 0.8, e = 30, \lambda_1 = \lambda_2 = \lambda_3 = 0.9, \\ \pi(n_1^1) &= 0.6, \pi(n_2^1) = 0.4, \\ \pi(n_1^2) &= 0.3, \pi(n_2^2) = 0.3, \pi(n_3^2) = 0.3, \pi(n_4^2) = 0.1, \\ \pi(n_1^3) &= 0.1, \pi(n_2^3) = 0.2, \pi(n_3^3) = 0.1, \pi(n_4^3) = 0.2, \\ \pi(n_5^3) &= 0.05, \pi(n_6^3) = 0.25, \pi(n_7^3) = 0.05, \pi(n_8^3) = 0.05. \end{aligned}$$

Using (6) and (8), we compute the cooperative controls for each possible subgame and for the whole game. The cooperative state trajectory is given by

$$\begin{array}{cccccccc} x^*(n^0) & x^*(n_1^1) & x^*(n_2^1) & x^*(n_1^2) & x^*(n_2^2) & x^*(n_3^2) & x^*(n_4^2) & x^*(n_1^3) \\ 0 & 55.177 & 55.177 & 63.4188 & 63.4188 & 79.0484 & 79.0484 & 72.8623 \\ x^*(n_2^3) & x^*(n_3^3) & x^*(n_4^3) & x^*(n_5^3) & x^*(n_6^3) & x^*(n_7^3) & x^*(n_8^3) & \\ 72.8623 & 88.6125 & 88.6125 & 75.0804 & 75.0804 & 94.7123 & 94.7123 & \end{array}$$

We use (7) to compute the players' costs in each of the subgames and report them in Table 1.

Table 1: Players' costs in the cooperative control profile.

Time period	$t = 0$			$t = 1$			$t = 2$	
Node	$n^0$	$n_1^1$	$n_2^1$	$n_1^2$	$n_2^2$	$n_3^2$	$n_4^2$	
$W_1(\mathbf{u}(\cdot), x^*(\cdot))$	249.371	213.06	242.807	122.401	161.771	149.748	201.349	
$W_2(\mathbf{u}(\cdot), x^*(\cdot))$	381.031	344.644	394.864	214.701	279.02	267.534	351.835	
$W_3(\mathbf{u}(\cdot), x^*(\cdot))$	517.495	479.554	550.625	308.23	398.087	386.625	504.397	
	$t = 3$							
	$n_1^3$	$n_2^3$	$n_3^3$	$n_4^3$	$n_5^3$	$n_6^3$	$n_7^3$	$n_8^3$
	53.0891	53.0891	78.5218	78.5218	56.3706	56.3706	89.7041	89.7041
	106.178	106.178	157.044	157.044	112.741	112.741	179.408	179.408
	159.267	159.267	235.565	235.565	169.112	169.112	269.112	269.112

For each player  $j \in M$  and any node  $n_i^t \in \mathcal{N}^t$ ,  $t = 0, \dots, T-1$  we need to solve the optimization problem defined in (16), with min instead of max. Once we obtain these costs, we compute the differences with the cooperative payoffs and give them in Table 2. Based on these differences, we determine the values  $\varepsilon_j(n_i^t)$  for  $j \in M$  and  $n_i^t \in \mathcal{N}^t$ ,  $t = 0, 1, 2$  (see Table 3).

From Table 3 we see that at root node  $n^0$ , no player benefits from deviating from cooperation. At time  $t = 1$ , only player 1 can gain by deviating, whereas at terminal period 2, all players gain by deviating, which is expected. The largest benefits from deviation are realized simultaneously for all players in node  $n_4^2$ . Finally, we note that, in this example,  $\varepsilon$  is equal to 39.5753.

Table 2: Maximum benefits from deviation for any subgame.

Time period	$t = 0$	$t = 1$		$t = 2$			
Node	$n^0$	$n_1^1$	$n_2^1$	$n_1^2$	$n_2^2$	$n_3^2$	$n_4^2$
Player 1	-5.16799	23.3117	31.4627	23.4217	34.6419	24.8694	39.5753
Player 2	-113.551	-29.5407	-22.8292	16.4602	24.3455	17.4776	27.8126
Player 3	-198.86	-72.3691	-67.1232	10.2658	15.1836	10.9003	17.346

Table 3: Values of  $\varepsilon_j(\cdot)$  for any subgame and any player.

Time period	$t = 0$	$t = 1$		$t = 2$			
Node	$n^0$	$n_1^1$	$n_2^1$	$n_1^2$	$n_2^2$	$n_3^2$	$n_4^2$
$\varepsilon_1$	0	23.3117	31.4627	23.4217	34.6419	24.8694	39.5753
$\varepsilon_2$	0	0	0	16.4602	24.3455	17.4776	27.8126
$\varepsilon_3$	0	0	0	10.2658	15.1836	10.9003	17.346

The game at hand belongs to the class of environmental games with negative externalities. Given the general result that cooperation is hard to achieve in such setting (see, e.g., the survey in Jørgensen et al. (2010)), the results in the above short-time horizon example come at no surprise. To get a hint of the impact of having a longer time horizon, we now let the set of periods be  $\mathcal{T} = \{0, 1, \dots, 10\}$ , with everything else being equal.

The event tree is depicted in Figure 2. It is a binary tree, i.e., each node in periods  $t = 0, \dots, 9$  has two successors. The conditional probability of realization of the upward successor of any node is  $\frac{1}{3}$  and is  $\frac{2}{3}$  for a downward successor.. So, for instance, we have probabilities  $\pi(n_1^1) = \frac{1}{3}$  and  $\pi(n_2^1) = \frac{2}{3}$  in period 1, and probabilities  $\pi(n_1^2) = \frac{1}{9}$ ,  $\pi(n_2^2) = \frac{2}{9}$ ,  $\pi(n_3^2) = \frac{2}{9}$ ,  $\pi(n_4^2) = \frac{4}{9}$  for  $t = 2$ . The root node  $n^0$  and all upward (or left-handed) nodes have the low rate  $\underline{\delta}$  of pollution absorption by nature, and all downward (or right-handed) nodes have the high level  $\bar{\delta}$  of pollution absorption.

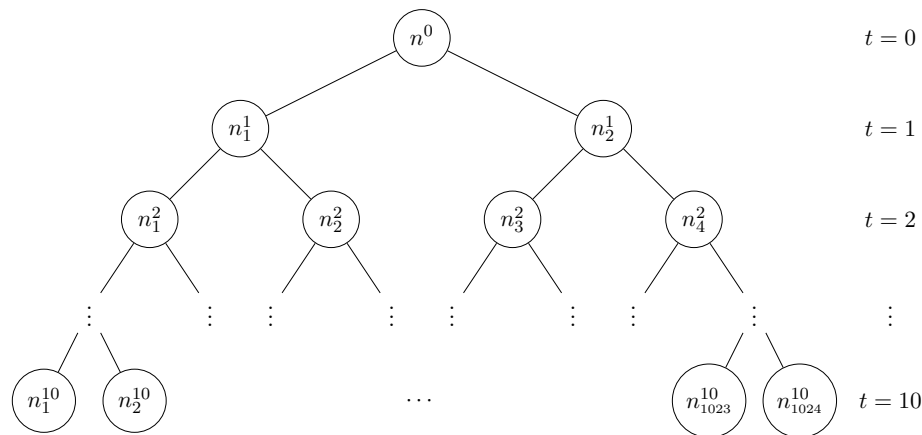


Figure 2: Event tree graph for  $T = 10$ .

As in the previous example, we compute the cooperative payoffs as well as the benefits from deviating from the cooperative solution in all subgames. To save on space, in Tables 4 and 5, we only print the results regarding the benefits of cheating on the agreement and the values of  $\varepsilon_j(n_i^t)$ . (More precisely, we only show the max values for each time period.) These tables show that the first time a player could benefit from deviating from cooperation is player 1 in period 7. Interestingly, the other two players would only deviate in the last period. The value of  $\varepsilon$  is equal to 45.1047. As under a cooperative regime, the total accumulated pollution is lower than under noncooperation. The results obtained here are encouraging, not only from an economic point of view, but also from environmental one.



## 5 Concluding remarks

We showed in this paper how to construct an  $\varepsilon$ -cooperative equilibrium for the class of dynamic games played over event trees. As this class offers a natural modeling framework in many areas, such as renewable resources and environmental management, and given that cooperation is often desirable to, e.g., sustain the resource or reduce emissions of pollutants, our developments can help in the design of long-term sustainable agreements.

Here we suggested using minmax strategies to deter defection. The same approach can be followed to implement other punishing strategies. For instance, the players can decide to punish a deviator by playing their Nash strategies. If it is the case, then it suffices to specify this in the trigger mode of the behavior strategy. Clearly, however, we would expect to obtain a higher value for  $\varepsilon$ , and it is likely that the players will start deviating earlier than they would under minmax.

## References

- [1] Angelova, V., Bruttel, L.V., Güth, W., Kamecke, U. (2013). Can subgame perfect equilibrium threats foster cooperation? An experimental test of finite-horizon folk theorems. *Economic Inquiry* 51, 1345–1356.
- [2] Aumann, R.J. and Shapley, L.S. (1994). Long-term competition: a game-theoretic analysis. In Megiddo N. (ed.), *Essays in Game Theory in Honor of Michael Maschler*. Springer, 1–15.
- [3] Benoit, J.-P. and Krishna, V. (1985). Finitely repeated games. *Econometrica* 53(4), 905–922.
- [4] Chistyakov, S. and Petrosyan, L. (2013). Strong strategic support of cooperative solutions in differential games. *Advances in Dynamic Games, Annals of the International Society of Dynamic Games* 12, 99–107.
- [5] Dutta P. (1995). A folk theorem for stochastic games. *Journal of Economic Theory* 66, 1–32.
- [6] Eswaran, M. and Lewis, T. (1986) Collusive behaviour in finite repeated games with bonding. *Economics Letters* 20, 213–216.
- [7] Flesh, J., Kuipers, J., Mashiah-Yaakovi, A., Shmaya, E., Shoenmakers, G., Solan, E., and Vrieze, K. (2014). Nonexistence of subgame-perfect-equilibrium in perfect information games with infinite horizon. *International Journal of Game Theory* 43, 945-951.
- [8] Flesh, J., Predtethinski, A. (2015). On refinements of subgame perfect  $\epsilon$ -equilibrium. *International Journal of Game Theory*, <http://dx.doi.org/10.1007/s00182-015-0468-8>.
- [9] Genc, T., Reynolds, S.S. and Sen, S. (2007). Dynamic oligopolistic games under uncertainty: A stochastic programming approach. *Journal of Economic Dynamics & Control* 31, 55–80.
- [10] Genc, T. and Sen, S. (2008). An analysis of capacity and price trajectories for the Ontario electricity market using dynamic Nash equilibrium under uncertainty. *Energy Economics* 30, 173–191.
- [11] Germain, M., Toint, P., Tulkens, H., and de Zeeuw A. (2003). Transfers to sustain dynamic core-theoretic cooperation in international stock pollutant control. *Journal of Economic Dynamics & Control* 28(1), 79–99.
- [12] Haurie, A., Krawczyk, J.B. and Zaccour, G. (2012). *Games and Dynamic Games*, Scientific World, Singapore.
- [13] Haurie, A., Zaccour, G. and Smeers, Y. (1990). Stochastic equilibrium programming for dynamic oligopolistic markets. *Journal of Optimization Theory and Applications* 66(2), 243–253.
- [14] Kalai, E. (1987). Bounded rationality and strategic complexity in repeated games. *Game theory and applications, Pro. Int. Conf., Columbus/OH (USA)*, 131-157.
- [15] Mailath, G. J., Postlewaite, A., Samuelson, L. (2005) Contemporaneous perfect epsilon-equilibria. *Games and Economic Behavior* 53, 126–140.
- [16] Parilina, E. (2014). Strategic stability of one-point optimality principles in cooperative stochastic games. *Mathematical Game Theory and Applications* 6(1), 56–72 (in Russian).
- [17] Parilina, E., and Zaccour, G. (2015). Node-consistent core for games played over event trees. *Automatica* 55, 304–311, 2015.
- [18] Pineau, P.-O. and Murto, P. (2003). An Oligopolistic Investment Model of the Finnish Electricity Market, *Annals of Operations Research* 121, 123–148.
- [19] Pineau, P.-O., Rasata, H. and Zaccour, G. (2011). Impact of some parameters on investments in oligopolistic electricity markets. *European Journal of Operational Research* 213(1), 180–195.
- [20] Radner, R. (1980) Collusive behavior in noncooperative epsilon-equilibria of oligopolies with long but finite lives. *Journal of Economic Theory* 22, 136–154.

- 
- [21] Reddy, P.V., Shevkoplyas E. and Zaccour, G. (2013). Time-consistent Shapley value for games played over event Trees. *Automatica* 49(6), 1521–1527.
  - [22] Rosen, J.B. (1965). Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica* 33(3), 520–534.
  - [23] Solan, E., Vieille, N. (2003). Deterministic multi-player Dynkin games. *Journal of Mathematical Economics* 39, 911-929.
  - [24] Tolwinski, B., Haurie, A., and Leitmann, G. (1986). Cooperative equilibria in differential games. *Journal of Mathematical Analysis and Applications* 119(1–2), 182–202.
  - [25] Yeung, D.W.K. and Petrosyan, L. (2005). *Cooperative Stochastic Differential Games*, Springer, New York, NY.
  - [26] Zaccour, G. (1987). *Théorie des jeux et marchés énergétiques: marché européen de gaz naturel et échanges d'électricité*, Ph.D. Thesis, HEC Montréal.
  - [27] Zaccour, G. (2008). Time consistency in cooperative differential games: A Tutorial. *INFOR* 46(1), 81–92.