Les Cahiers du GERAD ISSN: 0711–2440

Robust dynamic games played over event trees

E. Parilina, G. Zaccour

G–2025–60

September 2025

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, 2025
– Bibliothèque et Archives Canada, 2025

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, 2025
– Library and Archives Canada, 2025

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Robust dynamic games played over event trees

Elena Parilina [a, b]

Georges Zaccour [c, d]

[a] *Saint Petersburg State University, Saint Petersburg, Russia*

[b] *Qingdao University, School of Mathematics and Statistics, Qingdao 266071, PR China*

[c] *Chair in Game Theory and Management, Montréal (Qc), Canada, H3T 1J4*

[d] *Department of Decision Sciences, HEC Montréal & GERAD, Montréal, (Qc), Canada, H3T 2A7*

e.parilina@spbu.ru
georges.zaccour@gerad.ca

September 2025
Les Cahiers du GERAD
G–2025–60

Abstract : We characterize cooperative and non-cooperative solutions of a dynamic game played over event tree when the transition probability are not given. We assume that the players implement feedback strategies that are adapted to the nodes of the event tree and to the state variable. A robust control is adopted, that is, an approach where the players consider that nature will choose the worst probability distribution with respect to their payoffs. We illustrate our results with an environmental management example.

Keywords : Stochastic games; event tree; S-adapted strategies; robust optimization

Résumé : Nous caractérisons les solutions coopératives et non coopératives d'un jeu dynamique joué sur un arbre d'événements lorsque la probabilité de transition n'est pas donnée. Nous supposons que les joueurs mettent en œuvre des stratégies de rétroaction adaptées aux nœuds de l'arbre d'événements et à la variable d'état. Un contrôle robuste est adopté, c'est-à-dire une approche où les joueurs considèrent que la nature choisira la pire distribution de probabilités par rapport à leurs gains. Nous illustrons nos résultats par un exemple de gestion environnementale.

Mots clés : Jeux stochastiques; arbre d'événements; stratégies S-adaptées; optimisation robuste

Acknowledgements: The work of the second author was supported by RGPIN–2021–02462, NSERC, Canada.

1 Introduction

Many problems in economics, engineering, and management involve (i) a small number of agents who compete, or cooperate, repeatedly over time; (ii) an accumulation process, e.g., production capacity, pollution stock, brand reputation, that affects the players' current and future payoffs; and (iii) an uncertainty about the model's parameter values. These features point out toward stochastic dynamic games as a natural methodological framework to determine the outcomes. When the game is in discrete time and the players' actions do not affect the transition probabilities of the stochastic process, then an event tree becomes an attractive tool to model the uncertainty. A game that encompasses these features is called a dynamic game played over event tree (DGPET).

The class of DGPETs was initially developed in Zaccour (1987) and Haurie et al. (1990) to model long-term competition among the few exporters of natural gas to some European markets represented by stochastic demands. Pineau and Murto (2003) and Pineau eta l. (2011) adopted a DGPET model to analyze strategic investments in the electricity industry in Finland, and Genc et al. (2007) in Ontario, Canada. The information structure, that is, the piece of information used by the players to make their decisions, was termed S-adapted information structure, where S stands for sample value of the stochastic process. The resulting equilibrium was referred to as S-adapted open-loop equilibrium to highlight that the strategies are adapted to the nodes, but do not depend on the state variables. Haurie and Moresino (2001, 2002) dealt with the computation of S-adapted open-loop equilibria in oligopolistic markets.

Some more recent developments focused on cooperative DGPETs. Reddy et al. (2013) proposed a node-consistent Shapley value and Parilina and Zaccour (2015a) a node-consistent core to sustain cooperation over the whole duration of the game. Parilina and Zaccour (2015b) constructed an ε-cooperative equilibrium and Parilina et al. (2017) determined the price of anarchy in linear-state DGPETs. Parilina and Zaccour (2017) extended the framework to a case where the terminal time is random, and Parilina and Zaccour (2022) to a scenario where the players have asymmetric beliefs.[1]

In all contributions to DGPETs, the authors assumed (i) that the transition probabilities in the event tree are given and adopted, and (ii) that the players implement S-adapted open-loop strategies with the exception of Reddy and Zaccour (2019) who considered S-adapted feedback strategies. The objective of this paper is to determine noncooperative (Nash equilibrium) and cooperative (joint optimization) solutions in S-adpated feedback strategies in DGPETs when the set of possible values of the stochastic parameters is given at each node of the event tree, but not the transition probabilities from one node to its successors. Ellsberg (1961) refers to such uncertainty by *ambiguity*, which is defined explicitly in Arend (2020) as follows: "*Ambiguity exists when the set of future possible outcomes is known, as is the set of actions possible to take, as are the payoffs for each action in each outcome, but the (absolute) probabilities of those outcomes are all unknowable prior to having to make the decision.*" Further, as they lead to subgame-perfect equilibria, S-adpated feedback strategies are conceptually more appealing than S-adpated open-loop strategies.

One way to deal with ambiguity is to consider all possible scenarios (a finite set), keep their probabilities unknown, and use a robust-control approach to solve the problem assuming that Nature "plays" against the decision maker by minimizing her payoff (Başar and Bernhard (2008)). In this framework, a certain measure of robustness is sought against uncertainty that can be represented as deterministic variability in the value of the parameters in the state dynamics or cost functions. Bauso et al. (2016) implement this idea of robustness in a mean-field game with identical players, and Bauso and Timmer (2012) apply the robust optimization technique to coalitional games.

An alternative approach to deal with uncertainty is "scenario analysis", which aims at reducing the number of possible values of the unknown variable up to a finite number and assume the probabilities of these scenarios (Høyland and Wallace (2001); Kaut (2021)). This approach requires statistical data

[1]See the textbook by Parilina et al. (2022) for a general introduction to cooperative and noncooperative DGPETs.

and involves experts in the modeling, which may render its implementation complicated in situations where the aggregation of subjective information is an issue, as in, e.g., a noncooperative game. Parilina and Zaccour (2022) considered three procedures for aggregating players' beliefs (subjective probabilities) in a cooperative DGPET, namely, average weighting of beliefs, geometric average of beliefs, and multiplicative belief aggregation rule, and obtained that the aggregation rule has a significant impact on the results. For a comparison of scenarios analysis and robust-control design in a single player context, see Calafiore and Campi (2006).

The closest contributions to ours are Jank and Kun (2002), and Engwerda (2017, 2022). Jank and Kun (2002) provided a definition of a robust (or worst-case) Nash equilibrium in linear-quadratic differential games (LQDGs), and Engwerda (2017) stated the necessary and sufficient conditions for open-loop Nash equilibria for this class of games. Further, Engwerda (2022) designed equilibrium control policies in a LQDG where the uncertainty is affected by the players who are assumed to be risk-averse. Our contribution with respect to these references is twofold. First, in Engwerda (2017, 2022) and Jank and Kun (2002), the uncertainty appears additively in the cost function and in the right-hand side of the state dynamics. In our paper, the uncertainty is not additive and is embedded directly in the objective functional and in the function governing the evolution of the state. Second, and more importantly, we consider feedback and not open-loop strategies. Feedback strategies are state-dependent and, as recalled above, lead to subgame-perfect equilibria. To the best of our knowledge, this is the first attempt to determine robust Nash equilibrium in feedback strategies in nonzero-sum games.

The rest of the paper is organized as follows: In Section 2, we introduce a DGPET with ambiguity. In Sections 3.1 and 3.2, we derive a robust cooperative and noncooperative solution, respectively. We provide an illustrative example in Section 4, and briefly conclude in Section 5.

2 A DGPET with ambiguity

In this section, we introduce a DGPET in a setup where the structure of the event tree is known but not the transition probabilities of passing through nodes.

Let $\mathbb{T} = \{0, 1, \ldots, T\}$ be the set of time periods. The event tree is represented by a finite set of nodes $\mathbf{n}_t$ in period $t \in \mathbb{T}$ with a root node n_0 in period 0. Any node $n_t \in \mathbf{n}_t$ represents a possible realization of the history of the game process up to time t. We denote by $n_t^- \in \mathbf{n}_{t-1}$ the unique predecessor of node $n_t \in \mathbf{n}_t$. The set of all possible direct successors of node n_t is denoted by $\mathbf{n}_t^+ \subseteq \mathbf{n}_{t+1}$, and the set of all nodes of the event tree having n_t as their root node by $\mathbf{n}_t^{++}$. Consequently, the entire event tree is denoted by $\mathbf{n}_0^{++}$. When node n_t is reached at time t, the subtree $\mathbf{n}_t^{++}$ emanating from node n_t represents the residual uncertainty. A *scenario* is a path $[n_0, n_T]$ from the root node to node n_T, at which the game terminates. We denote by $[n_0, n]$ the set of all nodes encountered along the unique sample path starting from n_0 and ending at n. Let the set of all scenarios in the event tree be denoted by $\mathcal{S}$, that is, $\mathcal{S} = \{[n_0, n_T] : n_T \in \mathbf{n}_T\}$, and the set of scenarios in the event subtree emanating from node $n \in \mathbf{n}_0^{++}$ be denoted by $\mathcal{S}(n)$, that is, $\mathcal{S}(n) = \{[n, n_T] : n_T \in \mathbf{n}_T\}$.

Passing through node n_t, the stochastic process transits to one of the successor nodes from the set $\mathbf{n}_t^+$. The process certainly starts from node n_0 and terminates in one of the nodes from the set $\mathbf{n}_T$. The transition from node $n_t \in \mathbf{n}_t$, $t \in \mathbb{T} \setminus \{T\}$ to one of its successors belonging to the set $\mathbf{n}_t^+$ is stochastic and defined by a discrete probability measure $\mathbf{p}_{n_t} = (p_{n_t}^{n_{t+1}} : n_{t+1} \in \mathbf{n}_t^+)$. (It is discrete because the set of successors is a finite set of nodes.) Let the set of such probability measures be denoted by $\mathcal{P}_{n_t}$. The transition probabilities for different nodes are independent, and are independent of players' actions. Therefore, the probability of a scenario, say, $s = [n_0, n_3] = [n_0, n_1, n_2, n_3]$, where $n_1 \in \mathbf{n}_0^+$, $n_2 \in \mathbf{n}_1^+$, $n_3 \in \mathbf{n}_2^+$, is $P(s) = p_{n_0}^{n_1} \cdot p_{n_1}^{n_2} \cdot p_{n_2}^{n_3}$. Using transition probability measures we can compute the probability of passing through a node n_t, that is, the probability of the unique path $[n_0, n_1, \ldots, n_{t-l}, n_t]$, which is equal to $\pi_{n_t} = p_{n_0}^{n_1} \cdot \ldots \cdot p_{n_{t-1}}^{n_t}$.

When the transition probabilities are known, the game can be classified as *a game with uncertainty*, where uncertainty is given by known probability measure P_n for any $n \in \mathbf{n}_0^{++} \backslash \mathbf{n}_T$. When the transition probability measure P_n is unknown for any $n \in \mathbf{n}_0^{++} \setminus \mathbf{n}_T$, then the uncertainty is referred to as *ambiguity* in the stochastic programming literature (Epstein and Schneider (2007), Iyengar (2005), Nilim and El Ghaoui (2005)). Adopting this terminology, we can refer to our game as DGPET with ambiguity.

Denote by $M = \{1, 2, \ldots, m\}$ the set of players. Now, we introduce the remaining elements of a DGPET with ambiguity:

Control variables. Denote by $u^i_{n_t} \in U^i_{n_t} \subseteq \mathbb{R}^{m_i}$ the vector of control variables of Player i at node $n_t \in \mathbf{n}_t$, $t \in \mathbb{T} \backslash \{T\}$, where $U^i_{n_t}$ is the set of control values of Player i of dimension m_i. Denote by $\mathbf{U}^i$ the product of decision variable sets of Player i, i.e., $\prod_{\nu \in \mathbf{n}_0^{++} \backslash \mathbf{n}_T} U^i_\nu$ and by $U_{n_t} = \prod_{i \in M} U^i_{n_t}$ the joint decision set for all players at node n_t;

State variables. Denote by $x_{n_t} \in X \subset \mathbb{R}^q$ the state vector at node $n_t \in \mathbf{n}_t$, $t \in \mathbb{T}$, where X is the set of states of dimension q. The evolution of the state over nodes is given by the transition function $f^{n_t}(\cdot, \cdot) : X \times U_{n_t} \mapsto X$ defined for any node n_t. The state equations are given by

$$x_{n_t} = f^{n_t^-}\left(x_{n_t^-}, u_{n_t^-}\right), \qquad x_{n_0} = x_0, \tag{1}$$

where $u_{n_t^-} \in U_{n_t^-}$, $n_t \in \mathbf{n}_t$, $t \in \mathbb{T} \setminus \{0\}$, and x_0 is the initial state given for the root node n_0. Note that the dynamics in (1) imply that all successor nodes of $n_t \in \mathbf{n}_t$, that is, $n_{t+1} \in \mathbf{n}_t^+$, have the same state values;

Rewards. At any nonterminal node $n_t \in \mathbf{n}_t$, $t = 0, \ldots, T-1$, the reward to Player i is $\phi_i(n_t, x_{n_t}, u_{n_t})$ and it is $\Phi_i(n_T, x_{n_T})$ for any terminal node $n_T \in \mathbf{n}_T$. If the players choose the control

$$\mathbf{u} = \{u^i_n \in U^i_n : n \in \mathbf{n}_0^{++} \setminus \mathbf{n}_T,\ i \in M\},$$

then the state trajectory $\mathbf{x}$ is defined as a unique solution of (1), i.e.,

$$\mathbf{x} = \{x_n : n \in \mathbf{n}_0^{++}\},$$

and Player $i \in M$ obtains in a particular scenario $s = [n_0, n_T] \in \mathcal{S}$ the payoff

$$J^s_i(n_0, \mathbf{x}, \mathbf{u}) = \sum_{t=0}^{T-1} \rho^t \sum_{n_t \in \mathbf{n}_t \cap s} \phi_i(n_t, x_{n_t}, u_{n_t}) + \rho^T \Phi_i(n_T, x_{n_T}), \tag{2}$$

where $n_T \in \mathbf{n}_T$ is a node at which scenario s terminates, and $\rho \in (0, 1)$ is the discount factor. Player i's payoff in the subgame starting from node $n_t \in \mathbf{n}_t$, $t \in \mathbb{T} \setminus \{T\}$, in scenario $s' = [n_t, n_T] \in \mathcal{S}(n_t)$ is given by

$$J^{s'}_i(n_t, \mathbf{x}(n_t), \mathbf{u}(n_t)) = \sum_{\tau=t}^{T-1} \rho^{\tau-t} \sum_{\substack{n_\tau \in \\ \mathbf{n}_\tau \cap s'}} \phi_i(n_\tau, x_{n_\tau}, u_{n_\tau}) + \rho^{T-t} \Phi_i(n_T, x_{n_T}), \tag{3}$$

where the state and control trajectories $\mathbf{x}(n_t)$ and $\mathbf{u}(n_t)$ are defined as

$$\mathbf{x}(n_t) = \{x_n : n \in \mathbf{n}_t^{++}\}, \tag{4}$$

$$\mathbf{u}(n_t) = \{u^i_n \in U^i_n : n \in \mathbf{n}_t^{++} \setminus \mathbf{n}_T,\ i \in M\}, \tag{5}$$

and $\mathbf{x}(n_t)$ satisfies state dynamics equation (1) with initial state $x_{n_t} = \bar{x}_{n_t}$.

Information structure. An S-adapted feedback (or Markovian) information structure. That means that the players use the position of the game (n_t, x_{n_t}) as information basis to make their decisions at node n_t.

Strategies. Let Ψ_i be the set of strategies of Player i. An S-adapted feedback strategy $\psi_i(n_t, x_{n_t}) \in \Psi_i$ is a decision rule that defines Player i's control $u^i_{n_t} \in U^i_{n_t}$ for any node n_t, $t \in \mathbb{T} \setminus \{T\}$, and any state $x_{n_t} \in X$. Denote by $\psi(n_t, x_{n_t}) = (\psi_i(n_t, x_{n_t}) : i \in M)$ the strategy profile at node n_t and by $u(n_t) = (u_i(n_t) : i \in M)$ the corresponding control profile.

Remark 1. We are dealing with ambiguity/uncertainty about probability measures defining the transitions from one node to another in the event tree by using a robust approach and finding the solution with Bellman equation. To use it, we must satisfy the *rectangularity assumption*, which can be formulated for our event tree problem as follows: For any scenario $s \in \mathcal{S}$ in the game starting from node n_0, the set of probability measures defining the probability of the scenario s realization satisfies the condition $\mathcal{P} = \prod_{n \in s} \mathcal{P}_n$. This condition is clearly satisfied in our case since the transition probabilities are independent along any scenario.

We notice that the players choose their strategies at any node from set $\mathbf{n}_0^{++} \setminus \mathbf{n}_T$ while facing an ambiguity about a forthcoming node. We use a worst-case or robust approach frequently applied in optimization, control, and game-theoretical problems (Başar and Bernhard (2008); Iyengar (2005); Engwerda (2017); Jank and Kun (2002)). In this approach, the players choose their strategies assuming the "worst" realization of the stochastic process, i.e., at any intermediate node, the players consider all probability measures defining probabilities of realization of all possible scenarios emanating from this node and choose their strategies to maximize the payoff functional along the worst scenario or under the worst probability measure. When the players follow a robust approach to solve the game, their behavior can be called "pessimistic" as they assume that Nature defining transitions along the tree always choose the worst ones that minimize their payoffs.

3 Solutions

In this section, we determine cooperative and noncooperative solutions in DGPETs with ambiguity.

3.1 Cooperative solution

We assume that the players agree to form a grand coalition and choose their strategies $\psi_i(n_t, x_{n_t}) \in \Psi_i$, for all $i \in M$, $n_t \in \mathbf{n}_t$, and $t = 1, \ldots, T-1$ that maximize their joint payoff in a robust-optimization sense. Taking into account that the scenario in the subgame starting from any node n_t is stochastic and defined by the probability measure $P(n_t) \in \mathcal{P}(n_t)$, the total payoff of the grand coalition in the subgame starting from node n_t in time t with state x_{n_t} under strategy profile $\psi(n_t, x_{n_t})$ is given by

$$\begin{aligned} W(n_t, x_{n_t}, \psi(n_t, x_{n_t})) &= \min_{P(n_t) \in \mathcal{P}(n_t)} \mathbf{E}^{P(n_t)} \left[\sum_{i \in M} J_i(n_t, \mathbf{x}(n_t), \mathbf{u}(n_t)) \right] \\ &= \min_{P(n_t) \in \mathcal{P}(n_t)} \mathbf{E}^{P(n_t)} \left[\sum_{\tau=t}^{T-1} \rho^{\tau-t} \sum_{n_\tau \in \mathbf{n}_\tau} \sum_{i \in M} \phi_i(n_\tau, x_{n_\tau}, u_{n_\tau}) + \rho^{T-t} \sum_{i \in M} \Phi_i(n_T, x_{n_T}) \right], \end{aligned} \tag{6}$$

s. t. dynamics equation (1), where $\mathbf{E}^{P(n_t)}$ denotes the expectation over stochastic scenario $[n_t, n_T]$, that is, a collection of nodes $[n_t, n_{t+1}, \ldots, n_{T-1}, n_T]$ under probability measure $P(n_t) \in \mathcal{P}(n_t)$. Equation (6) defines the payoff of grand coalition M in the subgame starting from node n_t under profile of decision rules $\psi(n_t, x_{n_t})$ as a minimum expected payoff over all probability measures from set $\mathcal{P}(n_t)$, while in equation (3) the payoff of M in a particular scenario s' is defined. The first term in the square brackets of the last equality in (6) is Player i's payoff along a particular scenario starting from time t to $T-1$, and the second term is the terminal payoff of Player i. We should mention that the scenario $[n_t, n_T]$ is stochastic and it is realised under the probability measure $P(n_t)$. The robust optimization problem of coalition M in the subgame starting from node $n_t \in \mathbf{n}_t$, $t \in \mathbb{T} \setminus \{T\}$, with state x_{n_t} is

$$\begin{aligned} W^*(n_t, x_{n_t}) &= \max_{\psi(n_t, x_{n_t}) \in \Psi} W(n_t, x_{n_t}, \psi(n_t, x_{n_t})) \\ &= \max_{\psi(n_t, x_{n_t}) \in \Psi} \min_{P(n_t) \in \mathcal{P}(n_t)} \mathbf{E}^{P(n_t)} \left[\sum_{\tau=t}^{T-1} \rho^{\tau-t} \sum_{n_\tau \in \mathbf{n}_\tau} \sum_{i \in M} \phi_i(n_\tau, x_{n_\tau}, u_{n_\tau}) + \rho^{T-t} \sum_{i \in M} \Phi_i(n_T, x_{n_T}) \right], \end{aligned} \tag{7}$$

s. t. dynamics equation (1), and an optimal strategy profile is $\psi^*(n_t, x_{n_t})$ at which the maximum in (7) is achieved. This is a *cooperative robust strategy profile* for subgame staring at node n_t. Let the strategy profile $\psi^*(n_t, x_{n_t})$ generate in this subgame the collection of *cooperative robust controls* $\mathbf{u}^*(n_t) = (\mathbf{u}_i^*(n_t) : i \in M)$, where $\mathbf{u}_i^*(n_t) = \{u_n^i(n_t) : n \in \mathbf{n}_t^{++} \setminus \mathbf{n}_T\}$ and corresponding states $\mathbf{x}^*(n_t) = \{x_n^*(n_t) : n \in \mathbf{n}_t^{++}\}$ found as a unique solution of equations (1) with initial state $x_{n_t}(n_t) = \bar{x}_{n_t}$ substituting cooperative robust controls.

Assuming that the probability measure solving the minimization problem (6) defines a deterministic scenario, say scenario $s' = [n_t, n_T] \in \mathcal{S}(n_t)$, Player i's payoff in the game starting at node n_0 with state x_0 under strategy profile $\psi^*(n_t, x_{n_t})$ when scenario s' appears is defined by

$$J_i^{s'}(n_t, \mathbf{x}^*(n_t), \mathbf{u}^*(n_t)) = \sum_{\tau=t}^{T-1} \rho^{\tau-t} \sum_{n_\tau \in \mathbf{n}_\tau \cap s'} \phi_i(n_\tau, x_{n_\tau}^*(n_t), u_{n_\tau}^*(n_t)) + \rho^T \Phi_i(n_T, x_{n_T}^*(n_t)).$$

We make the following assumption about the existence and uniqueness of solutions of the optimization problem in (7):

A1: There exist a unique solution of maximization problem $\max_{\psi(n_t, x_{n_t}) \in \Psi(n_t)} W(n_t, x_{n_t}, \psi(n_t, x_{n_t}))$ for the subgame staring from any intermediate node n_t with state x_{n_t}.

This assumption guarantees that there is a unique cooperative strategy profile and that players' strategies are known and well defined along each scenario of the game. If there are many solutions of the maximization problem in (7), we need to consider one that the players have agreed to realize to avoid nonuniqueness in the choice of cooperative control.

The following theorem characterizes the profile of robust strategies that solves (7) for any subgame starting at node n_t with state x_{n_t}.

Theorem 1. The collection of functions $\{W^*(n_t, x_{n_t}) : n_t \in \mathbf{n}_t, t \in \mathbb{T}\}$ satisfies the following Bellman equation:

$$W^*(n_T, x_{n_T}) = \sum_{i \in M} \Phi_i(n_T, x_{n_T}),\ n_T \in \mathbf{n}_T,$$

$$W^*(n_t, x_{n_t}) = \max_{u_{n_t} \in U_{n_t}} \left[\sum_{i \in M} \phi_i(n_t, x_{n_t}, u_{n_t}) + \rho \min_{\mathbf{p}_{n_t} \in \mathcal{P}_{n_t}} \mathbf{E}^{\mathbf{p}_{n_t}} W^*(n_{t+1}, x_{n_{t+1}}) \right], \quad n_t \in \mathbf{n}_t, t \in \mathbb{T} \setminus \{T\},$$

s.t. state equation (1), where $\mathbf{p}_{n_t} \in \mathcal{P}_{n_t}$ is a discrete probability measure $\mathbf{p}_{n_t} = (p_{n_t}^{n_{t+1}} : n_{t+1} \in \mathbf{n}_t^+)$ defining transition probabilities from n_t to its successors belonging to set $\mathbf{n}_t^+$.

Proof. Consider any node $n_t \in \mathbf{n}_t, t \in \mathbb{T} \setminus \{T\}$. Since transition probabilities are independent for different stages, from (7) it follows that

$$\begin{aligned} W^*(n_t, x_{n_t}) &= \max_{\psi(n_t, x_{n_t}) \in \Psi} \min_{P(n_t) \in \mathcal{P}(n_t)} \mathbf{E}^{P(n_t)} \left[\sum_{\tau=t}^{T-1} \rho^{\tau-t} \sum_{n_\tau \in \mathbf{n}_\tau} \sum_{i \in M} \phi_i(n_\tau, x_{n_\tau}, u_{n_\tau}) + \rho^{T-t} \sum_{i \in M} \Phi_i(n_T, x_{n_T}) \right] \\ &= \max_{\psi(n_t, x_{n_t}) \in \Psi} \min_{\left(\mathbf{p}_{n_t}, P(n_{t+1})\right) \in \mathcal{P}_{n_t} \times \mathcal{P}(n_{t+1})} \mathbf{E}^{\left(\mathbf{p}_{n_t}, P(n_{t+1})\right)} \left[\sum_{\tau=t}^{T-1} \rho^{\tau-t} \sum_{n_\tau \in \mathbf{n}_\tau} \sum_{i \in M} \phi_i(n_\tau, x_{n_\tau}, u_{n_\tau}) \right. \\ &\quad \left. + \rho^{T-t} \sum_{i \in M} \Phi_i(n_T, x_{n_T}) \right]. \end{aligned}$$

As $\sum_{i \in M} \phi_i(n_t, x_{n_t}, u_{n_t})$ does not depend on probability measure $(\mathbf{p}_{n_t}, P(n_{t+1}))$, we have:

$$W^*(n_t, x_{n_t}) = \max_{\psi(n_t, x_{n_t}) \in \Psi} \left[\sum_{i \in M} \phi_i(n_t, x_{n_t}, u_{n_t}) \right.$$

$$
\begin{aligned}
&+ \min_{(\mathbf{p}_{n_t}, P(n_{t+1}))\in\mathcal{P}_{n_t}\times\mathcal{P}(n_{t+1})} \mathbf{E}^{(\mathbf{p}_{n_t}, P(n_{t+1}))} \Bigg\{ \sum_{\tau=t+1}^{T-1} \rho^{\tau-t} \cdot \\
&\cdot \sum_{n_\tau\in\mathbf{n}_\tau} \sum_{i\in M} \phi_i(n_\tau, x_{n_\tau}, u_{n_\tau}) + \rho^{T-t} \sum_{i\in M} \Phi_i(n_T, x_{n_T}) \Bigg\} \Bigg] \\
= & \max_{\psi(n_t, x_{n_t})\in\Psi} \Bigg[\sum_{i\in M} \phi_i(n_t, x_{n_t}, u_{n_t}) \\
&+ \min_{\mathbf{p}_{n_t}\in\mathcal{P}_{n_t}} \mathbf{E}^{\mathbf{p}_{n_t}} \Bigg(\min_{P(n_{t+1})\in\mathcal{P}(n_{t+1})} \mathbf{E}^{P(n_{t+1})} \Bigg\{ \sum_{\tau=t+1}^{T-1} \rho^{\tau-t} \cdot \\
&\cdot \sum_{n_\tau\in\mathbf{n}_\tau} \sum_{i\in M} \phi_i(n_\tau, x_{n_\tau}, u_{n_\tau}) + \rho^{T-t} \sum_{i\in M} \Phi_i(n_T, x_{n_T}) \Bigg\} \Bigg) \Bigg] \\
= & \max_{\psi(n_t, x_{n_t})\in\Psi} \Bigg[\sum_{i\in M} \phi_i(n_t, x_{n_t}, u_{n_t}) \\
&+\rho \min_{\mathbf{p}_{n_t}\in\mathcal{P}_{n_t}} \mathbf{E}^{\mathbf{p}_{n_t}} W(n_{t+1}, x_{n_{t+1}}, \psi(n_{t+1}, x_{n_{t+1}})) \Bigg],
\end{aligned}
$$

where the last equality follows from (6). We proceed equalities using the independence of transition probabilities on strategy profiles:

$$
\begin{aligned}
W^*(n_t, x_{n_t}) = & \max_{u_{n_t}\in U_{n_t}} \Bigg[\sum_{i\in M} \phi_i(n_t, x_{n_t}, u_{n_t}) \\
&+\rho \min_{\mathbf{p}_{n_t}\in\mathcal{P}_{n_t}} \mathbf{E}^{\mathbf{p}_{n_t}} \max_{\psi(n_{t+1}, x_{n_{t+1}})\in\Psi} W(n_{t+1}, x_{n_{t+1}}, \psi(n_{t+1}, x_{n_{t+1}})) \Bigg] \\
= & \max_{u_{n_t}\in U_{n_t}} \Bigg[\sum_{i\in M} \phi_i(n_t, x_{n_t}, u_{n_t}) + \rho \min_{\mathbf{p}_{n_t}\in\mathcal{P}_{n_t}} \mathbf{E}^{\mathbf{p}_{n_t}} W^*(n_{t+1}, x_{n_{t+1}}) \Bigg],
\end{aligned}
$$

subject to state dynamics equation (1).

The boundary condition for the Bellman equation is

$$
W^*(n_T, x_{n_T}) = \sum_{i\in M} \Phi_i(n_T, x_{n_T}),
$$

for any terminal node $n_T \in \mathbf{n}_T$. □

Remark 2. The proof of Theorem 1 is based on solving the Bellman equation taking into account the stochastic nature of the game process. The reason why we can use the Bellman equation is an assumption made in Remark 1, i.e., that the discrete probability measures defining the transitions are independent along each scenario. This allows us to use the Bellman equation to solve the problem. The robust approach is implemented in the Bellman equation when we consider the worst case over all probability measures defining a particular scenario in the subgame starting from any node n_t.

Remark 3. We use the worst-case scenario, but one can assume the best-case scenario (optimistic approach) to define the solution. In this case, the grand coalition solves the following optimization problem:

$$
\max_{\psi(n_0, x_0)\in\Psi(n_0)} \max_{P\in\mathcal{P}} \mathbf{E}^P \Big[\sum_{t=0}^{T-1} \rho^t \sum_{n_t\in\mathbf{n}_t} \sum_{i\in M} \phi_i(n_t, x_{n_t}, u_{n_t}) + \rho^T \sum_{i\in M} \Phi_i(n_T, x_{n_T}) \Big], \tag{8}
$$

s. t. dynamics equation (1).

3.2 Robust Nash equilibrium

In this section, we consider a noncooperative mode of play and seek a Nash equilibrium. In our robust-optimization setup, the total payoff of Player $i \in M$ in the game starting from node n_0 with state x_0 under feedback strategy profile $\psi(n_0, x_0)$ is given by

$$\begin{aligned} W_i(n_0, x_0, \psi(n_0, x_0)) &= \min_{P_i(\mathbf{u}) \in \mathcal{P}} \mathbf{E}^{P_i(\mathbf{u})} J_i(n_0, \mathbf{x}, \mathbf{u}) \\ &= \min_{P_i(\mathbf{u}) \in \mathcal{P}} \mathbf{E}^{P_i(\mathbf{u})} \Big[\sum_{t=0}^{T-1} \rho^t \sum_{n_t \in \mathbf{n}_t} \phi_i(n_t, x_{n_t}, u_{n_t}) + \rho^T \Phi_i(n_T, x_{n_T}) \Big], \end{aligned} \tag{9}$$

s. t. dynamics equation (1). Equation (9) defines the worst-case payoff to Player i under the profile of strategies $\psi(n_0, x_0)$ as the minimum expected payoff over all probability measures from the set $\mathcal{P}$. The worst-case probability measure, under which the minimum in (9) is reached, defines the worst-case scenario for Player i, and it may be different from the worst-case scenario for other players. Therefore, the definition of the Nash/worst-case equilibrium in DGPETs is given in two steps: (i) we define the worst-case probability measure for each player, and consequently the worst-case scenario; and (ii) we find the Nash equilibrium assuming that each player chooses her strategies to maximize her total payoff under her worst-case scenario. This definition is inspired from Engwerda (2017) and Jank and Kun (2002).

Definition 1. Consider a DGPET starting at node n_0 with state x_0. Let $\mathbf{u} \in \mathbf{U}$:

1. The worst-case scenario for Player $i \in M$ under control profile $\mathbf{u}$ is given by probability distribution $\hat{P}_i(\mathbf{u}) \in \mathcal{P}$ such that

$$\mathbf{E}^{\hat{P}_i(\mathbf{u})} J_i(n_0, \mathbf{x}, \mathbf{u}) \leq \mathbf{E}^{P} J_i(n_0, \mathbf{x}, \mathbf{u}), \tag{10}$$

   for any $P \in \mathcal{P}$. The worst-case payoff $\mathbf{E}^{\hat{P}_i(\mathbf{u})} J_i(n_0, \mathbf{x}, \mathbf{u})$ is denoted by $W_i(n_0, x_0, \psi(n_0, x_0))$ in equation (9).

2. Assume for any $\mathbf{u} \in \mathbf{U}$, there exists a worst-case scenario $\hat{P}_i(\mathbf{u})$ for Player i. The control profile $\mathbf{u}^{ne} = (\mathbf{u}_i^{ne} : i \in M) \in \mathbf{U}$ is a Nash/worst-case equilibrium if for any $i \in M$ it holds that

$$\mathbf{E}^{\hat{P}_i(\mathbf{u}^{ne})} J_i(n_0, \mathbf{x}, \mathbf{u}^{ne}) \geq \mathbf{E}^{\hat{P}_i(\mathbf{u}_i, \mathbf{u}_{-i}^{ne})} J_i(n_0, \mathbf{x}, (\mathbf{u}_i, \mathbf{u}_{-i}^{ne})), \tag{11}$$

   for any control $\mathbf{u}_i \in \mathbf{U}^i$ and corresponding worst-case probability measure $\hat{P}_i(\mathbf{u}_i, \mathbf{u}_{-i}^{ne})$ for control profile $(\mathbf{u}_i, \mathbf{u}_{-i}^{ne})$.

In Definition 1, we assume that the players choose their strategies knowing the payoff functions of other players, and consequently, the worst-case probability measures of any other player. According to Definition 1 every player wants to secure against worst-case scenario or probability measure with respect to her interests.

Remark 4. The definitions of the Nash/worst-case equilibrium given for differential games by Engwerda (2017) and Jank and Kun (2002) differ in one assumption. In Jank and Kun (2002), the existence of the worst-case scenario $\hat{P}_i(\mathbf{u})$ for Player i is assumed for any $(\mathbf{u}_i, \mathbf{u}_{-i}^{ne}) \in \mathbf{U}$, while in Engwerda (2017) it is assumed for any $\mathbf{u} \in \mathbf{U}$. Therefore, in Engwerda (2017) such a Nash equilibrium is called global. We keep the latter assumption but omit "global" in Definition 1 for simplicity.

Definition 2. The $\mathcal{S}$-adapted feedback strategy profile $\psi^{ne}(n_0, x_0) = (\psi_i^{ne}(n_0, x_0) : i \in M) \in \Psi$ is a Nash/worst-case equilibrium if for any $i \in M$ it holds that

$$W_i(n_0, x_0, \psi^{ne}(n_0, x_0)) \geq W_i(n_0, x_0, (\psi_i(n_0, x_0), \psi_{-i}^{ne}(n_0, x_0))) \tag{12}$$

for any admissible $\mathcal{S}$-adapted strategy $\psi_i(n_0, x_0) \in \Psi_i$, where $\psi_{-i}^{ne}(n_0, x_0) = (\psi_j^{ne}(n_0, x_0) : j \in M, j \neq i)$ denotes the collection of $\mathcal{S}$-adapted feedback strategies of all players except Player i, and the worst-case payoff $W_i(n_0, x_0, \psi(n_0, x_0))$ under strategy profile $\psi(n_0, x_0)$ is defined by (9).

In a robust Nash equilibrium, each player's *pessimistic* strategy is a best response to other players' *pessimistic* strategies.

For any intermediate node, we define the Nash/worst-case equilibrium in the subgame starting in time $t = 1, \ldots, T-1$ from node $n_t \in \mathbf{n}_t$. Player i's feedback strategy in this subgame is denoted by $\psi_i(n_t, x_{n_t}) \in \Psi(n_t)$ defining controls in any node $n_\tau \in \mathbf{n}_\tau$ in time $\tau = t, \ldots, T-1$ such that $n_\tau \in \mathbf{n}_t^{++}$. We define the worst-case payoff for Player i in the subgame starting from node n_t in time t with state x_{n_t} under strategy profile $\psi(n_t, x_{n_t})$ as

$$\begin{aligned} W_i(n_t, x_{n_t}, \psi(n_t, x_{n_t})) &= \min_{P_i(n_t, \mathbf{u}(n_t)) \in \mathcal{P}(n_t)} \mathbf{E}^{P_i(n_t, \mathbf{u}(n_t))} \left[J_i(n_t, \mathbf{x}(n_t), \mathbf{u}(n_t)) \right] \\ &= \min_{P_i(n_t, \mathbf{u}(n_t)) \in \mathcal{P}(n_t)} \mathbf{E}^{P_i(n_t, \mathbf{u}(n_t))} \left[\sum_{\tau=t}^{T-1} \rho^{\tau-t} \sum_{n_\tau \in \mathbf{n}_\tau} \phi_i(n_\tau, x_{n_\tau}, u_{n_\tau}) + \rho^{T-t} \Phi_i(n_T, x_{n_T}) \right], \end{aligned} \tag{13}$$

s. t. dynamics equation (1), where $\mathbf{E}^{P_i(n_t, \mathbf{u}(n_t))}$ denotes the expectation over realization of scenario $[n_t, n_T]$, that is, a collection of nodes $[n_t, n_{t+1}, \ldots, n_{T-1}, n_T]$ under probability measure $P_i(n_t, \mathbf{u}(n_t)) \in \mathcal{P}(n_t)$. Equation (13) defines the worst-case payoff of Player i in the subgame starting from node n_t with state x_{n_t} under profile of decision rules $\psi(n_t, x_{n_t})$ as her minimum expected payoff over all probability measures from the set $\mathcal{P}(n_t)$.

Definition 3. The $\mathcal{S}$-adapted feedback strategy profile $\psi^{ne}(n_t, x_{n_t}) = (\psi_i^{ne}(n_t, x_{n_t}) : i \in M) \in \Psi(n_t)$ is a Nash/worst-case equilibrium in $\mathcal{S}$-adapted feedback strategies in the subgame of DGPET starting from node n_t with state x_{n_t} if for any $i \in M$ it holds that

$$W_i(n_t, x_{n_t}, \psi^{ne}(n_t, x_{n_t})) \geq W_i(n_t, x_{n_t}, (\psi_i(n_t, x_{n_t}), \psi_{-i}^{ne}(n_t, x_{n_t}))) \tag{14}$$

for any admissible $\mathcal{S}$-adapted strategy $\psi_i(n_t, x_{n_t}) \in \Psi_i(n_t)$, where the worst-case payoff $W_i(n_t, x_{n_t}, \psi(n_t, x_{n_t}))$ under strategy profile $\psi(n_t, x_{n_t})$ is defined by (13).

Let $(\psi_i^{ne}(n_t, x_{n_t}), \psi_{-i}^{ne}(n_t, x_{n_t}))$ be a feedback worst-case Nash equilibrium in the subgame starting from node n_t with state x_{n_t}, and $x^{ne}(n_t) = (x_\nu^{ne} : \nu \in \mathbf{n}_t^{++})$ is the associated state trajectory in this subgame emanating from (n_t, x_{n_t}). The value function for Player $i \in M$ is defined as

$$W_i^{ne}(n_T, x_{n_T}) = \Phi_i(n_T, x_{n_T}), \tag{15}$$

$$\begin{aligned} W_i^{ne}(n_t, x_{n_t}) = \min_{P_i(n_t, \mathbf{u}(n_t)) \in \mathcal{P}(n_t)} \mathbf{E}^{P_i(n_t, \mathbf{u}(n_t))} \Big[& \sum_{\tau=t}^{T-1} \rho^{\tau-t} \sum_{n_\tau \in \mathbf{n}_\tau} \phi_i(n_\tau, x_{n_\tau}^{ne}, (\psi_i^{ne}(n_\tau, x_{n_\tau}^{ne}), \psi_{-i}^{ne}(n_\tau, x_{n_\tau}^{ne}))) \\ & + \rho^{T-t} \Phi_i(n_T, x_{n_T}) \Big], \end{aligned} \tag{16}$$

s. t. dynamics equation (1). The value function $W_i^{ne}(n_t, x_{n_t})$ is the payoff to Player i if the worst-case Nash equilibrium in $\mathcal{S}$-adapted feedback strategies is played from initial node n_t with initial state x_{n_t}. The following theorem defines the equilibrium conditions over time (Bellman equations).

Theorem 2. The collection of value functions $\{W_i^{ne}(n_t, x_{n_t}) : n_t \in \mathbf{n}_t, t \in \mathbb{T}\}$, $i \in M$, defined in (15) and (16) satisfies the following recurrent backward in time (Bellman) equations:

$$W_i^{ne}(n_T, x_{n_T}^{ne}) = \Phi_i(n_T, x_{n_T}^{ne}), \tag{17}$$

$$\begin{aligned} W_i^{ne}(n_t, x_{n_t}^{ne}) = \max_{u_{n_t}^i \in U_{n_t}^i} \Big[& \phi_i(n_t, x_{n_t}^{ne}, (u_{n_t}^i, \psi_{-i}^{ne}(n_t, x_{n_t}^{ne}))) \\ & + \rho \min_{\substack{P^i(n_t) \\ \in \mathcal{P}_{n_t}}} \mathbf{E}^{P^i(n_t)} W_i^{ne}\big(n_{t+1}, f^{n_t}(x_{n_t}^{ne}, (u_{n_t}^i, \psi_{-i}^{ne}(n_t, x_{n_t}^{ne})))\big) \Big], n_t \in \mathbf{n}_t, t \in \mathbb{T} \setminus \{T\}, \end{aligned} \tag{18}$$

s.t. state equation (1), where $\mathbf{E}^{P^i(n_t)}$ denotes the expectation of realization of scenario $[n_t, n_T]$, that is, a collection of nodes $[n_t, n_{t+1}, \ldots, n_{T-1}, n_T]$ under probability measure $P^i(n_t) \in \mathcal{P}_{n_t}$.

Proof. We claim that for any player $i \in M$ the following equality holds true:

$$\psi_i^{ne}(n_t, x_{n_t}^{ne}) = \arg \max_{u_{n_t}^i \in U_{n_t}^i} \Big\{ \phi_i(n_t, x_{n_t}^{ne}, (u_{n_t}^i, \psi_{-i}^{ne}(n_t, x_{n_t}^{ne}))) + \rho \min_{P^i(n_t) \in \mathcal{P}_{n_t}} \mathbf{E}^{P^i(n_t)} W_i^{ne}(n_{t+1}, x_{n_{t+1}}^{ne}) \Big\}. \tag{19}$$

This means that the control assigned by a $\mathcal{S}$-adapted feedback worst-case equilibrium strategy of Player i at node n_t with the state $x_{n_t}^{ne}$ must be the best reply to the $\mathcal{S}$-adapted feedback equilibrium strategies of other players at node n_t with the state $x_{n_t}^{ne}$. The proof of the theorem can be easily done by recurrence. Given the boundary conditions (17), we consider all nodes $n_{T-1} \in n_T^-$ with state variable $x_{n_{T-1}}$. At any such node, the $\mathcal{S}$-adapted feedback equilibrium strategies $(\psi_i^{ne} : i \in M)$ cannot be improved by unilateral deviation by Player $j \in M$, which follows from equation (19). Therefore, the strategy profile $(\psi_i^{ne} : i \in M)$ is the worst-case/Nash equilibrium in $\mathcal{S}$-adapted feedback strategies at any node $n_{T-1} \in n_T^-$ with state variable $x_{n_{T-1}}$. We proceed in the same way from nodes $n_{T-2} \in \mathbf{n}_{T-2}$ to the root node n_0. □

Theorem 2 establishes necessary conditions for a worst-case/Nash equilibrium in $\mathcal{S}$-adapted feedback strategies.

4 Illustrative example

We illustrate our theoretical results with an example of transboundary pollution control.[2]

Denote by $M = \{1, \ldots, m\}$ the set of players representing countries and by $\mathbb{T} = \{0, \ldots, T\}$ the set of periods. Economic activities in each country generate revenues and, as a by-product, pollutant emissions, e.g., CO_2. Denote by $Q_{n_t}^i$ Player i's production of goods and services at node $n_t \in \mathbf{n}_t$, $t \in \mathbb{T} \setminus \{T\}$, and by $u_{n_t}^i$ the resulting emissions, with $u_{n_t}^i = h^i\left(Q_{n_t}^i\right)$ and $h^i(0) = 0$. Assuming a monotone increasing relationship between production and revenues, we can express the revenues R_i as function of emissions. To keep it simple, we adopt the following specification:

$$R_i\left(u_{n_t}^i\right) = \alpha u_{n_t}^i - \frac{1}{2}\left(u_{n_t}^i\right)^2, \tag{20}$$

where α is a positive parameter. We constrain the revenues to be nonnegative for any $n_t \in \mathbf{n}_0^{++} \setminus \mathbf{n}_T$ and any player i.

Denote by $u_{n_t} = (u_{n_t}^1, \ldots, u_{n_t}^m)$ the vector of countries' emissions at node $n_t \in \mathbf{n}_t$, $t \in \mathbb{T} \setminus \{T\}$, and by x_{n_t} the stock of pollution at this node. The evolution of this stock is governed by the following difference equation:

$$x_{n_t} = (1 - \delta_{n_t^-}) x_{n_t^-} + \sum_{i \in M} u_{n_t^-}^i, \qquad n_t \in \mathbf{n}_0^{++} \setminus \{n_0\}, \tag{21}$$

$$x_{n_0} = x_0, \tag{22}$$

where the initial state x_0 is given for the root node n_0, and δ_{n_t} $(0 < \delta_{n_t} < 1)$ is a stochastic rate of pollution absorption by Mother Nature at node n_t. We suppose that δ_{n_t} can take two possible values, that is, $\delta_{n_t} \in \left\{\underline{\delta}, \overline{\delta}\right\}$, with $\underline{\delta} < \overline{\delta}$.

Each country suffers an environmental damage cost due to pollution accumulation. We assume that this cost is an increasing convex function in the pollution stock and retain the quadratic form $D_i(x_{n_t}) = c_i x_{n_t}^2$, $i \in M$, where c_i is a strictly positive parameter.

[2] This example is considered in Chapter 7 in Parilina et al. (2022) with given transition probabilities.

Denote by $\rho \in (0,1)$ the discount factor. Player i's payoff along scenario $s = [n_0, n_T]$ is given by

$$J_i^s(\mathbf{x},\mathbf{u}) = \sum_{t=0}^{T-1} \rho^t \sum_{n_t \in \mathbf{n}_t \cap s} \left(R_i\left(u_{n_t}\right) - D_i(x_{n_t})\right) + \rho^T \Phi_i(n_T, x_{n_T}),$$

subject to (21), (22), and $u_{n_t}^i \geq 0$ for all $i \in M$ and any $n_t \in \mathbf{n}_t$, $t = 0, \ldots, T-1$. Suppose that the payoff function $\Phi_i(n_T, x_{n_T})$ of Player i at a terminal node n_T is given by

$$\Phi_i(n_T, x_{n_T}) = -g_i x_{n_T}.$$

Substituting for $\Phi_i(n_T, x_{n_T}), R_i\left(u_{n_t}\right)$, and $D_i(x_{n_t})$ by their values along scenario s, we get

$$J_i^s(\mathbf{x},\mathbf{u}) = \sum_{t=0}^{T-1} \rho^t \sum_{n_t \in \mathbf{n}_t \cap s} \left(\alpha u_{n_t}^i - \frac{1}{2}\left(u_{n_t}^i\right)^2 - c_i x_{n_t}^2\right) - \rho^T g_i x_{n_T}.$$

We use the following parameter values in the numerical illustration:

$$\begin{aligned}
&M = \{1,2,3\},\ \mathbb{T} = \{0,1,2,3\},\ x_0 = 0,\ \rho = 0.9,\\
&c_1 = 0.15,\ c_2 = 0.10,\ c_3 = 0.05,\ \underline{\delta} = 0.10,\ \overline{\delta} = 0.45,\\
&g_1 = 0.15,\ g_2 = 0.10,\ g_3 = 0.05,\ \alpha = 35.
\end{aligned}$$

Note that the players are symmetric in terms of revenues, but face different damage costs.

The event tree is depicted in Figure 1, where nodes n_1^2, n_2^2, n_2^4 correspond to the high level of pollution reduction $\overline{\delta} = 0.45$, and nodes n_0, n_1^1, n_2^1, n_2^3 correspond to the low level of pollution reduction $\underline{\delta} = 0.10$.

We apply Theorem 1 to find the solution of the joint optimization problem in robust feedback S-adapted strategies, i.e., $(\psi_i(n_t, x_{n_t}) : i \in M)$. To obtain this solution, we solve the problem starting from the terminal nodes backward until we reach the root node n_0. Similarly, we use Theorem 2 to obtain the worst-case Nash equilibrium. The emissions profiles in both modes of play are given in Figure 1, and the corresponding state values are provided in Table 1.

Table 1: Pollution stocks in NE and cooperative solution in robust feedback strategies.

Time period	$t=0$	$t=1$		$t=2$			
Node	n_0	n_1^1	n_1^2	n_2^1	n_2^2	n_2^3	n_2^4
Coop	0.0000	15.886	15.886	41.052	41.052	39.185	39.185
NE	0.0000	46.803	46.803	95.392	95.392	84.755	84.755

	$t=3$							
Node	n_3^1	n_3^2	n_3^3	n_3^4	n_3^5	n_3^6	n_3^7	n_3^8
Coop	141.137	141.137	126.769	126.769	139.406	139.406	125.711	125.711
NE	190.583	190.583	157.196	157.196	181.010	181.010	151.345	151.345

Figure 2 gives the individual profits at each node in both modes of play and the total discounted payoff for each scenario. These results call for the following comments:

1. The total emissions in the noncooperative game exceed their counterparts in the cooperative game. This result is fully expected as each country internalizes the impact of own emissions on the other countries in the cooperative case, whereas it behaves selfishly in the worst-case Nash equilibrium. As a consequence, the cumulative pollution is lower in the cooperative case (see Table 1).

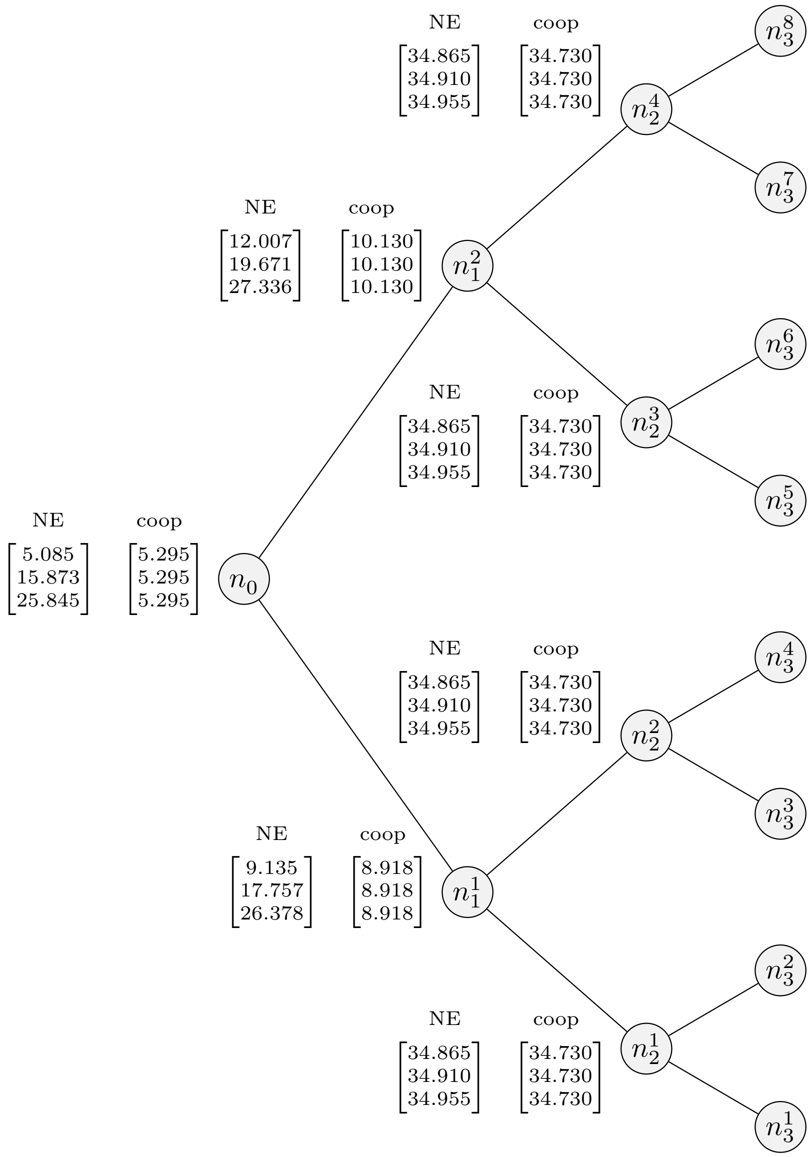

Figure 1: Emissions in NE and cooperative solution in robust feedback strategies.

2. In both solutions, the controls are the same for $t = 2$ because they are independent of the state and the parameter values are the same. This result is a by-product of the linear salvage value, which implies that the marginal value of the terminal stock of pollution is constant.
3. The total discounted profits of the three players is (as expected) higher when they jointly optimize their profits than when they act noncooperatively. Player 3 realizes a lower profit under cooperation, which requires the implementation of a side-payment mechanism to reach a Pareto-optimal solution, that is, an arrangement in which all players are better off under a cooperative mode of play. The total discounted payoffs along the eight scenarios are as follows:

Scenarios	Total coop payoffs	Total NE payoffs
$[n_0, n_3^7]$ and $[n_0, n_3^8]$	$2,353.341$	$1,567.134$
$[n_0, n_3^5]$ and $[n_0, n_3^6]$	$2,350.346$	$1,560.638$
$[n_0, n_3^3]$ and $[n_0, n_3^4]$	$2,232.252$	$1,000.977$
$[n_0, n_3^1]$ and $[n_0, n_3^2]$	$2,229.110$	995.892
$\frac{\text{Highest payoff}}{\text{Lowest payoff}}$	$\frac{2{,}353.341}{2{,}229.110} = 1.056$	$\frac{1{,}567.134}{995.892} = 1.574$

Whereas there is a relatively small variation in the sum of payoffs along the different scenarios in the cooperative case, there is a notable difference in the noncooperative case between the best and worst scenarios in terms of absorption of pollution.

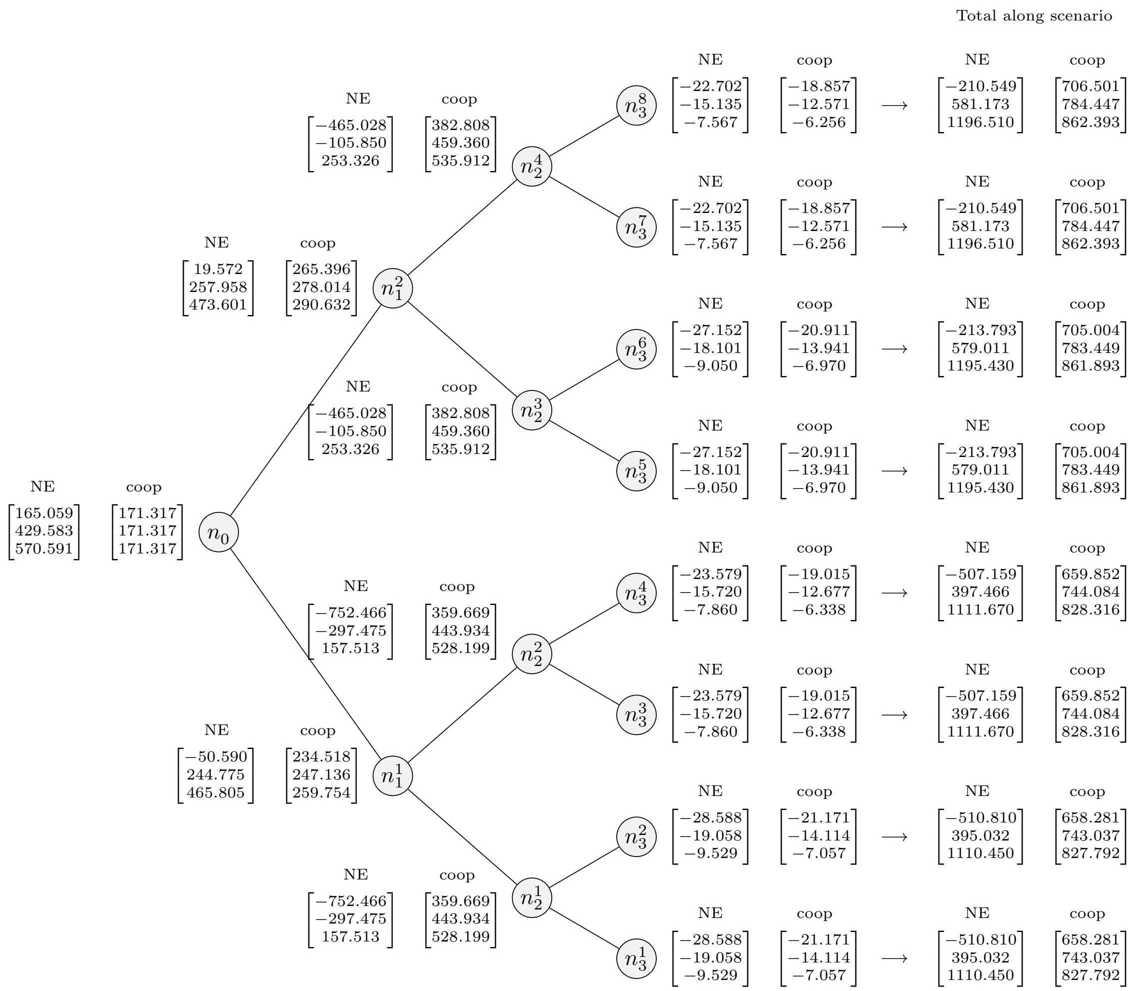

Figure 2: Profits in NE and cooperative solution in robust feedback strategies.

5 Conclusion

In this paper, we introduced the class of DGPETs with ambiguity and characterized cooperative and noncooperative robust solutions in which the players' strategies are adapted to the node and depend on the state variable. The following extensions are worth considering:

1. Under what circumstances, if any, it is in the best interest of a player to deviate from a robust strategy when all others implement such strategy? Simulation of the values of stochastic parameters could help in addressing this question.
2. Our numerical example shows that the cooperative solution is not Pareto optimal with respect to the noncooperative equilibrium. A relevant research question is then what type of side-payments could be implemented to remedy to this problem and lead to durable cooperation?
3. Alternatively to our solution approach, one could consider a minimax regret optimization methodology to solve a DGPET with ambiguity. The aim would be to obtain a solution (cooperative or noncooperative) that minimizes the maximum deviation, over all possible scenarios, between the payoff at the solution and the maximal payoff for the corresponding scenario.

References

Arend, R.J. (2020). Strategic decision-making under ambiguity: a new problem space and a proposed optimization approach. Business Research, 13, 1231–1251 https://doi.org/10.1007/s40685-020-00129-7

Bauso, D., Tembine, H., and Başar, T. (2016). Robust mean-field games. Dynamic Games and Applications, 6(3), 277–300.

Bauso, D., Timmer, J. (2012). On robustness and dynamics in (un)balanced coalitional games, Automatica, 48(10), 2592–2596.

Başar, T., Bernhard P. (2008). H^∞-Optimal Control and Related Minimax Design Problems. A Dynamic Game Approach. Second Edition. Berlin, Springer.

Calafiore, G. C., Campi, M. C. (2006). The scenario approach to robust control design, IEEE Transactions on Automatic Control, 51(5), 742–753.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. Quarterly Journal of Economics 75(4), 643–669.

Engwerda, J. C. (2017). Robust open-loop Nash equilibria in the noncooperative LQ game revisited. Optimal Control Applications and Methods, 38, 795-813.

Engwerda, J. (2022). Min-Max Robust Control in LQ-Differential Games. Dynamic Games Applications, 12, 1221–1279. https://doi.org/10.1007/s13235-021-00421-z

Epstein, L. G., Schneider, M. (2007). Learning under ambiguity. The Review of Economic Studies, 74(4), 1275–1303.

Genc, T.S., Reynolds, S.S., and Sen, S. (2007). Dynamic oligopolistic games under uncertainty: A stochastic programming approach, Journal of Economic Dynamics and Control, 31(1), 55-80.

Haurie, A., Moresino, F. (2001). Computation of S-adapted equilibria in piecewise deterministic games via stochastic programming methods. In: Altman, E., Pourtallier, O. (eds) Advances in Dynamic Games and Applications. Annals of the International Society of Dynamic Games, vol 6. Birkhäuser, Boston, MA. https://doi.org/10.1007/978-1-4612-0155-7_14

Haurie, A., Moresino, F. (2002). S-adapted oligopoly equilibria and approximations in stochastic variational inequalities. Annals of Operations Research, 114, 183–201. https://doi.org/10.1023/A:1021018421126

Haurie, A., Zaccour, G., and Smeers, Y. (1990). Stochastic equilibrium programming for dynamic oligopolistic markets. Journal of Optimization Theory and Applications, 66, 243–253.

Høyland, K., Wallace, S. W. (2001). Generating scenario trees for multistage decision problems. Management Science, 47(2), 295–307. http://www.jstor.org/stable/2661576

Iyengar, G. N. (2005). Robust dynamic programming. Mathematics of Operations Research, 30(2), 257–280.

Jank, G., Kun, G. (2002). Optimal control of disturbed linear-quadratic differential games. European Journal of Control, 8(2), 152-162.

Kaut, M. (2021). Scenario generation by selection from historical data. Computional Management Science, 18, 411–429 . https://doi.org/10.1007/s10287-021-00399-4

Nilim, A., El Ghaoui, L. (2005). Robust control of Markov decision processes with uncertain transition matrices. Operations Research, 53(5), 780–798.

Parilina, E., Reddy, P.V., and Zaccour, G. (2022). Theory and Applications of Dynamic Games. A Course on Noncooperative and Cooperative Games Played over Event Trees. Springer Cham.

Parilina, E., Sedakov, A., and Zaccour, G. (2017). Price of anarchy in a linear-state stochastic dynamic game. European Journal of Operational Research, 258, 790–800.

Parilina, E., Zaccour, G. (2015a). Node-consistent core for games played over event trees. Automatica, 53, 304–311.

Parilina, E., Zaccour, G. (2015b). Approximated cooperative equilibria for games played over event trees. Operations Research Letters, 43(5), 507-513.

Parilina, E., Zaccour, G. (2017). Node-consistent Shapley value for games played over event trees with random terminal time. Journal of Optimization Theory and Applications, 175(1), 236–254.

Parilina, E.M., Zaccour, G. (2022). Sustainable cooperation in dynamic games on event trees with players' asymmetric beliefs. Journal of Optimization Theory and Applications, 194(1), 92–120.

Pineau, P.-O., Murto, P. (2003). An oligopolistic investment model of the Finnish electricity market. Annals of Operations Research, 121, 123–148.

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.

Reddy, P.V., Shevkoplyas, E., and Zaccour, G. (2013). Time-consistent Shapley value for games played over event trees. Automatica, 49(6), 1521–1527.

Reddy, P.V., Zaccour, G. (2019). Open-loop and feedback Nash equilibria in constrained linear-quadratic dynamic games played over event trees. Automatica, 107, 162–174.

Zaccour, G. (1987). Théorie des jeux et marchés énergétiques: marché européen de gaz naturel et échanges d'électricité, Ph.D. Thesis, HEC Montréal.