Les Cahiers du GERAD ISSN: 0711–2440

Stability and negotiation of long-term agreements in cooperative difference games with nontransferable utility

E. M. Parilina, P. V. Reddy, G. Zaccour

G–2020–77

December 2020

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, 2020
– Bibliothèque et Archives Canada, 2020

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, 2020
– Library and Archives Canada, 2020

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

Stability and negotiation of long-term agreements in cooperative difference games with nontransferable utility

Elena M. Parilina [a]

Puduru Viswanadha Reddy [b]

Georges Zaccour [c]

[a] *Saint Petersburg State University, Saint Petersburg 199034, Russia*

[b] *Indian Institute of Technology Madras, Chennai, India*

[c] *GERAD and Department of Decision Sciences, HEC Montréal, Montréal, Canada, H3T 2A7*

e.parilina@spbu.ru
vishwa@ee.iitm.ac.in
georges.zaccour@gerad.ca

December 2020
Les Cahiers du GERAD
G–2020–77

Abstract: In this paper, we study the time consistency of cooperative agreements in dynamic games with nontransferable utility. An agreement designed at the outset of a game is time consistent (or sustainable) if it remains in place for the entire duration of the game, that is, if no player would benefit from switching to his Nash equilibrium strategy. The literature has highlighted that, since side payments are not allowed, the design of such an agreement is very challenging. To address this issue, we introduce different notions for the temporal stability of an agreement and characterize the agreement's intrinsic longevity. We illustrate our general results with a linear-quadratic difference game and show that an agreement's longevity can be easily assessed using the problem data. We also study the effect of information structure on the longevity of the agreement. We illustrate our results with a numerical example.

Keywords: Time consistency, cooperative dynamic games, nontransferable utility, linear-quadratic games

Résumé : Dans cet article, nous étudions la question de la cohérence temporelle des accords de coopération dans les jeux dynamiques à utilité non transférable. Un accord conçu au début d'un jeu est cohérent dans le temps (ou durable) s'il reste en place pendant toute la durée du jeu, c'est-à-dire si aucun joueur ne bénéficierait d'un passage à sa stratégie non coopérative. Comme les paiements latéraux ne sont pas autorisés, la conception d'un tel accord est très difficile. Pour résoudre ce problème, nous introduisons différentes notions de stabilité temporelle d'un accord et caractérisons la longévité intrinsèque de l'accord. Nous illustrons nos résultats généraux avec un jeu linéaire-quadratique et montrons que la longévité d'un accord peut être facilement évaluée en utilisant les données du problème.

Mots clés : Cohérence dynamique, jeux dynamiques coopératifs, utilité non transférable, jeux linéaires-quadratiques

Acknowledgements: The first author's research is supported by Saint Petersburg State University, grant no. 41128961. This research was partially conducted during the research stay of the first author at GERAD, HEC Montreal, Canada. The second author's research is supported by the IIT Madras, New Faculty Initiative Grant ELE1718849NFIGPUVR. The third author's research is supported by NSERC Canada, grant RGPIN-2016-04975. He also thanks the Department of Electrical Engineering, IIT Madras, India, for its hospitality.

1 Introduction

In many instances, agents (e.g., firms, spouses, countries) opt for long-term contracts instead of short-term ones for essentially two main reasons. First, a long-term agreement avoids the possibly high renegotiation cost (in dollars and time). Second, if today's decisions influence tomorrow's payoff, then short-term optimization and cost-benefit analysis lead to bad decisions. To illustrate, consider climate change. If the players only account for the impact of pollution on today's health and wealth, they will most likely only postpone painful but (ultimately) unavoidable decisions, which will exacerbate the problems in the longer term, leading to an eventually higher damage cost.

One main concern with long-term agreements is their sustainability, that is, how to ensure that the concerned parties will honor their commitments as time goes by. It is an empirical fact that some long-term contracts break down before their maturity. A few examples of such contract breaches include the high observed rate of divorce; Canada leaving the Kyoto Protocol; and the US, the Paris Agreement.

The literature on dynamic games has addressed this issue following two approaches. The first aims at embedding the cooperative solution with an equilibrium property that renders the agreement stable by construction. The early contributions in (state-space) dynamic games include [12], [24], [11], and [9]. The books [4] and [10] provide a comprehensive introduction to cooperative equilibria in differential games. The second stream, to which this paper belongs, seeks to build a time-consistent agreement, that is, one in which, at each instant of time, each player finds it optimal to continue with the agreement rather than switching to his noncooperative strategy. The concept of time consistency in differential games was proposed by [20].

In dynamic games with transferable utility (TU), one can use intertemporal payment transfers to implement a time-consistent agreement; see [29], [30], [14] and [21] for comprehensive reviews. In dynamic games with non-transferable utility (NTU), such transfers are not allowed, which makes the design of time-consistent agreements a formidable challenge, and explains the sparsity of the literature. Leitmann [17], Dockner and Jørgensen [3], Hämäläinen et al. [7], Yeung and Petrosyan [26], Yeung et al. [25], de-Paz et al. [2], and Marín-Solano [18] studied some cooperative differential games with non-transferable payoffs. Haurie [8] analyzed the time-consistency property of the Nash bargaining solution in NTU cooperative differential games. Sorger [23] and Yeung and Petrosyan [27] studied these games in a discrete-time setting.

The objective of this paper is to study the existence of time-consistent individually rational (TCIR) cooperative agreements in discrete-time dynamic games with non-transferable utility. A series of cooperative game solutions and bargaining procedures have been developed to deal with NTU games, e.g., the Nash bargaining solution [19], the Kalai-Smorodinsky bargaining solution [16], the Kalai proportional solution [15], and the core by Edgeworth [5]. These bargaining solutions have two properties in common, namely, individual rationality (i.e., no player would accept an agreement that leaves him worse than staying out of it) and Pareto optimality. In a static game, all the Pareto solutions can be generated by solving a weighted sum optimization problem, and in a dynamic game, by solving a weighted sum optimal-control problem [6] with the weight vectors belonging to a unit simplex. Then, a cooperative contract signed at an initial date means, in particular, that the players have selected and agreed on a specific weight vector. In [26], the authors studied the existence of TCIR agreements in a class of NTU stochastic differential games. One conclusion is that such agreements may not exist because the time-consistency requirements are too restrictive (see also [28]). If one allows for variations in the weight vector over time, then the task becomes easier. However, a change in the weight vector means, tautologically, that the initial agreement is not time consistent. This approach of varying the players' weights over time was studied in [23] and [18], while [27] provides a dynamic programming based procedure for finding subgame-consistent cooperative solutions in discrete-time NTU games with varying weights. For a review of dynamic cooperative NTU games, see [28].

In this paper, considering a fairly general class of discrete-time NTU games, we introduce four different notions of inter-temporal stability related to the time consistency of an agreement. We show that the space of individually rational agreements can be canonically decomposed into a union of disjoint sets of TCIR agreements that break down before their maturity date and those that persist through the full duration of the game. This result implies that the players can assess, from the outset, the intrinsic longevity of any individually rational agreement that could be signed at an initial date. Further, we show that, for the class of linear-quadratic difference games, these sets can be easily computed from the problem data merely by testing for the positive semi-definiteness of the matrices. Interestingly, for any of the classical fairness based bargaining solutions, e.g., Nash, Kalai-Smorodinsky, and egalitarian solutions, we can determine if the agreement will remain in place until the end of the game, and if not, when it will break down. Another practical implication is that if designing a TCIR contract having the same horizon as the game itself is out of reach, then the players can settle for the longest feasible time span. After, they can renegotiate a new agreement, that is, adopt another vector of weights, as in, e.g., [18], [23], and [27].

The paper is organized as follows. In Section 2, we introduce the cooperative dynamic game model and review the concept of Pareto optimality and provide results on the necessary and sufficient conditions for the existence of Pareto-optimal solutions in discrete-time dynamic games. In Section 3, we introduce four different notions of stability of cooperative agreements in the time-consistency sense. We provide a canonical decomposition of individually rational Pareto solutions. In Section 4, we specialize these results for the class of linear-quadratic difference games to obtain a semi-analytic procedure for constructing the sets corresponding to time-consistent individually rational agreements when players use open-loop and feedback information structures. In Section 5, we provide an example to illustrate our results, and we conclude in Section 6.

1.1 Notation

We shall use the following notation. The n-dimensional Euclidean space is denoted by $\mathbb{R}^n$, $n \geq 1$. A' denotes the transpose of a matrix A. $A_1 \oplus A_2 \oplus \cdots \oplus A_n$ represents the block diagonal matrix obtained by taking the matrices $A_1, A_2, \cdots, A_n$ as diagonal elements in this sequence. The $n \times n$ identity matrix is represented by I, and the ith column of I is denoted by the vector $\mathbf{e}_i \in \mathbb{R}^n$. We denote a positive semi-definite (definite) matrix A as $A \succeq 0$ ($\succ 0$). We denote the quadratic term $x'Ax$ as $||x||^2_A$, where $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$ is a symmetric matrix. The n-dimensional unit simplex is denoted by $\mathcal{P}$, which is defined as $\mathcal{P} := \{(\alpha^1, \alpha^2, \cdots, \alpha^n) \in \mathbb{R}^n \bigm| \alpha^i \in (0,1),\ i = 1,2,\cdots,n,\ \sum_{i=1}^n \alpha^i = 1\}$. The symbol $\boldsymbol{\alpha}$ denotes the weight vector $(\alpha^1, \alpha^2, \cdots, \alpha^n)$ in $\mathcal{P}$.

2 Discrete-time cooperative NTU games and Pareto-optimality

Consider a multi-stage finite-horizon nonzero-sum discrete-time dynamic game with T stages. We denote by $\mathcal{N} = \{1, 2, \cdots, N\}$ the set of players and by $\mathcal{T} = \{0, 1, 2, \cdots, T\}$ the set of decision instants. The evolution of the state is governed by the following difference equation:

$$x_{t+1} = f_t(x_t, u^1_t, u^2_t, \cdots, u^N_t),\ x_0 \text{ is given}, \tag{1}$$

where $x_t \in \mathbb{R}^n$ represents the state of the system and $u^i_t \in U^i_t \subset \mathbb{R}^{m_i}$ is the control action of Player i at time period t in m_i-dimensional control action set. Denote by $\mathbf{u}_t := (u^1_t,\ u^2_t, \cdots, u^N_t)$ the joint action of the players at time t and $\mathbf{U}_t := \Pi^N_{i=1} U^i_t \subset \mathbb{R}^m,\ m = \sum_i m_i$, the joint action set of the players. Denote by $\tilde{u}^i := (u^i_0, u^i_1, \cdots, u^i_{T-1})$ a strategy, i.e., a profile of actions, of Player i, and by $\tilde{\mathbf{u}} := (\tilde{u}^1, \tilde{u}^2, \cdots, \tilde{u}^N)$ a joint strategy of the players. Let $\tilde{U}^i = U^i_0 \times U^i_1 \times \cdots \times U^i_{T-1}$ be the control set of Player i and $\mathcal{U} = \tilde{U}^1 \times \cdots \times \tilde{U}^N$ be the joint strategy set. Each player $i \in \mathcal{N}$ aims at minimizing the objective

$$J^i(\tilde{\mathbf{u}}) = h^i(x_T) + \sum_{t=0}^{T-1} g^i_t(x_t, \mathbf{u}_t), \tag{2}$$

where $h^i(x_T)$ is Player i's salvage value at T and $g^i_t(x_t, \mathbf{u}_t)$ is the running cost at t. We assume that the vector function $f_t : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, the scalar functions $g^i_t : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, and $h^i : \mathbb{R}^n \to \mathbb{R}$ are continuous for all $t \in \mathcal{T}$ and for all $i \in \mathcal{N}$, and that the partial derivatives of these functions exist and are continuous in their arguments.

Remark 1 *A complete description of a dynamic game includes the specification of the information structure, that is, the piece of information used by the players when they make their decisions. As our construction* *is independent of this structure, and to avoid usage of additional notation, we postpone the discussion on information structure to Section 4.*

We consider the situation where the players are willing to cooperate to improve their objectives. If the utility functions (2) are transferable, then the players collectively solve a joint sum optimal-control problem. In our paper, the utility functions are non-transferable, and if the players agree to cooperate, then they will seek a Pareto-optimal agreement defined as a Pareto-optimal strategy profile. A Pareto-optimal solution has the property that the cost incurred by any single player cannot be further reduced without increasing the cost of at least one other player. Put differently, a solution that cannot be improved upon by all the players simultaneously is a Pareto (optimal) solution.

Definition 1 *(Pareto solution) The strategy $\tilde{\mathbf{u}}^* \in \mathcal{U}$ is said to be Pareto efficient if there does not exist $\tilde{\mathbf{u}} \in \mathcal{U}$ such that for all $i \in \mathcal{N}$, $J_i(\tilde{\mathbf{u}}) \leq J_i(\tilde{\mathbf{u}}^*)$ and for some $j \in \mathcal{N}$ $J_i(\tilde{\mathbf{u}}) < J_i(\tilde{\mathbf{u}}^*)$. The corresponding point $(J^1(\tilde{\mathbf{u}}^*), J^2(\tilde{\mathbf{u}}^*), \cdots, J^N(\tilde{\mathbf{u}}^*))$ is called a Pareto solution. The set of all Pareto solutions is called a Pareto frontier.*

It is well-known that the problem of finding Pareto solutions is closely related to solving a weighted sum optimal-control problem with weights belonging to the unit simplex.

Lemma 1 *(Leitmann [17]) Let $\boldsymbol{\alpha} \in \mathcal{P}$ and assume that $\tilde{\mathbf{u}}^* \in \mathcal{U}$ is such that*

$$\tilde{\mathbf{u}}^* = \arg\min_{\tilde{\mathbf{u}} \in \mathcal{U}} \sum_{i \in \mathcal{N}} \alpha_i J^i(\tilde{\mathbf{u}}),$$

then $\tilde{\mathbf{u}}^*$ is Pareto-optimal.

Being a sufficient condition, from the above lemma it is not clear whether we obtain all Pareto-optimal controls solving a weighted sum optimization problem. The following lemma provides both a necessary and sufficient characterization of Pareto solutions. It states that Pareto-optimal solutions can be obtained by solving N constrained optimization problems; see also [22, Chapter 22].

Lemma 2 *The joint strategy $\tilde{\mathbf{u}}^* \in \mathcal{U}$ is Pareto optimal, if, and only if, for all i, $\tilde{\mathbf{u}}^*$ minimizes $J^i(\tilde{\mathbf{u}})$ on the constrained set*

$$\mathcal{U}^i := \left\{ \tilde{\mathbf{u}} \in \mathcal{U} \mid J^j(\tilde{\mathbf{u}}) \leq J^j(\tilde{\mathbf{u}}^*),\ \forall j \in i^- \right\}, \tag{3}$$

where $i^- = \mathcal{N} \backslash i$.

Proof. Assume that $\tilde{\mathbf{u}}^*$ is Pareto-optimal. This implies $\tilde{\mathbf{u}}^* \in \mathcal{U}^i$ for all $i \in \mathcal{N}$, so $\mathcal{U}^i \neq \emptyset$. If $\tilde{\mathbf{u}}^*$ does not minimize $J^i(\tilde{\mathbf{u}})$ on the constrained set $\mathcal{U}^i$ for some i, then there exists $\tilde{\mathbf{v}}$ such that $J^j(\tilde{\mathbf{v}}) \leq J^j(\tilde{\mathbf{u}}^*)$ for all $j \neq i$ and $J^i(\tilde{\mathbf{v}}) < J^i(\tilde{\mathbf{u}}^*)$. Clearly, $\tilde{\mathbf{v}}$ dominates $\tilde{\mathbf{u}}^*$, which contradicts the Pareto-optimality of $\tilde{\mathbf{u}}^*$.

Suppose $\tilde{\mathbf{u}}^*$ minimizes every $J^i(\tilde{\mathbf{u}})$ on $\mathcal{U}^i$ and assume that $\tilde{\mathbf{u}}^*$ is not Pareto-optimal. Then, there exists $\tilde{\mathbf{v}}$ and a player index j such that $J^k(\tilde{\mathbf{v}}) \leq J^k(\tilde{\mathbf{u}}^*)$ for all $k \neq j$ and $J^j(\tilde{\mathbf{v}}) \leq J^j(\tilde{\mathbf{u}}^*)$. This is contradictory to the minimality of $\tilde{\mathbf{u}}^*$ on $\mathcal{U}^j$. □

We observe that, for a fixed player i the constraint set $\mathcal{U}^i$ defined in (3) depends on the entries of the Pareto-optimal value that represents the loss of all players in i^-. Therefore, all Pareto solutions can be obtained by solving N constrained optimal-control problems. Using the above lemma, a necessary condition for the existence of Pareto-optimal solutions for the dynamic game defined by (1–2) can be obtained. In particular, it can be shown that corresponding to every Pareto-optimal control $\tilde{\mathbf{u}}^*$ there exists a weight vector $\boldsymbol{\alpha} \in \mathcal{P}$. In other words, it is possible to obtain all the candidate Pareto solutions by solving the set of necessary conditions resulting from a related weighted sum optimal-control problem. In the theorem below we formally state this result without a proof as the theorem is a discrete time analog of the continuous time version provided by Engwerda [6, Theorem 2.7].

Theorem 1 *Assume* $(J^1(\tilde{\mathbf{u}}^*), J^2(\tilde{\mathbf{u}}^*), \cdots, J^N(\tilde{\mathbf{u}}^*))$ *is a Pareto solution for problem (1–2). Then, there exist an* $\boldsymbol{\alpha} \in \mathcal{P}$*, which is a unit simplex on* $\mathcal{N}$*, and costate variables* $\lambda_t \in \mathbb{R}^n$*, such that with Hamiltonian* $H_t(x_t, \mathbf{u}_t, \lambda_{t+1}) = \sum_{i=1}^N \alpha^i g_t^i(x_t, \mathbf{u}_t) + \lambda'_{t+1} f_t(x_t, \mathbf{u}_t)$*,* $\tilde{\mathbf{u}}^*$ *satisfies*

$$H_t(x_t^*, \mathbf{u}_t^*, \lambda_{t+1}) \leq H_t(x_t^*, \mathbf{u}_t, \lambda_{t+1}), \ \forall \mathbf{u}_t \in \mathbf{U}_t \tag{4a}$$

$$\lambda_t = \left(\frac{\partial f_t}{\partial x_t}\right)' \lambda_{t+1} + \sum_i \ \alpha^i \ \frac{\partial g_t^i}{\partial x_t}, \ \lambda_T = \frac{\partial \left(\sum_{i=1}^N \alpha^i \ h^i\right)}{\partial x_T} \tag{4b}$$

$$x_{t+1}^* = f_t(x_t^*, \mathbf{u}_t^*), \ x_0^* = x_0. \tag{4c}$$

Next, we provide conditions under which the necessary conditions for Pareto optimality given by (4) are also sufficient, that is, conditions under which the solutions of (4) will be Pareto efficient.

Theorem 2 *Assume that there exist* $\alpha^i \in (0,1)$ *with* $\sum_{i=1}^N \alpha^i = 1$*, and the collection of vectors* $\{\lambda_t^*, \mathbf{u}_t^*, x_t^*\}$*, that satisfy (4). Assume that Hamiltonian* $H_t(x_t, \mathbf{u}_t, \lambda_{t+1}) = \sum_i \alpha^i g^i(x_t, u_t) + \lambda'_{t+1} f_t(x_t, \mathbf{u}_t))$ *has a minimum with respect to* $\mathbf{u}_t$ *for all* $t \in \mathcal{T} \backslash \{T\}$*. Let the minimized Hamiltonian be given by* $H_t^*(x_t, \lambda_{t+1}) = \min_{\mathbf{u}_t} H_t(x_t, \ \mathbf{u}_t, \lambda_{t+1})$*. Then, If* $H_t^*(x_t, \lambda_{t+1})$ *is convex with respect to* x_t *for all* $t \in \mathcal{T} \backslash \{T\}$ *and* $h(x_T) = \sum_{i=1}^N \alpha^i h^i(x_T)$ *is convex with respect to* x_T*, then* $\tilde{\mathbf{u}}^*$ *is Pareto efficient.*

Proof. For any $\tilde{\mathbf{u}} \in \mathcal{U}$, we consider the difference

$$\begin{aligned}\sum_i \alpha^i J^i(\tilde{\mathbf{u}}) - \sum_i \alpha^i J^i(\tilde{\mathbf{u}}^*) &= \sum_i \alpha^i \left(h^i(x_T) - h^i(x_T^*)\right) + \sum_{t=0}^{T-1} \sum_i \alpha^i \left(g_t^i(x_t, \mathbf{u}_t) - g_t^i(x_t^*, \mathbf{u}_t^*)\right) \\ &= \sum_{t=1}^{T-1} \left(H_t(x_t, \mathbf{u}_t, \lambda_{t+1}^*) - H_t(x_t^*, \mathbf{u}_t^*, \lambda_{t+1}^*) - \lambda_t^{*\prime} (x_t - x_t^*)\right) \\ &\quad + H_0(x_0, \mathbf{u}_0, \lambda_1^*) - H_0(x_0^*, \mathbf{u}_0^*, \lambda_1^*) + h(x_T) - h(x_T^*) - \lambda_T^{*\prime} (x_T - x_T^*) .\end{aligned} \tag{5}$$

From the convexity of the minimized Hamiltonian and from the necessary condition (4b) we have

$$\begin{aligned}H_t(x_t, \mathbf{u}_t, \lambda_{t+1}^*) - H_t(x_t^*, \mathbf{u}_t^*, \lambda_{t+1}^*) &\geq H_t^*(x_t, \lambda_{t+1}^*) - H_t^*(x_t^*, \lambda_{t+1}^*) \\ &\geq \frac{\partial H_t}{\partial x_t}' (x_t^*, \lambda_{t+1}^*)(x_t - x_t^*) = \lambda_t^{*\prime}(x_t - x_t^*).\end{aligned} \tag{6}$$

Again from the convexity of the function $h(.)$, and from (4b) we have

$$h(x_T) - h(x_T^*) \geq \frac{\partial h}{\partial x_T}' (x_T^*)(x_T - x_T^*) = \lambda_T^{*\prime}(x_T - x_T^*). \tag{7}$$

Using (6) and (7) in (5), and from $x_0 = x_0^*$, we get

$$\sum_i \alpha_i J^i(\tilde{\mathbf{u}}) - \sum_i \alpha_i J^i(\tilde{\mathbf{u}}^*) \geq H_0(x_0, \mathbf{u}_0, \lambda_1^*) - H_0(x_0, \mathbf{u}_0^*, \lambda_1^*) \geq 0.$$

So, $\tilde{\mathbf{u}}^*$ minimizes the objective $\sum_i \alpha_i J^i(\tilde{\mathbf{u}}^*)$, and this implies from Lemma 1 that $\tilde{\mathbf{u}}^*$ is Pareto efficient. □

Remark 2 *From Theorem 1 and Theorem 2 it is clear that all the Pareto solutions associated with the NTU game described by (1)–(2) can be obtained, under a few convexity assumptions, by solving a weighted sum optimal-control problem, with weight vectors belonging to* $\mathcal{P}$*. However, in general, it is not clear if there exists a one-to-one relationship between the set of weight vectors* $\mathcal{P}$ *and the set of all Pareto solutions.*

3 Time-consistent individually rational (TCIR) Pareto-optimal solutions

In a non-transferable utility setting, we have seen in the previous section that when the players agree to cooperate they seek to obtain Pareto-optimal solutions. Further, we have shown that all the Pareto solutions can be obtained when the players jointly solve a weighted sum optimal-control problem with the weights belonging to the set $\mathcal{P}$. The selection of a weight vector at the outset of the game reflects the players' individual bargaining power and is interpreted as the agreement made by the players,[1] and understood to remain in place for the duration of the game. However, as the selected Pareto solution is not in general an equilibrium, there is no mechanism embedded in the agreement that ensures its sustainability over time; we rule out the non-credible idea that it suffices to state that the agreement is binding to avoid it breaking down before maturity. The issue of the durability of cooperation is the topic of this section, and main contribution of this paper. In particular, we address the following two questions related to the sustainability of long-term agreements when side-payments are not allowed:

1. In the set of Pareto solutions, are there agreements, which are negotiated at the outset of the game, that will remain in place for the entire duration of the game?
2. Is there a way to measure the intrinsic longevity of an agreement? In other words, what is the maximum duration of the agreement, which is made at the outset of the game, that can sustain without breaking down before the maturity date?

Remark 3 *In [27], the authors consider situations where agreements, which are made at the outset of the game, break down before maturity date, and provide a re-negotiation mechanism using a dynamic programming approach. As the agreements correspond to a weight vector in the set $\mathcal{P}$ their work entails varying weights at a later stage of the game. Our work differs from [27] as our focus is towards measuring the intrinsic logetivity of a long-term agreement.*

Whether the players honor a cooperative agreement or not depends on what players will do once the agreement breaks down. More importantly, whether the players agree to sign a long-term contract also depends on their bargaining strengths, that is, what they can achieve on their own either individually or through formation of coalitions. The following is a standing assumption on the behavior of the players, which will be used throughout the paper.

Assumption 1 *If the agreement breaks down at any state t, then each player $i \in \mathcal{N}$ individually minimizes (without forming coalitions) his objective in the subgame starting at t.*

Let $\tilde{\mathbf{u}}^\alpha$ be the Pareto-optimal control corresponding to the weight vector $\boldsymbol{\alpha} \in \mathcal{P}$. Let $\tilde{x}^\alpha$ be the state trajectory generated by the Pareto-optimal control $\tilde{\mathbf{u}}^\alpha$. Suppose that the players have been following the Pareto-optimal control trajectory $\tilde{\mathbf{u}}^\alpha$ until stage $t-1$, and let x_t^α be the state value at stage t. At stage t they may reconsider sustainability of a Pareto-optimal agreement by comparing the cost-to-go in the subgame starting at t with state x_t^α with the non-cooperative outcomes in this subgame. More formally, in the sugame starting at stage t the player i's objective is given by

$$sJ^i(\tilde{\mathbf{u}}|_t) = h^i(x_T) + \sum_{\tau=t}^{T-1} g_\tau^i(x_\tau, \mathbf{u}_\tau),$$

and the state trajectory evolves according to

$$x_{\tau+1} = f_\tau(x_\tau, \mathbf{u}_\tau),\ x_t = x_t^\alpha,\ \tau = t, t+1, \cdots, T-1,$$

where $\tilde{\mathbf{u}}|_t := \{\mathbf{u}_t, \mathbf{u}_{t+1}, \cdots, \mathbf{u}_{T-1}\}$ denote the strategies of the players in the subgame starting at stage t. We denote by $\tilde{u}^i|_t$ and $\tilde{u}^{i-}|_t$ the strategies used by player i and the players in i^- in the subgame, respectively.

Following Assumption 1 when the players reevaluate the Pareto-optimal agreement at any stage t they compare their individual Pareto cost-to-go to the non-cooperative cost-to-go in the subgame starting from

[1] that is, an outcome of the game when players can communicate.

stage t. Let $\tilde{\mathbf{u}}^{\diamond}|_t$ denote the control strategies used by the players in the non-coopeartive subgame game starting at time t when the agreement breaks down at time t. Let $\{x_m^{\diamond},\ m = t, t+1, \cdots, T\}$ denote the corresponding state trajectory. If every player finds the Pareto-optimal cost lower than their own noncooperative cost in the subgame starting at (t, x_t^{α}), then no player will deviate from the agreement $\boldsymbol{\alpha} \in \mathcal{P}$ at stage t. The cost incurred by Player i when the players use the Pareto-optimal control $\tilde{\mathbf{u}}^{\alpha}$ in the subgame starting at (t, x_t^{α}) is given by

$$W_c^i(t, x_t^{\alpha}) := h^i(x_T^{\alpha}) + \sum_{m=t}^{T-1} g_m^i(x_m^{\alpha}, \mathbf{u}_m^{\alpha}), \tag{8}$$

where the subscript c stands for cooperation. If the game is played noncooperatively in the subgame starting from (t, x_t^{α}), then Player i receives his Nash-equilibrium outcome, that is,

$$W_{nc}^i(t, x_t^{\alpha}) := h^i(x_T^{\diamond}) + \sum_{m=t}^{T-1} g_m^i(x_m^{\diamond}, \mathbf{u}_m^{\diamond}), \tag{9}$$

where the subscript nc stands for noncooperation. Using the cost-to-go functions (8) and (9), we introduce the following four notions of inter-temporal stability of an agreement.

Definition 2 *A Pareto solution corresponding to a weight vector $\boldsymbol{\alpha} \in \mathcal{P}$ is individually rational, if the following condition holds true for every $i \in \mathcal{N}$ and for a given $x_0 \in \mathbb{R}^n$:*

$$W_c^i(0, x_0) \leq W_{nc}^i(0, x_0). \tag{10}$$

We denote by $\mathbf{I}$ the set of all individually rational weight vectors, that is,

$$\mathbf{I} := \left\{ \boldsymbol{\alpha} \in \mathcal{P} \;\middle|\; W_c^i(0, x_0) \leq W_{nc}^i(0, x_0),\ \forall i \in \mathcal{N} \right\}. \tag{11}$$

An individually rational Pareto solution corresponding to the weight vector $\boldsymbol{\alpha} \in \mathcal{P}$ ensures that every player receives a cost lower than the noncooperative cost in the dynamic game starting at $(0, x_0)$. The condition in (10) can be interpreted as a necessary condition for cooperation to take place. If initially the Pareto-optimal cost-to-go is not lower than its noncooperative counterpart, then there is no reason to enter into an agreement. Note that the condition implicitly assumes that the players will indeed implement the Pareto-optimal control $\tilde{\mathbf{u}}^{\alpha}$ throughout the duration of the game.

Definition 3 *A Pareto solution corresponding to a weight vector $\boldsymbol{\alpha} \in \mathcal{P}$ is individually rational at stage l $(0 \leq l \leq T)$, if the following condition holds true for every $i \in \mathcal{N}$ and for a given $x_0 \in \mathbb{R}^n$:*

$$W_c^i(l, x_l^{\alpha}) \leq W_{nc}^i(l, x_l^{\alpha}). \tag{12}$$

We denote by $\mathbf{I}_l$ the set of all individually rational weight vectors at stage l, that is,

$$\mathbf{I}_l := \left\{ \boldsymbol{\alpha} \in \mathcal{P} \;\middle|\; W_c^i(l, x_l^{\alpha}) \leq W_{nc}^i(l, x_l^{\alpha}),\ \forall i \in \mathcal{N} \right\}. \tag{13}$$

Suppose that the game has been played cooperatively from the outset until stage $l-1$. This means that the players agreed on a weight vector $\alpha \in \mathcal{P}$ at stage 0 and that the state value at stage l is x_l^{α}. If at stage l the cooperative cost-to-go to Player i, for all $i \in \mathcal{N}$, is at most equal to his noncooperative payoff-to-go, then it is individually rational to cooperatively play the subgame starting in position (l, x_l^{α}).

Definition 4 *A Pareto solution corresponding to a weight vector $\boldsymbol{\alpha} \in \mathcal{P}$ is stage l $(1 \leq l \leq T)$ individually rational for a given $x_0 \in \mathbb{R}^n$ if the following condition holds true:*

$$W_c^i(t, x_t^{\alpha}) \leq W_{nc}^i(t, x_t^{\alpha}) \text{ for all } i \in \mathcal{N},\ 0 \leq t \leq l-1, \tag{14a}$$

$$W_c^i(l, x_l^\alpha) > W_{nc}^i(l, x_l^\alpha) \text{ for at least one } i \in \mathcal{N}. \tag{14b}$$

We denote by $\mathbf{TI}_l$ *the set of all time-consistent individually rational weight vectors until stage* l*, that is,*

$$\mathbf{TI}_l := \left\{ \boldsymbol{\alpha} \in \mathcal{P} \;\middle|\; (14) \text{ holds true} \right\}. \tag{15}$$

The above definition characterizes the conditions under which a cooperative agreement, which is individually rational until stage $l-1$, breaks down at stage l. The number $l-1$ corresponds to the longest individually rational Pareto agreement that can be sustained with the weight vector $\boldsymbol{\alpha}$ chosen at the start of the game. The set $\mathbf{TI}_l$ $(1 \le l \le T)$ defines the long-term agreements that break down before their maturity date.

Definition 5 *A Pareto solution corresponding to a weight vector* $\boldsymbol{\alpha} \in \mathcal{P}$ *is time-consistent individually rational (TCIR), if the following condition holds true for every* $i \in \mathcal{N}$*, for all* $t \in \mathcal{T}$*, and for all* $x_0 \in \mathbb{R}^n$*:*

$$W_c^i(t, x_t^\alpha) \le W_{nc}^i(t, x_t^\alpha).$$

We denote by $\mathbf{TI}$ *the set of all time-consistent individually rational weight vectors, that is,*

$$\mathbf{TI} := \left\{ \boldsymbol{\alpha} \in \mathcal{P} \;\middle|\; W_c^i(t, x_t^\alpha) \le W_{nc}^i(t, x_t^\alpha),\ \forall i \in \mathcal{N},\ \forall t \in \mathcal{T} \right\}. \tag{16}$$

A TCIR agreement is a Pareto solution that is individually rational in all subgames starting along the Pareto-optimal state trajectory. If an agreement is in the set $\mathbf{TI}$, then it will remain in place until its maturity date. Note that the set $\mathbf{TI}$ may be empty.

Remark 4 *The set of all individually rational weight vectors given by* (11) *is equivalent to the imputation set of the NTU game obtained by transforming the dynamic game described by* (1)–(2) *into a static game.*

The following theorem characterizes the relationships between the different agreements specified in Definitions 2–5. In particular, we show that the set of all individually rational Pareto agreements given by (11) admits a certain canonical decomposition.

Theorem 3 *Consider the NTU dynamic game described by* (1)–(2)*. The following relations hold true:*

- *(a)* $\mathbf{I} = \mathbf{I}_0$.
- *(b)* $\mathbf{I}_T = \mathcal{P}$.
- *(c)* $\mathbf{TI}_l = \left(\cap_{t=0}^{l-1} \mathbf{I}_t\right) \cap \bar{\mathbf{I}}_l,\ 1 \le l \le T$.
- *(d)* $\mathbf{TI}_T = \emptyset$.
- *(e)* $\mathbf{TI}_l \cap \mathbf{TI}_{l+1} = \emptyset,\ 1 \le l \le T-1$.
- *(f)* $\mathbf{TI} = \cap_{t=0}^{T} \mathbf{I}_t$.
- *(g)* $\mathbf{TI} \cap \mathbf{TI}_l = \emptyset$, $1 \le l \le T-1$.
- *(h)* $\mathbf{I} = \left(\cup_{t=1}^{T} \mathbf{TI}_t\right) \cup \mathbf{TI}$.

Proof.

- (a) Follows directly from Definitions 2–3.
- (b) Along the Pareto-optimal state trajectory $\tilde{x}^\alpha$, the cost-to-go of the players, from (8) and (9), satisfy $W_c^i(T, x_T^\alpha) = h^i(x_T^\alpha) = W_{nc}^i(T, x_T^\alpha)$. This implies all the weight vectors $\boldsymbol{\alpha} \in \mathcal{P}$ are individually rational at time T.
- (c) Let $\alpha \in \mathbf{TI}_l$. Then from Definition 4, $W_c^i(t, x_t^\alpha) \le W_{nc}^i(t, x_t^\alpha)$ for all $i \in \mathcal{N}$ and $0 \le t \le l-1$ and $W_c^i(l, x_l^\alpha) > W_{nc}^i(l, x_l^\alpha)$ for at least one $i \in \mathcal{N}$. This implies $\alpha \in \left(\cap_{t=0}^{l-1} \mathbf{I}_t\right) \cap \bar{\mathbf{I}}_l$, where $\bar{\mathbf{I}}_l$ is a compliment of set $\mathbf{I}_l$. Therefore, $\mathbf{TI}_l \subseteq \left(\cap_{t=0}^{l-1} \mathbf{I}_t\right) \cap \bar{\mathbf{I}}_l$. Next, let $\alpha \in \left(\cap_{t=0}^{l-1} \mathbf{I}_t\right) \cap \bar{\mathbf{I}}_l$, i.e., the weight vectors satisfy conditions (14), and this implies $\alpha \in \mathbf{TI}_l$. So, $\left(\cap_{t=0}^{l-1} \mathbf{I}_t\right) \cap \bar{\mathbf{I}}_l \subseteq \mathbf{TI}_l$.

(d) From (b) we have $\bar{\mathbf{I}}_T = \emptyset$, then from (c) we have that $\mathbf{TI}_T = \left(\cap_{t=0}^{T-1}\mathbf{I}_t\right) \cap \bar{\mathbf{I}}_T = \emptyset$.

(e) Let $A = \cap_{t=0}^{l-1}\mathbf{I}_t$, $B = \mathbf{I}_l$ and $C = \mathbf{I}_{l+1}$; then we have $\mathbf{TI}_l = A \cap \bar{B}$ and $\mathbf{TI}_{l+1} = (A \cap B) \cap \bar{C}$. Since $(A \cap B) \cap (A \cap \bar{B}) = A \cap (B \cap \bar{B}) = \emptyset$, we have that $\mathbf{TI}_l \cap \mathbf{TI}_{l+1} = \emptyset$.

(f) Follows from Definition 5 and using the same reasoning as in (c).

(g) Using the notation in (e) and letting $D = \cap_{t=l+1}^{T}\mathbf{I}_t$, we have $\mathbf{TI} \cap \mathbf{TI}_l = (A \cap B \cap D) \cap (A \cap \bar{B}) = A \cap (B \cap \bar{B}) \cap D = \emptyset$.

(h) First, assume $\mathbf{I} = \emptyset$. Since $\mathbf{I}_0 = \mathbf{I}$, $\mathbf{TI}_t \subseteq \mathbf{I}_0$ and $\mathbf{TI} \subseteq \mathbf{I}_0$, we have that $\left(\cup_{t=1}^{T}\mathbf{TI}_t\right) \cup \mathbf{TI} = \emptyset$. The property holds true. Next, consider $\mathbf{I} \neq \emptyset$; then from (a), any $\alpha \in \mathbf{I}$ is individually rational at stage 0. However, this α may or may not be time-consistent individually rational, and this implies either that there exists $l \leq T-1$ and $l \geq 1$ such that $\alpha \in \mathbf{TI}_l$ or that $\alpha \in \mathbf{TI}$. So, $\alpha \in \left(\cup_{t=1}^{T}\mathbf{TI}_t\right) \cup \mathbf{TI}$, and this implies $\mathbf{I} \subseteq \left(\cup_{t=1}^{T}\mathbf{TI}_t\right) \cup \mathbf{TI}$. Since, $\mathbf{TI} \subseteq \mathbf{I}_0 = \mathbf{I}$ and $\mathbf{TI}_t \subseteq \mathbf{I}_0 = \mathbf{I}$ for all $1 \leq t \leq T$, we have $\left(\cup_{t=1}^{T}\mathbf{TI}_t\right) \cup \mathbf{TI} \subseteq \mathbf{I}$.

□

When $\mathbf{TI}_l \neq \emptyset$, it implies that there exist agreements $\boldsymbol{\alpha} \in \mathcal{P}$ in the game that are time-consistent individually rational until stage $l-1$ and breaks down at stage l. Using Theorem 3 the agreement space can be partitioned into regions corresponding to agreements that are time-consistent individually rational until stage l. Properties (e) and (g) are natural consequences of Definition 4. Property (h) says that all individually rational agreements are a disjoint union of time-consistent individually rational agreements that remain in place until the end of the game and those that break down before the maturity date.

Remark 5 *The framework presented in this section is general and existence of time-consistent agreements can be verified in dynamic games with non-linear dynamics and objectives. Further, the framework can be easily adapted to continuous time and stochastic settings. Moreover, Assumption 1 about players' behavior can be extended to a group rational setting resulting in time-consistent agreements, which have stronger inter-temporal stability properties.*

Remark 6 *Following Assumption 1, Nash equilibrium strategies can be taken as the non-cooperative strategies to be used by the players in the subgame when the agreement breaks down. In some dynamic games, the uniqueness of the Nash equilibrium in any subgame follows from the functional forms of the reward functions and the state dynamics. However, if the Nash equilibrium is not unique, then one faces the issue of equilibrium selection and in the subsequent computation of cost-to-go functions (8) and (9). This difficult problem is way beyond our objective in this paper. To uniquely define the players' equilibrium costs in any subgame, we assume that the players select one Nash equilibrium for every subgame on Pareto-optimal state trajectory.*

4 TCIR Pareto solutions in linear-quadratic games

In this section, we show that the sets introduced in Definitions 2–5 can be computed from the problem data when the dynamic game is of the linear-quadratic variety. To this end, we begin by introducing the discrete-time finite-horizon linear-quadratic difference games (LQDGs). The dynamic interaction environment of the players evolves according to the following difference equation:

$$x_{t+1} = A_t x_t + \sum_{i \in \mathcal{N}} B_t^i u_t^i, \quad t \in \mathcal{T}\backslash\{T\}, \quad x_0 \text{ is given,} \tag{17}$$

where $A_t \in \mathbb{R}^{n \times n}$ and $B_t^i \in \mathbb{R}^{n \times m_i}$. Player $i \in \mathcal{N}$ uses his strategy $\tilde{u}^i$ to minimize the following quadratic objective function:

$$J^i(\tilde{\mathbf{u}}) = \frac{1}{2} x_T' Q_T^i x_T + \sum_{t=0}^{T-1} \left(\frac{1}{2}\, x_t' Q_t^i x_t + \frac{1}{2}\, u_t^{i\prime} R_t^i u_t^i \right), \tag{18}$$

where the matrices $Q_t^i \in \mathbb{R}^{n \times n}$ are symmetric for all $t \in \mathcal{T}$ and for all $i \in \mathcal{N}$, and the matrices R_t^i are symmetric and positive definite for all $t \in \mathcal{T}\backslash\{T\}$ and for all $i \in \mathcal{N}$.

4.1 Pareto solutions

From Theorem 1 and Theorem 2 all the Pareto solutions associated with the NTU game, defined by (17)–(18), can be obtained by minimizing the following weighted-sum optimal-control problem:

$$\min_{\tilde{\mathbf{u}}} \sum_i \alpha^i J^i(\tilde{\mathbf{u}}) \text{ subject to (17)}, \tag{19}$$

with $\boldsymbol{\alpha} \in \mathcal{P}$. For notational convenience we introduce for $\boldsymbol{\alpha} \in \mathcal{P}$ the next matrices $Q_t^\alpha := \sum_i \alpha^i Q_t^i$ for $t \in \mathcal{T}$, $\mathbf{R}_t^\alpha := \oplus_{i=1}^N \alpha^i R_t^i$ and $\mathbf{B}_t := [B_t^1 \ B_t^2 \ \cdots \ B_t^N]$ for $t \in \mathcal{T}\backslash\{T\}$. We introduce the Hamiltonian associated with the above problem as:

$$H_t(x_t, \mathbf{u}_t, \lambda_{t+1}) = \frac{1}{2} x_t' Q_t^\alpha x_t + \frac{1}{2} \mathbf{u}_t' \mathbf{R}_t^\alpha \mathbf{u}_t + \lambda_{t+1}' \left(A x_t + \mathbf{B}_t \mathbf{u}_t\right). \tag{20}$$

Following Theorem 1, the necessary conditions for $\tilde{\mathbf{u}}^\alpha$ to be Pareto-optimal are given by the following two-point boundary value problem:

$$\mathbf{u}_t^\alpha = -\left(\mathbf{R}_t^\alpha\right)^{-1} \mathbf{B}_t' \lambda_{t+1}, \tag{21a}$$

$$x_{t+1} = A_t x_t + \mathbf{B}_t \mathbf{u}_t^\alpha, \ x_0 \text{ is given} \tag{21b}$$

$$\lambda_t = A_t' \lambda_{t+1} + Q_t^\alpha x_t, \ \lambda_T = Q_T^\alpha x_T. \tag{21c}$$

Due to linearity of the above equations, the co-state vector can be taken as linear in the state variable, that is, $\lambda_t = K_t x_t$, and we get

$$\left(I + \mathbf{B}_t' \left(\mathbf{R}_t^\alpha\right)^{-1} \mathbf{B}_t' K_{t+1}\right) x_{t+1} = A_t x_t.$$

If the matrices $\Gamma_t := I + \mathbf{B}_t' \left(\mathbf{R}_t^\alpha\right)^{-1} \mathbf{B}_t' K_{t+1}$ are invertible for all $t \in \mathcal{T}\backslash\{T\}$, this means that the two-point boundary value problem (21a)–(21c) is uniquely solvable and the set of matrices $\{K_t, \ t \in \mathcal{T}\}$ are solutions of the following Ricatti difference equation:

$$K_t = Q_t^\alpha + A_t' K_{t+1} \Gamma_t^{-1} A_t, \quad K_T = Q_T^\alpha. \tag{22}$$

Remark 7 *If the backward difference Equation (22) admits a solution, that is, the matrices $\{\Gamma_t, \ t \in \mathcal{T}\backslash\{T\}\}$ are invertible, then the two-point boundary value problem (21) has a unique solution. To see this, let $\tilde{\lambda}_t = \lambda_t - K_t x_t$ be any other solution of (22). Then substituting this in (21), and after short calculations using (22), we write the two-point boundary value problem in $(x, \tilde{\lambda})$ coordinates as follows:*

$$x_{t+1} = \Gamma_t^{-1} \left(A_t x_t - \mathbf{B}_t (\mathbf{R}_t^\alpha)^{-1} \mathbf{B}_t' \tilde{\lambda}_{t+1}\right)$$
$$\tilde{\lambda}_t = A_t' \tilde{\lambda}_{t+1} - A_t' K_{t+1} \Gamma_t^{-1} \mathbf{B}_t (\mathbf{R}_t^\alpha)^{-1} \mathbf{B}_t' \tilde{\lambda}_{t+1}.$$

The above system of equations is decoupled. From the terminal conditions we have $\tilde{\lambda}_T = 0$, and as a result, we have $\tilde{\lambda}_t = 0$ for all $t \in \mathcal{T}$, and therefore the solution is unique.

As the Equations (21) are necessary conditions, we have that (21a) is a candidate Pareto-optimal control. For sufficiency, we know from Theorem 2 that the minimized Hamiltonian and the salvage value are required to be convex in the state variable. Since, the matrices $Q_t, \ t \in \mathcal{T}$, are not necessarily positive semi-definite we cannot ensure that the candidate Pareto-optimal control obtained by solving the necessary conditions (21) is indeed Pareto-optimal.

To obtain the required sufficient conditions, we transform the dynamic problem (19) into a static optimization problem by eliminating the state variable. Let the matrices $\Lambda_t := \mathbf{R}_t^\alpha + \mathbf{B}_t' M_{t+1} \mathbf{B}_t$ be invertible for all $t \in \mathcal{T}\backslash\{T\}$ with the matrices $M_t, \ t \in \mathcal{T}$ computed as the solution of the following symmetric backward Ricatti difference equation:

$$M_t = A_t' M_{t+1} A_t + Q_t^\alpha - A_t' M_{t+1} \mathbf{B}_t \Lambda_t^{-1} \mathbf{B}_t' M_{t+1} A_t, \ M_T = Q_T^\alpha \tag{23}$$

We define $\Delta_t = \frac{1}{2}x'_{t+1}M_{t+1}x_{t+1} - \frac{1}{2}x'_tM_tx_t$. Then, using the sum $\sum_{t=0}^{T-1}\Delta_t$ and (23), it is possible to write the weighted sum objective function (19) as

$$\sum_i \alpha^i J^i(\tilde{\mathbf{u}}) = \frac{1}{2}x'_0M_0x_0 + \sum_{t=0}^{T-1}\frac{1}{2}||\mathbf{u}_t + \Lambda_t^{-1}\mathbf{B}'_tM_{t+1}A_tx_t||^2_{\Lambda_t}. \tag{24}$$

The next lemma relates how the candidate Pareto-optimal control (21a) obtained from the solution of the two-point boundary value problem (21) is related to the minimizer of problem (24).

Lemma 3 *Let the set of matrices $\{\Lambda_t,\ t\in\mathcal{T}\backslash\{T\}\}$ be invertible and the solutions M_t of the symmetric matrix Riccati difference Equation (23) exist for $t\in\mathcal{T}$. If the two-point boundary value problem*

$$\bar{\lambda}_t = A'_t\bar{\lambda}_{t+1} + Q^\alpha_t\bar{x}_t,\ \ \bar{\lambda}_T = Q^\alpha_T\bar{x}_T \tag{25}$$

$$\bar{x}_{t+1} = A_t\bar{x}_t - \mathbf{B}_t(\mathbf{R}^\alpha_t)^{-1}\mathbf{B}'_t\bar{\lambda}_{t+1},\ \ \bar{x}_0 = x_0, \tag{26}$$

has a unique solution, then we set $\bar{\mathbf{u}}_t = -(\mathbf{R}^\alpha_t)^{-1}\mathbf{B}'_t\bar{\lambda}_t$ and $\gamma_t := \bar{\lambda}_t - M_t\bar{x}_t$. Then the sequences $\{\bar{x}_t, \mathbf{u}^\alpha_t, \bar{\lambda}_t, \gamma_t\}$ solve equations

$$\bar{\mathbf{u}}_t + \Lambda_t^{-1}\mathbf{B}'_tM_{t+1}A_t\bar{x}_t = 0 \tag{27}$$

$$\gamma_t = A'_t\gamma_{t+1} - A'_tM_{t+1}\mathbf{B}_t\Lambda_t^{-1}\mathbf{B}'_t\gamma_{t+1} \tag{28}$$

Proof. To prove (28) we have

$$\begin{aligned}
&A'_t\gamma_{t+1} - A'_tM_{t+1}\mathbf{B}_t\Lambda_t^{-1}\mathbf{B}'_t\gamma_{t+1} - \gamma_t\\
&\quad= A'_t\left(\bar{\lambda}_{t+1} - M_{t+1}\bar{x}_{t+1}\right) - A'_tM_{t+1}\mathbf{B}_t\Lambda_t^{-1}\mathbf{B}'_t\left(\bar{\lambda}_{t+1} - M_{t+1}\bar{x}_{t+1}\right) - \left(A'_t\bar{\lambda}_{t+1} + Q^\alpha_t\bar{x}_t - M_t\bar{x}_t\right)\\
&\quad= \left(M_t - Q^\alpha_t\right)\bar{x}_t + \left(-A'_tM_{t+1} + A'_tM_{t+1}\mathbf{B}_t\Lambda_t^{-1}\mathbf{B}'_tM_{t+1}\right)\bar{x}_{t+1} - A'_tM_{t+1}\mathbf{B}_t\Lambda_t^{-1}\mathbf{B}'_t\bar{\lambda}_{t+1}\\
&\quad= \left(M_t - Q^\alpha_t - A'_tM_{t+1}A_t + A'_tM_{t+1}\mathbf{B}_t\Lambda_t^{-1}\mathbf{B}'_tM_{t+1}A_t\right)\bar{x}_t\\
&\qquad + A'_tM_{t+1}\mathbf{B}_t\Lambda_t^{-1}\left(\Lambda_t(\mathbf{R}^\alpha_t)^{-1} - I - \mathbf{B}'_tM_{t+1}\mathbf{B}_t(\mathbf{R}^\alpha_t)^{-1}\right)\mathbf{B}'_t\bar{\lambda}_{t+1} = 0
\end{aligned}$$

Since, $\gamma_T = \bar{\lambda}_T - M_T\bar{x}_T = Q^\alpha_T\bar{x}_T - Q^\alpha_T\bar{x}_T = 0$, we have that $\gamma_t = 0$ for all $t\in\mathcal{T}$. To prove (27) we have

$$\begin{aligned}
\Lambda_t\bar{\mathbf{u}}_t + \mathbf{B}'_tM_{t+1}A_t\bar{x}_t &= -\Lambda_t(\mathbf{R}^\alpha_t)^{-1}\mathbf{B}'_t\bar{\lambda}_{t+1} + \mathbf{B}'_tM_{t+1}\left(\bar{x}_{t+1} + \mathbf{B}_t(\mathbf{R}^\alpha_t)^{-1}\mathbf{B}'_t\bar{\lambda}_{t+1}\right)\\
&= -\left(\mathbf{R}^\alpha_t + \mathbf{B}'_tM_{t+1}\mathbf{B}_t\right)(\mathbf{R}^\alpha_t)^{-1}\mathbf{B}'_t\bar{\lambda}_{t+1} + \mathbf{B}'_tM_{t+1}\left(\bar{x}_{t+1} + \mathbf{B}_t(\mathbf{R}^\alpha_t)^{-1}\mathbf{B}'_t\bar{\lambda}_{t+1}\right)\\
&= -\left(I + \mathbf{B}'_tM_{t+1}\mathbf{B}_t(\mathbf{R}^\alpha_t)^{-1}\right)\mathbf{B}'_t\bar{\lambda}_{t+1} + \mathbf{B}'_tM_{t+1}\left(\bar{x}_{t+1} + \mathbf{B}_t(\mathbf{R}^\alpha_t)^{-1}\mathbf{B}'_t\bar{\lambda}_{t+1}\right)\\
&= \mathbf{B}'_t\left(M_{t+1}\bar{x}_{t+1} - \bar{\lambda}_{t+1}\right) = -\mathbf{B}'_t\gamma_{t+1} = 0
\end{aligned}$$

□

In the next theorem we provide a necessary and sufficient condition for Pareto-optimality for the dynamic game (19).

Theorem 4 *For every $\boldsymbol{\alpha}\in\mathcal{P}$, let the backward recursive Equation (23) admits solutions such that the set of matrices $\{\Lambda_t, t\in\mathcal{T}\backslash\{T\}\}$ are positive definite. Further, if for every $\boldsymbol{\alpha}\in\mathcal{P}$ the backward recursive Equation (22) admits a solution $\{K_t,\ t\in\mathcal{T}\}$, then there exits a unique Pareto-optimal control given by*

$$\mathbf{u}^\alpha_t = -(\mathbf{R}^\alpha_t)^{-1}\mathbf{B}'_tK_{t+1}x^\alpha_{t+1},\ \ t\in\mathcal{T}\backslash\{T\}, \tag{29}$$

where x^α_t, $t\in\mathcal{T}$ is the Pareto-optimal state trajectory generated by the closed-loop system

$$x^\alpha_{t+1} = \bar{A}^\alpha_tx^\alpha_t,\ \ x^\alpha_0 = x_0, \tag{30}$$

where $\bar{A}_t^\alpha := \Gamma_t^{-1} A_t$. *Further, let the cost-to-go for player* i *evaluated along the Pareto-optimal state trajectory in the subgame starting at* (t, x_t^α) *be given by*

$$W_c^i(t, x_t^\alpha) = \frac{1}{2} x_t^\alpha P_c^i(t; \boldsymbol{\alpha}) x_t^\alpha, \tag{31}$$

where $P_c^i(t; \boldsymbol{\alpha})$ *satisfies the following symmetric backward recursive equation*

$$P_c^i(t; \boldsymbol{\alpha}) = Q_t^i + (\bar{A}_t^\alpha)' \Big(K'_{t+1} \mathbf{B}_t (\mathbf{R}_t^\alpha)^{-1} \mathbf{e}_i R_t^i \mathbf{e}'_i (\mathbf{R}_t^\alpha)^{-1} \mathbf{B}'_t K_{t+1} + P_c^i(t+1; \boldsymbol{\alpha}) \Big) \bar{A}_t^\alpha \tag{32}$$

with $P_c^i(T; \boldsymbol{\alpha}) = Q_T^i$.

Proof. From Remark 7, existence of the solution of (22) implies unique solvablitiy of the two-point boundary value problem. Then, from Lemma 3 we know that candidate Pareto-optimal control (27) satisfies the equation $\mathbf{u}_t^\alpha + \Lambda_t^{-1} \mathbf{B}'_t M_{t+1} A_t x_t^\alpha = 0$. Next, as the matrices Λ_t are positive definite the weighted sum objective function (24) is a strictly convex function in the decision variables $\mathbf{u}_t,\ t \in \mathcal{T} \backslash \{T\}$, with the unique minimizer obtained by setting $\mathbf{u}_t + \Lambda_t^{-1} \mathbf{B}'_t M_{t+1} A_t x_t = 0$. From the sufficient condition provided in Lemma 1 this implies that this control (as a function of the state variable) is a Pareto-optimal control. This implies that the candidate solution obtained by solving the necessary conditions is indeed Pareto-optimal. The cost-to-go of Player i along the Pareto-optimal state trajectory $x_t^\alpha,\ t \in \mathcal{T}$ for the subgame starting from (t, x_t^α) is given by

$$\begin{aligned} W_c^i(t, x_t^\alpha) = \frac{1}{2} x_t^{\alpha\prime} P_c^i(t; \boldsymbol{\alpha}) x_t^\alpha &= \frac{1}{2} x_T^{\alpha\prime} Q_T^i x_T^\alpha + \sum_{\tau=t}^{T-1} \frac{1}{2} x_\tau^{\alpha\prime} Q_\tau^i x_\tau^\alpha + \frac{1}{2} u_\tau^{i\,\prime} R_\tau^i u_\tau^i \\ &= \frac{1}{2} x_t^{\alpha\prime} Q_t^i x_t^\alpha + \frac{1}{2} u_t^{i\prime} R_t^i u_t^i + \frac{1}{2} x_{t+1}^\alpha{}' P_c^i(t+1; \boldsymbol{\alpha}) x_{t+1}^\alpha. \end{aligned}$$

The Pareto-optimal control of Player i is given by $u_t^i = \mathbf{e}'_i \mathbf{u}_t^\alpha$. Then using (29) in the above equation and equating the coefficients which are quadratic in x_t^α on both sides we obtain (32).

□

Remark 8 *It is easy to verify that if the matrices* $Q_t^i,\ t \in \mathcal{T}$ *are positive semi-definite, then the matrices* $P_c^i(t; \alpha)$ *are positive semi-definite for all* $i \in \mathcal{N}, t \in \mathcal{T}$.

4.2 Nash equilibrium

At any intermediate stage, the players can reconsider continuing with the agreement or not. At any stage $t \in \mathcal{T}$, the players will find it optimal to continue their cooperation if each player's Pareto-optimal cost-to-go is lower than his non-cooperative cost-to-go. Note that this comparison is carried out along the Pareto-optimal state trajectory x_t^α, which means that the players have implemented the Pareto-optimal control $\tilde{\mathbf{u}}^\alpha$ until stage $t-1$.

In this paper, we assume that players use Nash equilibrium strategies when they play non-cooperatively towards reevaluation of the cooperative agreement in the sub-games. A Nash equilibrium strategy profile $\tilde{\mathbf{u}}^\diamond|_t := (\tilde{u}^{i\diamond}|_t, \tilde{u}^{i^-\diamond}|_t)$ is such that, for every player $i \in \mathcal{N}$ solves the following optimal control problem

$$\min_{\tilde{u}^i|_t} \quad sJ^i(\tilde{u}^i|_t, \tilde{u}^{i^-\diamond}|_t) \tag{33}$$

subject to

$$x_{\tau+1} = A_\tau x_\tau + \sum_{j \in i^-} B_\tau^j u_\tau^{j\,\diamond} + B_\tau^i u_\tau^i,\ x_t = x_t^\alpha,$$

where

$$sJ^i(\tilde{u}^i|_t, \tilde{u}^{i^-\diamond}|_t) = \frac{1}{2} x'_T Q_T^i x_T + \sum_{\tau=t}^{T-1} \left(\frac{1}{2}\ x'_\tau Q_\tau^i x_\tau + \frac{1}{2}\ u_\tau^{i\,\prime} R_\tau^i u_\tau^i \right).$$

In multistage games the interaction environment is dynamic, which is embedded in state variables and their evolution. It is well known [1] that the Nash equilibrium solution varies with the information used by the players when making their decisions. In the literature (see, e.g., [1] and [10]) these information structures have been defined for dynamic games as open-loop, closed-loop and feedback information structures. In an open-loop information structure, the players design their strategies using only the knowledge of time t (and initial state x_0). Whereas in a closed-loop and feedback information structures, players design their equilibrium strategies using the knowledge of current time t and the current state variable x_t. This implies that the durability of the cooperative agreement depends on the information structure adopted by the players in their non-cooperative play when reevaluating the cooperative agreement. We have the following assumption on the behavior of players in the subgames.

Assumption 2 *When determining the Nash equilibrium strategies in the subgames starting at (t, x^α_t) players use the same information structure (either open-loop or feedback) at all times $t \in \mathcal{T}$.*

Associated with the subgame the open-loop Nash-equilibrium strategies of the players are obtained as follows. Before stating the theorem we assume that the solutions of the following matrix difference equations exist for every $i \in \mathcal{N}$ and $\tau = t, t+1, \cdots, T$:

$$\Psi_\tau = I + \sum_{j\in\mathcal{N}} B^j_\tau R^j_\tau B^{j\prime}_\tau K^j_{\tau+1} \tag{34a}$$

$$K^i_\tau = Q^i_\tau + A'_\tau K^i_{\tau+1}\Psi^{-1}_\tau A_\tau, \quad K^i_T = Q^i_T \tag{34b}$$

$$\Lambda^i_\tau = R^i_\tau + B^{i\prime}_\tau M^i_{\tau+1} B^i_\tau \tag{34c}$$

$$M^i_\tau = Q^i_\tau - A'_\tau M^i_{\tau+1} B^i_\tau {\Lambda^i_\tau}^{-1} B^{i\prime}_\tau M^i_{\tau+1} A_\tau \tag{34d}$$

$$+ A'_\tau M^i_{\tau+1} A_\tau, \quad M^i_T = Q^i_T \tag{34e}$$

We have the following theorem from [13] concerning the sufficient condition related to the existence and uniqueness of the open-loop Nash equilibrium.

Theorem 5 *[13, Theorem 2.5] Let the symmetric Riccati difference Equation (34e) admit solutions such that the matrices $R^i_\tau + B^{i\prime}_\tau M^i_{t+1} B^i_\tau$ are positive definite, for $i \in \mathcal{N}$, $\tau = t, t+1, \cdots, T$. If furthermore the discrete-time open-loop Nash Riccati difference Equation (34b) admit solutions K^i_τ, $i = 1, 2, \cdots, N$, $\tau = t, t+1, \cdots, T-1$, then there exists a unique open-loop Nash equilibrium for the subgame (33), and is given in the feedback form by*

$$u^{i\diamond}_\tau = E^i_\tau x^\diamond_\tau, \ i \in \mathcal{N}, \ \tau = t, t+1, \cdots, T-1, \tag{35}$$

where

$$E^i_\tau = -{R^i_\tau}^{-1} B^{i\prime}_\tau K^i_{\tau+1}\Psi^{-1}_\tau A_\tau \tag{36}$$

and $x^\diamond_\tau$ is a solution of the closed-loop system for $\tau \ge t$

$$x^\diamond_{\tau+1} = \bar{A}^o_\tau x^\diamond_\tau, \ \bar{A}^o_\tau = \Psi^{-1}_\tau A_\tau, \ x^\diamond_t = x^\alpha_t. \tag{37}$$

The open-loop Nash equilibrium state trajectory is obtained as $x^\diamond_\tau = \Phi^o(\tau, t)x^\diamond_t$, where $\Phi^o(\tau, t) = \bar{A}^o_{\tau-1}\bar{A}^o_{\tau-2}\cdots\bar{A}^o_{t+1}\bar{A}^o_t$ for $\tau > t$ and $\Phi^o(\tau, \tau) = I$. Player i's open-loop Nash equilibrium payoff in the subgame starting at (t, x^α_t) is given by

$$W^i_{nc}(t, x^\alpha_t) = \frac{1}{2}x^{\alpha\prime}_t P^{io}_t x^\alpha_t = \frac{1}{2}x^{\diamond\prime}_T Q^i_T x^\diamond_T + \sum_{\tau=t}^{T-1}\left(\frac{1}{2}x^{\diamond\prime}_\tau Q^i_\tau x^\diamond_\tau + \frac{1}{2}u^{i\diamond\prime}_\tau R^i_\tau u^{i\diamond}_\tau\right), \tag{38}$$

where

$$P^{io}_t = \Phi^o(T, t)' Q^i_T \Phi^o(T, t) + \sum_{\tau=t}^{T-1}\Phi^o(\tau, t)'\left(Q^i_\tau + E^{i\prime}_\tau R^i_\tau E^i_\tau\right)\Phi^o(\tau, t). \tag{39}$$

Next, associated with the subgame the feedback Nash equilibrium strategies of the players are obtained as follows. We introduce the following matrices in preparation for the next result. Let F_τ^i $i \in N$, $\tau = t, t+1, \cdots, T$ be the set of matrices satisfying the following linear matrix equations:

$$\left(R_\tau^i + B_\tau^{i\prime} N_{\tau+1}^i B_\tau^i\right) F_\tau^i + B_\tau^{i\prime} N_{\tau+1}^i \sum_{j \in i^-} B_\tau^j F_\tau^j = -B_\tau^{i\prime} N_{\tau+1}^i A_\tau,\ i \in \mathcal{N}, \tag{40}$$

where the matrices N_τ^i $(i \in \mathcal{N})$ are obtained recursively from

$$N_\tau^i = \left(A_\tau + \sum_i B_\tau^i F_\tau^i\right)' N_{\tau+1}^i \left(A_\tau + \sum_i B_\tau^i F_\tau^i\right) + Q_\tau^i + F_\tau^{i\prime} R_\tau^i F_\tau^i,\ N_T^i = Q_T^i. \tag{41}$$

Associated with (40) we define the matrices for $\tau = t, t+1, \cdots, T-1$

$$[\Omega_\tau]_{ij} = \begin{cases} R_\tau^i + B_\tau^{i\prime} N_{\tau+1}^i B_\tau^i & i = j \\ B_\tau^{i\prime} N_{\tau+1}^j B_\tau^j & i \neq j. \end{cases} \tag{42}$$

From [1, Corollary 6.1] and [1, Remark 6.4] we have the following result on the existence of feedback Nash equilibrium.

Theorem 6 *[1, Corollary 6.1] Let the set of matrices Ω_τ, $\tau = t, t+1, \cdots, T-1$ be invertible, and the set of matrices $R_\tau^i + B_\tau^{i\prime} N_\tau^i B_\tau^i$ be positive definite for $i \in \mathcal{N}$, $\tau = t, t+1, \cdots, T-1$, then there exists a unique feedback Nash equilibrium for the subgame (33), and is given by*

$$u_\tau^{i\diamond} = F_\tau^i x_\tau^\diamond,\ i \in \mathcal{N},\ \tau = t, t+1, \cdots, T-1, \tag{43}$$

where $x_\tau^\diamond$ is a solution of the closed-loop system for $\tau \geq t$

$$x_{\tau+1}^\diamond = \bar{A}_\tau^f x_\tau^\diamond,\ \bar{A}_\tau^f = \left(A_\tau + \sum_i B_\tau^i F_\tau^i\right),\ x_t^\diamond = x_t^\alpha. \tag{44}$$

The feedback Nash equilibrium state trajectory is obtained as $x_\tau^\diamond = \Phi^f(\tau, t) x_t^\diamond$, where $\Phi^f(\tau, t) = \bar{A}_{\tau-1}^f \bar{A}_{\tau-2}^f \cdots \bar{A}_{t+1}^f \bar{A}_t^f$ for $\tau > t$ and $\Phi^f(\tau, \tau) = I$. Player i's feedback Nash equilibrium payoff in the subgame starting at (t, x_t^α) is given by

$$W_{nc}^i(t, x_t^\alpha) = \frac{1}{2} x_t^{\alpha\prime} P_t^{if} x_t^\alpha = \frac{1}{2} x_T^{\diamond\prime} Q_T^i x_T^\diamond + \sum_{\tau=t}^{T-1} \left(\frac{1}{2} x_\tau^{\diamond\prime} Q_\tau^i x_\tau^\diamond + \frac{1}{2} u_\tau^{i\diamond\prime} R_\tau^i u_\tau^{i\diamond}\right), \tag{45}$$

where

$$P_t^{if} = \Phi^f(T, t)' Q_T^i \Phi^f(T, t) + \sum_{\tau=t}^{T-1} \Phi^f(\tau, t)' \left(Q_\tau^i + F_\tau^{i\prime} R_\tau^i F_\tau^i\right) \Phi^f(\tau, t). \tag{46}$$

4.3 TCIR Pareto-optimal solutions

In this section, we provide conditions for the existence of open-loop and feedback individually rational, and time-consistent individually rational Pareto-optimal solutions in linear-quadratic difference games described by (17)–(18) for an arbitrary initial state $x_0 \in \mathbb{R}^n$. In the following theorem, to save on notation we denote the matrices $P_t^i = P_t^{io}$, $t \in \mathcal{T}$ when players use open-loop information structure and by $P_t^i = P_t^{if}$, $t \in \mathcal{T}$ when players use feedback information structure in the subgames.

Theorem 7 *Let $\boldsymbol{\alpha} \in \mathcal{P}$, and let the sequence of matrices $\{P_c^i(t;\alpha),\ P_t^i,\ t \in \mathcal{T},\ i \in \mathcal{N}\}$ be generated by (32), (39) and (46). Then, for any arbitrary initial state $x_0 \in \mathbb{R}^n$, we have*

$$\mathbf{I}_l := \left\{\boldsymbol{\alpha} \in \mathcal{P} \mid P_l^i - P_c^i(l; \boldsymbol{\alpha}) \succeq 0,\ \forall i \in \mathcal{N}\right\}, \tag{47}$$

$$\mathbf{TI} := \left\{\boldsymbol{\alpha} \in \mathcal{P} \mid P_l^i - P_c^i(l; \boldsymbol{\alpha}) \succeq 0,\ \forall i \in \mathcal{N},\ 0 \leq l \leq T\right\} \tag{48}$$

where $\mathbf{I}_l$ is the set of individually rational Pareto solutions at stage l, and $\mathbf{TI}$ the set of time-consistent individually rational Pareto solutions.

Proof. Starting from an arbitrary initial state $x_0 \in \mathbb{R}^n$, let x_l^α be the state vector reached in stage l after using the Pareto-optimal control $\tilde{\mathbf{u}}^\alpha$ until stage $l-1$. If condition (47) is satisfied, then $x_l^{\alpha\prime} P_l^i x_l^\alpha \geq x_l^{\alpha\prime} P_c^i(l;\boldsymbol{\alpha}) x_l^\alpha$. Then, from (31), (38) and (45), we have $W_{nc}^i(l, x_l^\alpha) \geq W_c^i(l, x_l^\alpha)$. That is, for every player $i \in \mathcal{N}$, the Player i's cost when using the Pareto-optimal control is lower than the Nash equilibrium cost in the subgame starting from (l, x_l^α).

From Definition 3, we have that the Pareto solution corresponding to α is individually rational at stage l. Next, assume that starting from an arbitrary x_0, players use the Pareto-optimal control $\tilde{\mathbf{u}}^\alpha$ until stage $l-1$ and reach the state x_l^α. If condition (48) is satisfied for all $i \in \mathcal{N}$ and for all $l \in \mathcal{T}$, then $x_l^{\alpha\prime} P_l^i x_l \geq x_l^{\alpha\prime} P_c^i(l;\boldsymbol{\alpha}) x_l^\alpha$. Again, from (31), (38) and (45), for all $l \in \mathcal{T}$ and $i \in \mathcal{N}$. That is, starting from every subgame (l, x_l^α), $l \in \mathcal{T}$, for every Player $i \in \mathcal{N}$, the Player i's cost when using the Pareto-optimal control is lower than the Nash-equilibrium cost. From Definition 5, we have that the Pareto solution corresponding to the weight $\boldsymbol{\alpha}$ is time-consistent individually rational. □

Using the characterization provided in Theorem 7 and from the properties derived in Theorem 3, we can evaluate the intrinsic stability, in the sense of time-consistent individual rationality, of agreements or Pareto-optimal solutions in terms of the problem data. Further, using the results in Theorem 7, we can test if a single-valued bargaining solution, e.g., a Nash, Kalai-Smorodinsky, and egalitarian solution, is time-consistent individual rational or not. Recall that these solutions are Pareto-optimal and fair, and each corresponds to a unique weight vector $\alpha^* \in \mathcal{P}$.

Remark 9 *Implementing any of these bargaining solutions requires to define a threat point, which gives what each player gets in case negotiation fails. In our case, in every subgame starting from (t, x_t^α), the threat point corresponds to the vector of Nash equilibrium cost-to-go $(W_{nc}^i(t, x_t^\alpha),\ i \in \mathcal{N})$.*

5 Numerical illustration

To illustrate our results we analyze the trans-boundary pollution game studied in [4, Chapter 12.4]. We consider two players (e.g., countries) involved in economic activities which result in emissions. The emissions add to the stock of pollution according to the discrete time dynamics

$$S_{t+1} = \delta S_t + E_t^1 + E_t^2, \tag{49}$$

where $0 \leq \delta \leq 1$ accounts for the rate at which pollution is cleared naturally, E_t^i denotes emissions by Player i due to economic activity at time t, and S_0 is the initial stock of pollution. Player i derives a utility $U(E_t^i) = a_i E_t^i - \frac{c_i}{2} {E_t^i}^2$ from emissions, but receives a disutility $D_i(S_t) = \frac{1}{2} b_i S_t^2$, where a_i, b_i, c_i are positive constants for $i = 1, 2$. Player i seeks to minimize the net cost

$$J^i = \frac{w_i}{2} S_T^2 + \sum_{t=0}^{T-1} \frac{b_i}{2} S_t^2 - \left(a_i E_t^i - \frac{c_i}{2} {E_t^i}^2 \right), \tag{50}$$

where the first term on the right hand side is the salvage value with parameter $w_i > 0$. Clearly, the game described by (49)–(50) is a linear-quadratic difference game. As both players incur costs due to pollution stock (49), a negative externality, there is an incentive to cooperate. We seek to analyze the stability and intrinsic longevity, in the sense of TCIR, of the cooperative agreements made by the players. To proceed further, we first transform the dynamic game described by (49)–(50) into the standard form studied in Section 4.

We introduce a dummy variable $z_t = 1$ (a constant) and define the state variable as $x_t = [S_t \ z_t]'$. Next, we define the control variable for player i as

$$u_t^i = E_t^i - \begin{bmatrix} 0 & \frac{a_i}{c_i} \end{bmatrix} x_t.$$

Using these variables the dynamics (49) can be written as

$$x_{t+1} = \begin{bmatrix} \delta & \frac{a_1}{c_1} + \frac{a_2}{c_2} \\ 0 & 1 \end{bmatrix} x_t + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_t^1 + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_t^2, \ x_0 = \begin{bmatrix} S_0 \\ 1 \end{bmatrix}, \tag{51}$$

and the objectives (50) of the players $i = 1, 2$ are given by

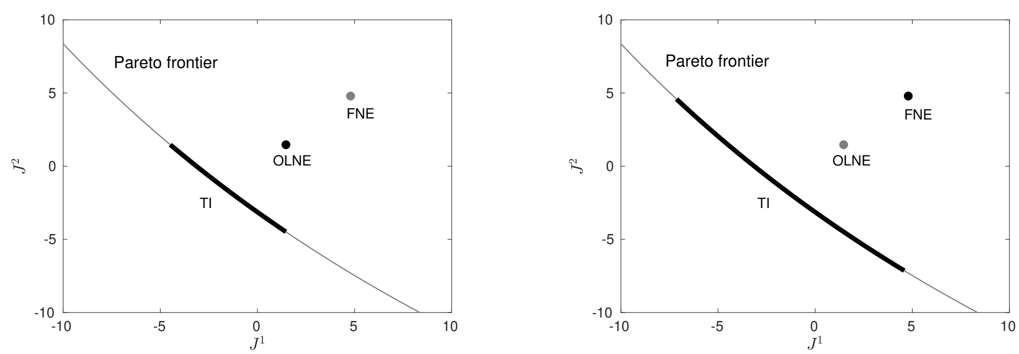
$$J^i = \frac{1}{2} x'_T \begin{bmatrix} w_1 & 0 \\ 0 & 0 \end{bmatrix} x_T + \sum_{t=0}^{T-1} \frac{1}{2} x'_t \begin{bmatrix} b & 0 \\ 0 & -\frac{a_i^2}{c_i} \end{bmatrix} x_t + \frac{1}{2} c_i u_i^2. \tag{52}$$

We illustrate the results of the previous sections. The parameters for the baseline scenario are set as follows:

$$\delta = 0.75, \quad a_1 = a_2 = 5, \quad c_1 = c_2 = 5, \quad b = 1,$$
$$w_1 = 1, \quad w_2 = 1, \quad T = 7, \quad S_0 = 1.$$

We compare our results under two behavioral settings, that is when players use open-loop and feedback information structures in the non-cooperative play in the subgames when reevaluating the cooperative agreement. As players have access to the state variable (dynamic information) they react more aggressively in the feedback case compared to open-loop case. This leads to higher emissions and higher costs when players use feedback strategies. Our objective is to analyze the effect of information structure on the inter-temporal stability of the cooperative agreement.

Figure 1 illustrates the open-loop and feedback time-consistent individually rational Pareto-optimal agreements. In Table 1 we compute the time-consistent individually rational until stage l sets. We notice that in the open-loop case all the $\mathbf{TI}_l$ sets are empty. This implies that all the individually rational Pareto-optimal weight vectors are time-consistent. On the other hand in the feedback case, the time-consistent individually rational until stage l sets are non-empty, implying that there exist individually rational agreements that can break down before the maturity date. Further, from Figure 1 we observe that the open-loop costs of players strictly dominate the feedback costs. This implies that the set of open-loop individually rational Pareto solutions is strictly included in the set of feedback individually rational Pareto solutions. The weights of the Nash, Kalai-Smorodinsky, and egalitarian bargaining in the open-loop and feedback cases are $\alpha_o^N = \alpha_o^{KS} = \alpha^e = \alpha_f^N = \alpha_f^{KS} = \alpha^e = 0.5$. Clearly, all these agreements sustain till the date of maturity.



(a) with open-loop strategies **(b) with feedback strategies**

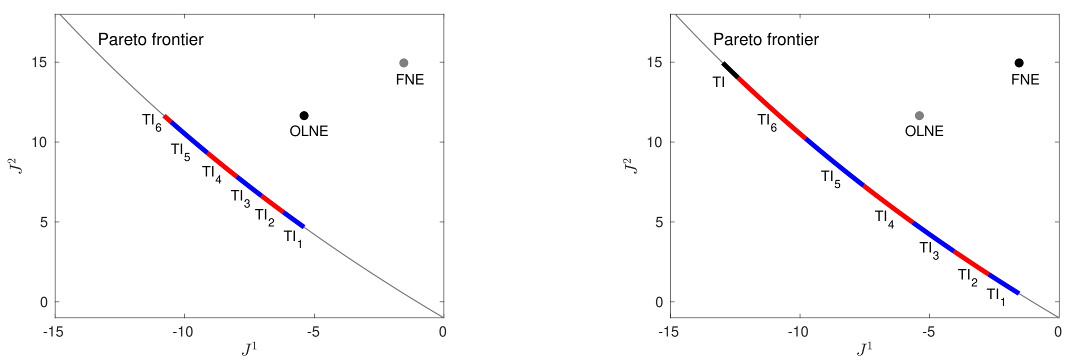
Figure 1: With $w_2 = 1$: The thick dark region on the Pareto frontier in panels (a) and (b) indicate the TCIR Pareto-optimal agreements.

Next, we increase the salvage value parameter of Player 2 to $w_2 = 2.3$. This implies that Player 2 incurs higher costs than Player 1 and as a result the bargaining power of Player 1 should be higher in the cooperative agreement. Since the players are more aggressive with feedback strategies, they receive higher costs compared to the open-loop strategies. Figure 2 illustrates this aspect as the open-loop costs strictly dominate the feedback costs. Further, Figure 2 illustrates the canonical decomposition of individually rational open-loop and feedback Pareto agreements on the Pareto frontier. Table 2 gives the canonical decomposition of individually rational Pareto-optimal weight vectors. The effect of the salvage value becomes prominent in the subgames closer to the maturity date. This implies that the individually rational until stage l weight vectors shift towards higher bargaining power to Player 1. However, due to higher costs of Player 2, Player 1 would be unwilling to share the costs in the subgames close to maturity date. This implies, there is a chance

Table 1: Canonical decomposition of individually rational Pareto-optimal agreements (or weight vectors) with $w_2 = 1$.

Stage	Open-loop weights		Feedback weights	
t	$\mathbf{I}_t$	$\mathbf{TI}_t$	$\mathbf{I}_t$	$\mathbf{TI}_t$
0	[0.47265, 0.52735]		[0.44456, 0.55544]	
1	[0.47064, 0.52936]	$\emptyset$	[0.44467, 0.55533]	[0.44456, 0.44467)∪(0.55533, 0.55544]
2	[0.46793, 0.53207]	$\emptyset$	[0.44485, 0.55515]	[0.44467, 0.44485)∪(0.55515, 0.55533]
3	[0.46427, 0.53573]	$\emptyset$	[0.44512, 0.55488]	[0.44485, 0.44512)∪(0.55488, 0.55515]
4	[0.45933, 0.54067]	$\emptyset$	[0.44549, 0.55451]	[0.44512, 0.44549)∪(0.55451, 0.55488]
5	[0.45295, 0.54705]	$\emptyset$	[0.44598, 0.55402]	[0.44549, 0.44598)∪(0.55402, 0.55451]
6	[0.44626, 0.55374]	$\emptyset$	[0.44626, 0.55374]	[0.44598, 0.44626)∪(0.55374, 0.55402]
7	[0, 1]	$\emptyset$	[0, 1]	$\emptyset$
TI	[0.47265, 0.52735]		[0.44626, 0.55374]	

that the agreement can breakdown in the later stages. This aspect is illustrated in Figure 2 for both the open-loop and feedback individually rational Pareto agreements. More specifically, in Figure 2(b) the small dark shaded region on the Pareto frontier illustrates the TCIR Pareto agreements. From Table 2, we observe that in the open-loop case there exists individually rational agreements that break down at stages $l = 1, 2, \cdots, 6$ and there does not exist agreements that can sustain till the date of maturity. On the other hand, in the feedback case, besides existence of agreements that breakdown before the date of maturity there also exist agreements that sustain for the entire duration. The weights of the Nash, Kalai-Smorodinsky, and egalitarian weights for the open-loop case are computed as $\alpha_o^N = 0.56342$, $\alpha_o^{KS} = 0.56342$ and $\alpha_o^E = 0.56691$ respectively. From Table 2 it is clear that they breakdown at stage 4. The weights of the Nash, Kalai-Smorodinsky, and egalitarian weights for the feedback case are $\alpha_f^N = 0.5577$, $\alpha_f^{KS} = 0.55767$, and $\alpha_f^E = 0.56454$ respectively. The Nash, and egalitarian bargaining agreements break down at stage 4 where as the Kalai-Smorodinsky agreement breaks down at stage 5.



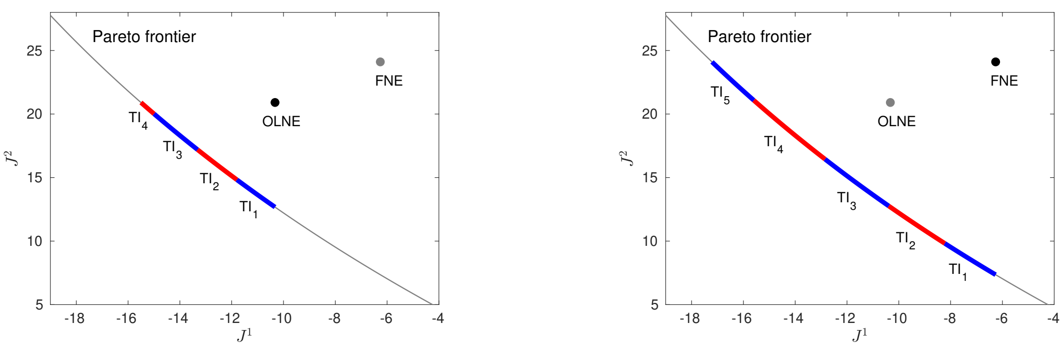
(a) with open-loop strategies **(b) with feedback strategies**

Figure 2: With $w_2 = 2.3$: In panels (a) and (b) decomposition of individually rational Pareto-optimal agreements which break down before the maturity date are shaded in blue and red colors for both open-loop and feedback cases. In panel (b) the dark shaded region on the Pareto frontier illustrates the set of feedback TCIR Pareto-optimal agreements.

Next, we increase the salvage value of Player 2 to $w_2 = 4$. Now, Player 2 costs are significantly higher than Player 1. This implies, as the terminal payoff's effect becomes prominent in the later stages, Player 1 will be unwilling to stay in cooperation in the later stages of the game. Figure 3 illustrates the decomposition of individually rational Pareto-optimal agreements in both the open-loop and feedback cases. Table 3 illustrates the canonical decomposition of individually rational Pareto-optimal weight vectors. In the open-loop case there exist agreements that break down at stages $l = 1, 2, 3, 4$ and in the feedback case there exist agreements that break down at stages $l = 1, 2, 3, 4, 5$. Further, in both information structures there does not exist a TCIR Pareto-optimal agreement, though in the feedback case the agreements are stable for one additional stage. The weights of the Nash, Kalai-Smorodinsky, and egalitarian weights for the open-loop case are $\alpha_o^N = 0.61383$, $\alpha_o^{KS} = 0.61382$, and $\alpha_o^E = 0.62027$ respectively. From Table 3 it is clear that the Nash and Kalai-Smorodinsky agreements breakdown at stage 2, whereas the egalitarian agreements breaks down at stage 3. The weights of the Nash, Kalai-Smorodinsky, and egalitarian weights for the feedback case are $\alpha_f^N = 0.60413$, $\alpha_f^{KS} = 0.60406$, and $\alpha_f^E = 0.61671$ respectively. The Nash and Kalai-Smorodinsky agreements breakdown at stage 3, whereas the egalitarian agreements breaks down at stage 4.

Table 2: Canonical decomposition of individually rational Pareto-optimal agreements (or weight vectors) with $w_2 = 2.3$.

Stage	Open-loop weights		Feedback weights	
t	$\mathbf{I}_t$	$\mathbf{TI}_t$	$\mathbf{I}_t$	$\mathbf{TI}_t$
0	[0.53665, 0.59025]		[0.50199, 0.61374]	
1	[0.54425, 0.60075]	[0.53665, 0.54425)	[0.51229, 0.62324]	[0.50199,0.51229)
2	[0.55232, 0.61271]	[0.54425, 0.55232)	[0.52432, 0.63390]	[0.51229,0.52432)
3	[0.56165, 0.62726]	[0.55232, 0.56165)	[0.53903, 0.64655]	[0.52432,0.53903)
4	[0.57299, 0.64559]	[0.56165, 0.57299)	[0.55728, 0.66166]	[0.53903,0.55728)
5	[0.58742, 0.66910]	[0.57299, 0.58742)	[0.57974, 0.67927]	[0.55728,0.57974)
6	[0.60726, 0.69902]	[0.58742, 0.59025]	[0.60726, 0.69902]	[0.57974,0.60726)
7	[0, 1]	$\emptyset$	[0, 1]	$\emptyset$
TI	$\emptyset$		[0.60726, 0.61374]	



(a) with open-loop strategies **(b) with feedback strategies**

Figure 3: With $w_2 = 4$: In panels (a) and (b) decomposition of individually rational Pareto-optimal agreements which break down before the maturity date are shaded in blue and red colors for both open-loop and feedback cases.

Table 3: Canonical decomposition of individually rational Pareto-optimal agreements (or weight vectors) with $w_2 = 4$.

Stage	Open-loop weights		Feedback weights	
t	$\mathbf{I}_t$	$\mathbf{TI}_t$	$\mathbf{I}_t$	$\mathbf{TI}_t$
0	[0.58635, 0.64143]		[0.54750, 0.66129]	
1	[0.60130, 0.65783]	[0.58635, 0.60130)	[0.56584, 0.67750]	[0.54750, 0.56584)
2	[0.61731, 0.67576]	[0.60130, 0.61731)	[0.58696, 0.69518]	[0.56584, 0.58696)
3	[0.63572, 0.69665]	[0.61731, 0.63572)	[0.61232, 0.71542]	[0.58696, 0.61232)
4	[0.65745, 0.72158]	[0.63572, 0.64143]	[0.64252, 0.73837]	[0.61232, 0.64252)
5	[0.68289, 0.75105]	$\emptyset$	[0.67661, 0.76254]	[0.64252, 0.66129]
6	[0.71080, 0.78350]	$\emptyset$	[0.71080, 0.78350]	$\emptyset$
7	[0, 1]	$\emptyset$	[0, 1]	$\emptyset$
TI	$\emptyset$		$\emptyset$	

6 Conclusions

In this paper, we examined the individually rational cooperative solutions of dynamic games with non-transferable utilities. Given that players cannot redistribute the joint costs in NTU games, they agree to choose the joint strategies that minimize the weighted sum of their cost functions. Solving this optimization problem, we obtain the Pareto solution corresponding to a particular vector of weights. The individual rationality of such a Pareto solution is initially satisfied, but may break down at some intermediate period. In this case, the individually rational cooperative solution is not time-consistent. We provide four definitions related to individual rationality and time consistency of Pareto solutions and examine the sets of weights guaranteeing satisfaction of these properties. We specify the results for linear-quadratic difference games defining the set of open-loop and feedback time-consistent individually rational Pareto solutions for this class of games. A numerical illustration shows the time inconsistency of well-known bargaining solutions (Nash, Kalai-Smorodinsky, and egalitarian solutions).

References

[1] T. Başar and G.J. Olsder. Dynamic Noncooperative Game Theory: Second Edition. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1999.

[2] A. de Paz, J. Marín-Solano, and J. Navas. Time-consistent equilibria in common access resource games with asymmetric players under partial cooperation. Environmental Modeling & Assessment, 18(2):171–184, 2013.

[3] E. Dockner and S. Jørgensen. Cooperative and non-cooperative differential game solutions to an investment and pricing problem. The Journal of the Operational Research Society, 35(8):731–739, 1984.

[4] E. J. Dockner, S. Jørgensen, N. V. Long, and G. Sorger. Differential Games in Economics and Management Science. Cambridge, Cambridge University Press, 2000.

[5] F.Y. Edgeworth. Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences. Reprints of economic classics. Kegan Paul, 1881.

[6] J. C. Engwerda. Necessary and Sufficient Conditions for Pareto Optimal Solutions of Cooperative Differential Games. SIAM Journal on Control and Optimization, 48(6):3859–3881, 2010.

[7] R. P. Hämäläinen, A. Haurie, and V. Kaitala. Equilibria and threats in a fishery management game. Optimal Control Applications and Methods, 6(4):315–333, 1985.

[8] A. Haurie. A note on nonzero-sum differential games with bargaining solution. Journal of Optimization Theory and Applications, 18(1):31–39, 1976.

[9] A. Haurie, J.B. Krawczyk, and M. Roche. Monitoring cooperative equilibria in a stochastic differential game. Journal of Optimization Theory and Applications, 81:73–95, 1994.

[10] A. Haurie, J.B. Krawczyk, and G. Zaccour. Games and Dynamic Games. Scientific World, Singapore, 2012.

[11] A. Haurie and M. Pohjola. Efficient equilibria in a differential game of capitalism. Journal of Economic Dynamics and Control, 11(1):65–78, 1987.

[12] A. Haurie and B. Tolwinski. Definition and properties of cooperative equilibria in a two-player game of infinite duration. Journal of Optimization Theory and Applications, 46(4):525–534, 1985.

[13] G. Jank and H. Abou-Kandil. Existence and uniqueness of open-loop nash equilibria in linear-quadratic discrete time games. IEEE Transactions on Automatic Control, 48(2):267–271, 2003.

[14] S. Jørgensen and G. Zaccour. Time Consistency in Cooperative Differential Games. In Georges Zaccour, editor, Decision & Control in Management Science: Essays in Honor of Alain Haurie, pages 349–366. Springer US, Boston, MA, 2002.

[15] E. Kalai. Proportional solutions to bargaining situations: Interpersonal utility comparisons. Econometrica, 45(7):1623–1630, 1977.

[16] E. Kalai and M. Smorodinsky. Other solutions to Nash's bargaining problem. Econometrica, 43(3):513–518, 1975.

[17] G. Leitmann. Cooperative and non-cooperative many players differential games: course held at the Department of Automation and Information, July 1973. Courses and lectures — International Centre for Mechanical Sciences. Springer- Verlag, 1974.

[18] J. Marín-Solano. Time-consistent equilibria in a differential game model with time inconsistent preferences and partial cooperation. In Josef Haunschmied, Vladimir M. Veliov, and Stefan Wrzaczek, editors, Dynamic Games in Economics, pages 219–238. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.

[19] J. Nash. The bargaining problem. Econometrica, 18(2):155–162, 1950.

[20] L. Petrosyan. Stable Solutions of Differential Games with Many Participants. Viestnik of Leningrad University, 19:46–52, 1977.

[21] L.A. Petrosyan and G. Zaccour. Cooperative differential games with transferable payoffs. In Tamer Başar and Georges Zaccour, editors, Handbook of Dynamic Game Theory, pages 595–632. Springer International Publishing, Cham, 2018.

[22] C.P. Simon and L. Blume. Mathematics for Economists. Norton, 1994.

[23] G. Sorger. Recursive nash bargaining over a productive asset. Journal of Economic Dynamics and Control, 30(12):2637–2659, 2006.

[24] B. Tolwinski, A. Haurie, and G. Leitmann. Cooperative Equilibria in Differential Games. Journal of Mathematical Analysis and Applications, 119:182–202, 1986.

[25] D. W. K. Yeung, L. Petrosyan, and P. M. Yeung. Subgame consistent solutions for a class of cooperative stochastic differential games with nontransferable payoffs. In Steffen Jørgensen, Marc Quincampoix, and Thomas L. Vincent, editors, Advances in Dynamic Game Theory: Numerical Methods, Algorithms, and Applications to Ecology and Economics, pages 153–170. Birkhäuser Boston, Boston, MA, 2007.

[26] D. W. K. Yeung and L. A. Petrosyan. Subgame consistent solutions of a cooperative stochastic differential game with nontransferable payoffs. Journal of Optimization Theory and Applications, 124(3):701–724, 2005.

[27] D. W. K. Yeung and L. A. Petrosyan. Subgame consistent cooperative solution for NTU dynamic games via variable weights. Automatica, 59:84–89, 2015.

[28] David W. K. Yeung and Leon A. Petrosyan. Nontransferable utility cooperative dynamic games. In Tamer Başar and Georges Zaccour, editors, Handbook of Dynamic Game Theory, pages 633–670. Springer International Publishing, Cham, 2018.

[29] D.W.K. Yeung and L.A. Petrosyan. Subgame Consistent Cooperation: A Comprehensive Treatise. Theory and Decision Library C. Springer Singapore, 2016.

[30] G. Zaccour. Time Consistency in Cooperative Differential Games: A Tutorial. INFOR: Information Systems and Operational Research, 46(1):81–92, 2008.