Les Cahiers du GERAD ISSN: 0711–2440

Feedback Nash equilibria in differential games with impulse control

U. Sadana, P.V. Reddy, G. Zaccour

G–2021–36

June 2021

La collection *Les Cahiers du GERAD* est constituée des travaux de recherche menés par nos membres. La plupart de ces documents de travail a été soumis à des revues avec comité de révision. Lorsqu'un document est accepté et publié, le pdf original est retiré si c'est nécessaire et un lien vers l'article publié est ajouté.

Citation suggérée : U. Sadana, P.V. Reddy, G. Zaccour (June 2021). Feedback Nash equilibria in differential games with impulse control, Rapport technique, Les Cahiers du GERAD G–2021–36, GERAD, HEC Montréal, Canada.

Avant de citer ce rapport technique, veuillez visiter notre site Web (https://www.gerad.ca/fr/papers/G-2021-36) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

The series *Les Cahiers du GERAD* consists of working papers carried out by our members. Most of these pre-prints have been submitted to peer-reviewed journals. When accepted and published, if necessary, the original pdf is removed and a link to the published article is added.

Suggested citation: U. Sadana, P.V. Reddy, G. Zaccour (June 2021). Feedback Nash equilibria in differential games with impulse control, Technical report, Les Cahiers du GERAD G–2021–36, GERAD, HEC Montréal, Canada.

Before citing this technical report, please visit our website (https://www.gerad.ca/en/papers/G-2021-36) to update your reference data, if it has been published in a scientific journal.

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, 2021
– Bibliothèque et Archives Canada, 2021

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, 2021
– Library and Archives Canada, 2021

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

Feedback Nash equilibria in differential games with impulse control

Utsav Sadana [a]

Puduru Viswanadha Reddy [b]

Georges Zaccour [a]

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

[b] *Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai 600 036, India*

utsav.sadana@mail.mcgill.ca

vishwa@ee.iitm.ac.in

georges.zaccour@gerad.ca

June 2021
Les Cahiers du GERAD
G–2021–36

Copyright © 2021 GERAD, Sadana, Reddy, Zaccour

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs. Les auteurs conservent leur droit d'auteur et leurs droits moraux sur leurs publications et les utilisateurs s'engagent à reconnaître et respecter les exigences légales associées à ces droits. Ainsi, les utilisateurs:

- Peuvent télécharger et imprimer une copie de toute publication du portail public aux fins d'étude ou de recherche privée;
- Ne peuvent pas distribuer le matériel ou l'utiliser pour une activité à but lucratif ou pour un gain commercial;
- Peuvent distribuer gratuitement l'URL identifiant la publication.

Si vous pensez que ce document enfreint le droit d'auteur, contactez-nous en fournissant des détails. Nous supprimerons immédiatement l'accès au travail et enquêterons sur votre demande.

The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*. Copyright and moral rights for the publications are retained by the authors and the users must commit themselves to recognize and abide the legal requirements associated with these rights. Thus, users:

- May download and print one copy of any publication from the public portal for the purpose of private study or research;
- May not further distribute the material or use it for any profit-making activity or commercial gain;
- May freely distribute the URL identifying the publication.

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Abstract: We study a class of deterministic finite-horizon two-player nonzero-sum differential games where players are endowed with different kinds of controls. We assume that Player 1 uses piecewise-continuous controls, while Player 2 uses impulse controls. For this class of games, we seek to derive conditions for the existence of feedback Nash equilibrium strategies for the players. More specifically, we provide a verification theorem for identifying such equilibrium strategies, using the Hamilton-Jacobi-Bellman (HJB) equations for Player 1 and the quasi-variational inequalities (QVIs) for Player 2. Further, we show that the equilibrium number of interventions by Player 2 is upper bounded. Furthermore, we specialize the obtained results to a scalar two-player linear-quadratic differential game. In this game, Player 1's objective is to drive the state variable towards a specific target value, and Player 2 has a similar objective with a different target value. We provide, for the first time, an analytical characterization of the feedback Nash equilibrium in a linear-quadratic differential game with impulse control. We illustrate our results using numerical experiments.

Keywords: Nonzero-sum differential games, feedback Nash equilibrium, linear-quadratic differential games, impulse controls, quasivariational inequalities

Résumé : Nous étudions une classe de jeux déterministes à horizon fini à deux joueurs à somme non nulle où les joueurs sont dotés de différents types de commandes. Nous supposons que le joueur 1 utilise un contrôle continu, tandis que le joueur 2 utilise un contrôle impulsionnel. Pour cette classe de jeux, nous cherchons à dériver des conditions d'existence de stratégies d'équilibre de Nash en rétroaction. Plus précisément, nous fournissons un théorème de vérification pour identifier de telles stratégies d'équilibre, en utilisant les équations de Hamilton-Jacobi-Bellman (HJB) pour le joueur 1 et les inégalités quasi-variationnelles (QVIs) pour le joueur 2. De plus, nous montrons, qu'à l'équilibre, le nombre d'interventions du joueur 2 est majoré. De plus, nous appliquons les résultats obtenus à un jeu différentiel linéaire-quadratique scalaire à deux joueurs. Dans ce jeu, l'objectif du joueur 1 est d'amener l'état vers une valeur cible spécifique, et le joueur 2 a un objectif similaire mais vise une cible différente. Nous proposons, pour la première fois, une caractérisation analytique de l'équilibre de Nash en rétroaction dans un jeu différentiel linéaire-quadratique avec contrôles impulsionnels. Nous illustrons nos résultats à l'aide d'expériences numériques.

Mots clés : Jeux différentiels à somme non nulle, équilibre de Nash en rétroaction, jeux différentiels linéaires-quadratiques, contrôles impulsionnels, inégalités quasi-variationnelles

Acknowledgements: We thank Herbert Dawid for providing valuable comments and suggestions that helped us improve the paper. The first author's research is supported by the FRQNT Doctoral research scholarship (B2X, 275596). The second author's research is supported by SERB, Government of India, grant MTR/2019/000771.

1 Introduction

Many real-world applications, such as industry regulation and cybersecurity, can be modeled as a two-player finite-horizon nonzero-sum differential game, where one player influences the evolution of the state variable continuously over time, whereas the other takes actions that introduce jumps in the state variable at certain strategically chosen discrete time instants. An example of such a setting is a game between an environmental regulation agency, which determines when and by how much to change the cap on pollution emissions, and a (representative) firm, which continuously makes production decisions that have emissions as a by-product.

Nash equilibrium in differential games varies with the information that is available to the players when they determine their strategies, which is also known as the information structure [1]. In our previous paper [2], we introduced a two-player nonzero-sum differential game with impulse controls to study the aforementioned interactions assuming an open-loop information structure, where the strategies of the players are functions of time and the initial state (which is a known parameter). It is well known that open-loop Nash equilibrium (OLNE) strategies are not strongly time consistent, that is, that the equilibrium strategies derived for a given initial state might not constitute the equilibrium of the subgame starting at an intermediate time instant during the game, if the state value at the start of the subgame deviates from the equilibrium state trajectory determined at the start of the game [3]. To address this limitation of open-loop strategies, the literature on differential game theory has focused on a feedback information structure, where players' actions at each instant of time during the game are determined by a strategy that depends on both the current state and the current time [4, 5, 6]. The resulting feedback strategies of the players are known to be strongly time consistent [3].

The objective of this paper is to study the class of games that we have considered in [2], but here under a feedback information structure. In [2], we have studied a class of differential games where Player 1 uses piecewise continuous controls and Player 2 uses impulse controls. The novelty of the present paper lies in providing conditions for the existence of a feedback Nash equilibrium (FNE) in this canonical class of differential games. We have studied these canonical games of minimal configuration for analytical tractability, and our model can be extended to the more general case where both players use both types of controls. FNE is obtained under the assumption that the impulse controls lie within the class of threshold policies, that is, Player 2 gives an impulse only when the state leaves her continuation region, which is characterized by using the Bensoussan Lions quasivariational inequalities (QVIs) [7, 8, 9]. Even for impulse optimal control problems, it is challenging to solve QVIs for a general class of impulse controls (see, e.g., the central bank intervention problems studied in [10] and [11]). Furthermore, threshold policies are quite natural for applications in industry regulation and cybersecurity.

Our contribution is threefold: First, we provide a verification theorem for a general class of differential games with impulse controls that can be used to characterize the FNE strategies. In particular, we show that the (value) functions that satisfy the Hamilton-Jacobi-Bellman equations for Player 1, coupled with a system of QVIs for Player 2, coincide with the respective payoffs of the players in the FNE. The novel feature of our model is that Player 1 can continuously change both the state trajectory and Player 2's continuation set, which is a collection of all time and state vectors for which it is optimal for Player 2 not to intervene in the system. This feature differentiates our work from the existing literature on differential games with impulse control (see [12] and [13]), where the continuous evolution of the state is exogenously given and all players shift the state from one level to another at discrete time instants. Since the FNE strategies obtained by using the verification theorem are a function of the current time and state pairs, they are strongly time consistent.

Second, we show that, under a few regularity assumptions, the equilibrium number of impulses is bounded by a value that is derived from the problem data.

Our third contribution lies in providing, for the first time, a complete analytical characterization of FNE in a scalar linear-quadratic differential game (LQDG) with impulse controls. LQDGs have been widely studied in engineering, economics, and management because they provide a tractable framework to model real-world problems involving nonconstant returns to scale, interactions between the players' control variables, as well as interactions between the state and control variables. LQDGs assume linear state dynamics, which can be seen as a locally reasonable approximation of nonlinear state dynamics. A comprehensive coverage of LQDGs can be found in, e.g., [1, 4, 5, 14], and [6]. However, these references provide existence and uniqueness results for classical differential games, where players only use ordinary controls and where there are no fixed costs in the game. To the best of our knowledge, the literature on differential games does not provide any theoretical or computational means to identify the FNE in nonzero-sum LQDGs with impulse controls.

The specialized linear-quadratic game we study in this paper involves Player 1 using piecewise-continuous controls to minimize the cost associated with the state deviating from her target value, while Player 2 uses impulse controls to instantaneously change the state from one level to another so as to keep the state close to her own target. This model is a multi-agent adaptation of the impulse optimal control problem (single player) studied in [11]. In particular, in our setting, Player 2's impulse optimal control problem is a modified version of the impulse control problem analyzed in [11]. Our regularity assumptions on the value function and impulse controls of Player 2 also follow from [11] where analytical solutions of the HJB equation are obtained in the continuation region by using a quadratic form on the value function; see also [15].

The remainder of the paper is organized as follows. In Section 1.1, we review the literature on impulse optimal control problems, differential games where at least one player uses piecewise-continuous controls, and impulse games where all players use impulse controls only. We introduce our model in Section 2. In Section 3, we provide a verification theorem for the existence of the FNE. In Section 4, we specialize our results to a scalar linear-quadratic game, and we solve this game in Section 5 for different problem parameters. Finally, concluding remarks are given in Section 6.

1.1 Literature review

One of the well-studied impulse control problems is the central bank intervention problem, where the bank intervenes in the foreign exchange market and continuously controls the domestic interest rate to keep the exchange rate close to a target value (see, e.g., [11] and [15]). The characterization of optimal impulse control in a one-decision-maker setting has been the topic of a long series of contributions in diverse fields, e.g, finance [16]; management [17, 18, 19, 20, 21]; and epidemiology [22]. In contrast, the literature in differential games with impulse controls has been very limited, and has predominantly dealt with zero-sum games (see, e.g., [23] and [24]). With the exception of our previous papers [2, 25, 26], the equilibrium solutions in nonzero-sum differential games with impulse controls have been obtained under the assumption that the impulse timing is known a priori [27].

In [2], we provided an algorithm for computing the open-loop Nash equilibrium in linear-quadratic dynamic games with impulse control. Reference [25] characterized the sampled-data Nash equilibrium for the class of games introduced in [2]. Further, [26] determined the FNE for a specialized case of linear-state differential games (LSDGs) with impulse controls, and showed, contrary to the case with ordinary controls, that the FNE and OLNE do not coincide when linear value functions are used to determine the FNE. By definition, LSDGs do not account for nonlinearities in the state variables or interactions between the state and control variables in the players' objective functionals, which limits their applications in practice. In this paper, we relax this restriction and consider a general class of differential games, and by the same token, push further the literature in nonzero-sum differential games.

Our work is closely related to the impulse games studied in [12, 13, 28], and [29] with a feedback information structure where, however, all players are assumed to make discrete-time interventions in the continuous-time stochastic processes. To illustrate, [28] studied a specialized pollution control game between a government that determines the regulatory constraints on emissions and a (representative) firm that takes discrete-time actions to expand its capacity. It is assumed that both the government and the firm use only impulse controls. In [12], the authors studied infinite-horizon nonzero-sum game problem assuming threshold-type impulse controls and showed that a system of QVIs gives sufficient conditions for a FNE if the value functions of both players satisfy certain regularity conditions. There are no piecewise-continuous controls in their model, which limit its applicability to many problems of interest in regulation and security. Reference [13] extended their two-player model to an N-player setting and analyzed the corresponding mean-field game. In [30], a game problem between an impulse player and a stopper is solved using the QVIs. The consideration of impulse controls makes it difficult to analytically characterize Nash equilibria for a general class of differential games, which explains why it is tempting to focus on tractable games. For instance, [12] determined closed-form solutions for symmetric linear-state impulse stochastic games.

2 Model

We consider a deterministic finite-horizon two-player nonzero-sum differential game where the two players can affect a continuously evolving state vector to minimize their individual costs. In our canonical game, the two players are equipped with different types of controls. In particular, Player 1 continuously affects the state vector using her piecewise continuous control $u(t) \in \Omega_1 \subset \mathbb{R}^{m_1}$ while Player 2 uses discrete-time actions to instantaneously change the state by using an impulse control $\tilde{v} = \{(\tau_i, \xi_i)\}_{i\geq 1}$ where τ_i denotes an intervention instant and $\xi_i \in \Omega_2 \subset \mathbb{R}^{m_2}$ denotes the size of the impulse at time τ_i. The sets Ω_1 and Ω_2 are assumed to be bounded and convex.

The state vector is controlled by Player 1 and evolves as follows:

$$\dot{x}(t) = f(x(t), u(t)),\ x(0^-) = x_0,\ \text{for } t \neq \tau_i, i \geq 1. \tag{1}$$

And at the impulse instant, τ_i, Player 2 introduces jumps that are given by

$$x(\tau_i^+) - x(\tau_i^-) = g(x(\tau_i^-), \xi_i), \tag{2}$$

where $f : \mathbb{R}^n \times \mathbb{R}^{m_1} \to \mathbb{R}^n$, $g : \mathbb{R}^n \times \mathbb{R}^{m_2} \to \mathbb{R}^n$, and τ_i^- and τ_i^+ denote the time instants just before and after the intervention time τ_i.

The cost functions of Player 1 and Player 2 are given by

$$J_1(0, x_0, u(.), \tilde{v}) = \int_0^T h_1(x(t), u(t))dt + \sum_{i\geq 1} \mathbb{1}_{0\leq\tau_i<T}\ b_1(x(\tau_i^-), \xi_i) + s_1(x(T)), \tag{3}$$

$$J_2(0, x_0, u(.), \tilde{v}) = \int_0^T h_2(x(t), u(t))dt + \sum_{i\geq 1} \mathbb{1}_{0\leq\tau_i<T}\ b_2(x(\tau_i^-), \xi_i) + s_2(x(T)), \tag{4}$$

where $h_i : \mathbb{R}^n \times \mathbb{R}^{m_1} \to \mathbb{R}$ is the running cost of Player i, $b_i : \mathbb{R}^n \times \mathbb{R}^{m_2} \to \mathbb{R}$ is the cost accrued by Player i at the time of impulse, and $s_i : \mathbb{R}^n \to \mathbb{R}$ is the terminal cost of Player i. Here, $\mathbb{1}_y$ denotes an indicator function of y, that is, $\mathbb{1}_y$ is equal to 1 if y holds; otherwise, it is equal to 0.

3 Feedback Nash equilibrium

We focus our attention on the derivation of Nash equilibrium strategies under a memoryless perfect state information structure, also referred to as feedback Nash or Markov-perfect equilibrium. For this information structure, players use strategies that are functions of the current time t and current state vector $x(t)$.

3.1 Strategy of Player 1 and Player 2

The strategy spaces of the players are described as follows: Let $\Sigma := \big\{(t, x) \,|\, t \in [0, T], x \in \mathbb{R}^n\big\}$ and let $\mathcal{T}$ denote the set of admissible impulse instants. Player 1 affects the continuously evolving state dynamics $x(t)$ using her piecewise-continuous state-feedback strategy $\gamma : [0, T] \times \mathbb{R}^n \to \Omega_1$, while

Player 2 exercises discrete-time actions given by her state-feedback intervention policy δ. Following the literature (see [11] and [12]) on impulse controls, the intervention policy δ involves determining a continuation set $\mathcal{C}$ and a continuous function ζ such that Player 2 gives an impulse if and only if $(t,x) \in \Sigma \backslash \mathcal{C}$, and when Player 2 gives an impulse, its magnitude is given by the function $\zeta : [0,T] \times \mathbb{R}^n \to \Omega_2$. The intervention set $\mathcal{I}$ is given by $\mathcal{I} = \Sigma \backslash \mathcal{C}$. For a given strategy pair (γ, δ), Player 1's control is given by $u(t) = \gamma(t,x)$ and Player 2's impulse control $\tilde{v}$ is a sequence $\{(\tau_i, \xi_i)\}_{i \geq 1}$ where τ_i is the impulse instant and ξ_i is the impulse level.

Remark 1 *We emphasize that the timing of the interventions are given in feedback form as the continuation set $\mathcal{C}$ depends on both the current time and the current state vector. In particular, the continuation and intervention sets will be characterized, in Section 3.3, by the QVIs associated with Player 2's optimal behavior.*

Remark 2 *Nash equilibria in zero-sum differential games with impulse controls have been obtained in the literature (see, e.g., [24] and [31]) assuming nonanticipative strategies [32] where each player determines her strategy as a function of her opponent's strategy in a way that the strategies do not depend on the future strategies of the opponent. For tractability, we focus on feedback strategies that are also considered in [12]. As mentioned in [12], the feedback strategies are dependent on the other player's strategies via the state vector, which can be affected by both the players.*

Remark 3 *The actions of the players associated with an admissible strategy pair (γ, δ) can be described as follows: Player 1 continuously controls the state trajectory using state feedback $\gamma(t,x)$ during the time that the state lies in the continuation set $\mathcal{C}$. When the state leaves set $\mathcal{C}$, Player 2 intervenes and gives an impulse of size $\zeta(t,x)$ to bring the state into set $\mathcal{C}$.*

Definition 1 *The sequence $\tilde{v} = \{(\tau_i, v_i)\}_{i \geq 1}$, is an admissible impulse control of Player 2 if the number of impulses is finite and the impulse instants lie in the set $\mathcal{T}$ given by*

$$\begin{aligned} &\mathcal{T} = \{\tau_i,\, i \geq 1, \quad | \; 0 \leq \tau_1 < \tau_2 < \cdots < T\}, \\ &\tau_n = \inf\{t > \tau_{n-1} : (t,x) \notin \mathcal{C}\},\, n \geq 1,\, \tau_0 := 0. \end{aligned}$$

The above definition ensures that Player 2 gives an impulse as soon as the state leaves the continuation set $\mathcal{C}$.

Next, we determine the cost-to-go functions for Player 1 and Player 2 for a given strategy pair (γ, δ) and for any starting position of the game (t,x). Suppose $\gamma_{[t,T]} \in \Gamma_{[t,T]}$ and $\delta_{[t,T]} \in \Delta_{[t,T]}$ are restrictions of γ and δ, respectively, to the interval $[t,T]$, and $\Gamma_{[t,T]}$ and $\Delta_{[t,T]}$ denote the strategy sets for Player 1 and Player 2, respectively, in the interval $[t,T]$. Then, the state evolution for any starting position of the game (t,x) is given by

$$\dot{x}(t) = f(x(t), \gamma(t, x(t))),\; x(t) = x,\; \text{for } (t,x) \in \mathcal{C}, \tag{5}$$
$$x(\tau_i^+) - x(\tau_i^-) = g(x(\tau_i^-), \xi_i),\; \text{for } (\tau_i, x(\tau_i)) \in \mathcal{I}, \tag{6}$$

and the cost-to-go functions are given by

$$J_1(t,x,\gamma_{[t,T]},\delta_{[t,T]}) = \int_t^T h_1(x(s), \gamma_{[t,T]}(s, x(s)))ds + \sum_{j \geq 1} \mathbb{1}_{t \leq \tau_j < T}\; b_1(x(\tau_j^-), \xi_j) + s_1(x(T)), \tag{7}$$
$$J_2(t,x,\gamma_{[t,T]},\delta_{[t,T]}) = \int_t^T h_2(x(s), \gamma_{[t,T]}(s, x(s)))ds + \sum_{j \geq 1} \mathbb{1}_{t \leq \tau_j < T}\; b_2(x(\tau_j^-), \xi_j) + s_2(x(T)), \tag{8}$$

The differential game described by (5)–(8) constitutes a nonstandard optimal control problem of Player 1 due to intervention costs and state jumps, and an impulse optimal control problem of Player 2.

The feedback Nash equilibrium is defined as follows:

Definition 2 *For the differential game described by (5–8) with a memoryless perfect state information pattern, the strategy profile* $(\gamma^*, \delta^*) \in \Gamma \times \Delta$ *constitutes a feedback Nash equilibrium solution if, for any* $(t, x) \in \Sigma$*, we have*

$$J_1(t, x, \gamma^*_{[t,T]}, \delta^*_{[t,T]}) \leq J_1(t, x, \gamma_{[t,T]}, \delta^*_{[t,T]}),\ \forall \gamma_{[t,T]} \in \Gamma_{[t,T]}, \tag{9a}$$

$$J_2(t, x, \gamma^*_{[t,T]}, \delta^*_{[t,T]}) \leq J_2(t, x, \gamma^*_{[t,T]}, \delta_{[t,T]}),\ \forall \delta_{[t,T]} \in \Delta_{[t,T]}. \tag{9b}$$

3.2 Verification theorem

In this section, we provide methods for identifying the FNE associated with the differential game described by (5–8). To this end, from (9a), we know that the FNE strategy γ^* of Player 1 provides the best response to Player 2's FNE strategy δ^*. Similarly, from (9b), Player 2's FNE strategy δ^* is the best response to Player 1's FNE strategy γ^*. Further, $V_1 : [t, T] \times \mathbb{R}^n \to \mathbb{R}$ and $V_2 : [t, T] \times \mathbb{R}^n \to \mathbb{R}$ denote the equilibrium cost-to-go of the players in the subgame starting at $(t, x) \in \Sigma$, and can be defined as follows using (9a) and (9b):

$$V_1(t, x) = \inf_{\gamma_{[t,T]} \in \Gamma_{[t,T]}} J_1(t, x, \gamma_{[t,T]}, \delta^*_{[t,T]}), \tag{10a}$$

$$V_2(t, x) = \inf_{\delta_{[t,T]} \in \Delta_{[t,T]}} J_2(t, x, \gamma^*_{[t,T]}, \delta_{[t,T]}). \tag{10b}$$

The following is a standing assumption on the value functions, which will be used throughout the paper.

Assumption 1 *The value function of Player 1,* $V_1(t, x)$*, is differentiable in both* t *and* x *when* $(t, x) \in \mathcal{C}$.

From (9a), the value function $V_1(t, x)$ associated with Player 1's optimal control problem satisfies the following Hamilton-Jacobi-Bellman (HJB) equation for a given impulse control $\{(\tau_i^*, \xi_i^*)\}_{i\geq 1}$ corresponding to Player 2's FNE strategy δ^*:

$$-\frac{\partial V_1(t, x)}{\partial t} = \min_{\varphi \in \Omega_1} \mathcal{H}_1\left(t, \varphi, \frac{\partial V_1(t, x)}{\partial x}\right),\ (t, x) \in \mathcal{C}, \tag{11a}$$

$$V_1(T, x(T)) = s_1(x(T)),\ \forall (T, x) \in \Sigma, \tag{11b}$$

$$V_1(\tau_i^{*-}, x(\tau_i^{*-})) = V_1(\tau_i^{*-}, x(\tau_i^{*-}) + g(x(\tau_i^{*-}), \xi_i^*)) + b_1(x(\tau_i^{*-}), \xi_i^*),\ (\tau_i^*, x(\tau_i^*)) \in \mathcal{I}, \tag{11c}$$

where

$$\mathcal{H}_1\left(t, \varphi, \frac{\partial V_1(t, x)}{\partial x}\right) = h_1(x, \varphi) + \left(\frac{\partial V_1(t, x)}{\partial x}\right)^T f(x, \varphi). \tag{11d}$$

The above conditions can be interpreted as follows. From Definition 4, an admissible impulse cannot occur at the terminal time, hence condition (11b) holds. In the continuation region $\mathcal{C}$, Player 2 does not give any impulse, and therefore, the value function of Player 1 satisfies the HJB Equation (11a). When an impulse occurs in the intervention region, that is, $(\tau_i^*, x(\tau_i^*)) \in \mathcal{I}$, then Player 1's cost-to-go is the sum of the additional cost, $b_1(x(\tau_i^{*-}), \xi_i^*)$, incurred due to the intervention by Player 2, and the cost-to-go from playing optimally afterwards.

Remark 4 *We remark that the discontinuities in Player 1's value function can occur only due to interventions by Player 2.*

3.3 Continuation and intervention set

Player 2 solves the impulse optimal control problem (9b) for a given equilibrium strategy $\gamma^*(t, x)$ of Player 1.

Assumption 2 *The value function of Player 2, $V_2(t,x)$, is differentiable in both t and x for almost all values of the state x.*[1]

The value function $V_2(t,x)$ associated with Player 2's impulse control problem satisfies the following system of (weak) QVIs

$$\forall t \in [0,T],\ \text{a.a. } x \in \mathbb{R}^n, \frac{\partial V_2(t,x)}{\partial t} + \mathcal{H}_2\left(x, \gamma^*(t,x), \frac{\partial V_2(t,x)}{\partial x}\right) \geq 0. \tag{12a}$$

$\forall (t,x) \in \Sigma,$ the following two relations hold

$$V_2(t,x) \leq \mathcal{R}V_2(t,x), \tag{12b}$$

$$(V_2(t,x) - \mathcal{R}V_2(t,x))\Big(\frac{\partial V_2(t,x)}{\partial t} + \mathcal{H}_2\left(x, \gamma^*(t,x), \frac{\partial V_2(t,x)}{\partial x}\right)\Big) = 0, \tag{12c}$$

$$\text{and } V_2(T,x) = s_2(x(T)), \forall (T,x) \in \Sigma, \tag{12d}$$

where the Hamiltonian operator $\mathcal{H}_2$ and intervention operator $\mathcal{R}$ are defined as follows:

$$\mathcal{H}_2\left(x, \gamma^*(t,x), \frac{\partial V_2(t,x)}{\partial x}\right) = h_2(x, \gamma^*(t,x)) + \left(\frac{\partial V_2(t,x)}{dx}\right)^T f(x, \gamma^*(t,x)), \tag{12e}$$

$$\mathcal{R}V_2(t,x) = \min_{\eta \in \Omega_2} V_2(t, x + g(x,\eta)) + b_2(x,\eta). \tag{12f}$$

Remark 5 *QVIs can be interpreted as follows:*

1. *Condition (12b) ensures that the value function $V_2(\cdot)$ evaluated at any $(t,x) \in \Sigma$ is at most equal to the minimum cost that Player 2 incurs from intervening at time t and playing optimally afterwards.*
2. *Player 2 does not intervene at a time t if the cost-to-go from giving an impulse at time t is strictly greater than the value function $V_2(\cdot)$ evaluated at $(t,x) \in \Sigma$. Hence, when $V_2(t,x) = \mathcal{R}V_2(t,x)$, Player 2 gives an impulse.*
3. *At any $(t,x) \in \Sigma$, condition (12c) ensures that either Player 2 waits so that the HJB-like Equation (12a) for Player 2 holds with equality or Player 2 gives an impulse.*

Remark 6 *The value function of Player 2, $V_2(t,x)$, can have kinks at those time instants when the state value is at the boundary of the continuation set $\mathcal{C}$. In (single-agent) impulse control problems, the value function is assumed to be differentiable throughout the time horizon (see [11, 12], and the references therein).*

Remark 7 *The condition $V_2(\tau,x) = \mathcal{R}V_2(\tau,x)$ results in the continuity of the value function of Player 2 at the impulse instant τ under the feedback information structure. For impulse control problems studied by using the Pontryagin maximum principle, the Hamiltonian continuity condition [18] gives the timing of interventions (see also [2], where differential games with impulse control are analyzed using the impulse version of the Pontryagin maximum principle).*

QVIs allow us to define the continuation and intervention sets for Player 2 as follows:

Definition 3 *The continuation and intervention sets are given by*

$$\mathcal{C} = \Big\{(t,x) \in \Sigma | V_2(t,x) < \mathcal{R}V_2(t,x), \frac{\partial V_2(t,x)}{\partial t} + \mathcal{H}_2\left(x, \gamma^*(t,x), \frac{\partial V_2(t,x)}{\partial x}\right) = 0\Big\}, \tag{13}$$

$$\mathcal{I} = \Big\{(t,x) \in \Sigma | V_2(t,x) = \mathcal{R}V_2(t,x), \frac{\partial V_2(t,x)}{\partial t} + \mathcal{H}_2\left(x, \gamma^*(t,x), \frac{\partial V_2(t,x)}{\partial x}\right) \geq 0\Big\}. \tag{14}$$

[1]This assumption is also made in [11] to define (weak) QVIs (see (12a)).

Remark 8 *In impulse games studied in [12] and [13], the system of QVIs for any player j has an additional intervention operator to account for impulses by the other player(s), while the Hamiltonian operator is not an explicit function of the strategies of other player(s). In our game problem, the Hamiltonian operator of Player 2 depends on the strategies of Player 1, which in turn continuously affects the continuation and intervention sets of Player 2. Further, in the infinite-horizon impulse game studied in [12], the continuation sets depend only on the current state.*

Assumption 3 *There exists a unique measurable function $\zeta : [0,T] \times \mathbb{R}^n \to \Omega_2$ such that*

$$\zeta(t,x) = \arg\min_{\eta \in \Omega_2} \{V_2(t, x + g(x,\eta)) + b_2(x,\eta)\}. \tag{15}$$

Here, (15) gives the optimal impulse level at any (t,x) since it minimizes the sum of the immediate cost ($b_2(x,\eta)$) incurred from giving an impulse of size η and the cost-to-go from playing optimally afterwards (see also [12], where a similar assumption is used to solve stochastic impulse games).

We have the following assumptions regarding the state dynamics (1)–(2) and the objective functions described by (3)–(4):

Assumption 4 *The state dynamics and objective functions of Player 1 and Player 2 satisfy the following conditions:*

1. *$f(x,u)$ is (uniformly) Lipschitz continuous in x, that is, there exists a constant $c_f > 0$, such that*
$$|f(x,u) - f(y,u)| \le c_f|x-y|,\ \forall x,y \in \mathbb{R}^n,\ u \in \Omega_1.$$
2. *$g(x,\xi)$ is (uniformly) Lipschitz continuous in x, such that, for $c_g > 0$, we have*
$$|g(x,\xi) - g(y,\xi)| \le c_g|x-y|\ \forall x,y \in \mathbb{R}^n,\ \xi \in \Omega_2.$$
3. *$\forall x \in \mathbb{R}^n$, $\inf_{\eta \in \Omega_2} b_2(x,\eta) = \mu > 0$.*
4. *The functions f, g, h_i, b_i and s_i are bounded for $i \in \{1,2\}$.*

Assumptions 1.1 and 1.2 ensure that there exists a unique state trajectory $x(\cdot)$ for any measurable $u(\cdot)$ and impulse sequence $\{(\tau_i,\xi_i)\}_{i\ge1}$. Assumption 1.3 ensures that Player 2 intervenes only a finite number of times in the game due to the fixed cost associated with each impulse (see [11], where similar assumptions are provided in the context of an impulse optimal control problem). Assumption 1.4 is used later to show that the value functions of Player 1 and Player 2 have an upper and lower bound that depend on the problem parameters.

The sufficient conditions to characterize the FNE of the differential game described in (5)–(8) are given in the next theorem.

Theorem 1 (Verification Theorem) *Let Assumptions 1–4 hold. Suppose there exist functions $V_i : [0,T] \times \mathbb{R}^n \to \mathbb{R} (i=1,2)$ such that $V_1(t,x)$ satisfies the HJB equations (11) and $V_2(t,x)$ satisfies the QVIs (12) for all $(t,x) \in \Sigma$. Suppose there exist strategies (γ^*,δ^*) with the following properties. Player 1's control $u^*(t) = \gamma^*(t,x)$ satisfies for all $t \in [0,T]$*

$$u^*(t) = \gamma^*(t,x) = \arg\min_{\varphi \in \Omega_1} \mathcal{H}_1\left(x, \varphi, \frac{\partial V_1(t,x)}{\partial x}\right), \tag{16a}$$

and Player 2's impulse control is a sequence $\{(\tau_j^*,\xi_j^*)\}_{j\ge1}$ where interventions occur at $\tau_j^* = t$ if $(t,x) \in \mathcal{I}$, that is, (t,x) satisfy

$$V_2(t,x) = \mathcal{R}V_2(t,x), \tag{16b}$$

and the corresponding impulse levels ξ_j^* are given by

$$\xi_j^* = \zeta(t,x) = \arg\min_{\eta \in \Omega_2} \left(V_2(t, x+g(x,\eta)) + b_2(x,\eta)\right). \tag{16c}$$

Then, (γ^*,δ^*) is a FNE of the differential game described by (5–8). Further, $V_i(t,x)$ is the equilibrium cost-to-go of Player i, $(i=1,2)$ for the subgame starting at $(t,x) \in \Sigma$ and defined over the horizon $[t,T]$.

Proof. From Definition 2, we have to show that

$$\begin{aligned}
&V_j(t,x) = J_j(t,x,\gamma^*_{[t,T]},\delta^*_{[t,T]}),\ j=\{1,2\},\\
&V_1(t,x) \le J_1(t,x,\gamma_{[t,T]},\delta^*_{[t,T]}), \forall \gamma_{[t,T]} \in \Gamma_{[t,T]},\\
&V_2(t,x) \le J_2(t,x,\gamma^*_{[t,T]},\delta_{[t,T]}), \forall \delta_{[t,T]} \in \Delta_{[t,T]}.
\end{aligned}$$

Suppose $x_1(\cdot)$ is the state trajectory generated by Player 1 using an arbitrary admissible strategy $\gamma_{[t,T]}$ and Player 2 using the strategy $\delta^*_{[t,T]}$ such that Player 1's control $u(t)$ is given by $u(s) = \gamma_{[t,T]}(s,x(s))$, $s \in [t,T]$. Using the total derivative of $V_1(\cdot)$ between the impulse instants (τ^*_{j-1},τ^*_j), integrating with respect to t from τ^*_{j-1} to τ^*_j, and taking the summation for all $j \ge 1$, we obtain

$$\begin{aligned}
V_1(T,x_1(T)) - V_1(t,x) = &\sum_{j\ge 1}\int_{\tau^{*+}_{j-1}}^{\tau^{*-}_j} \Big(\frac{\partial V_1}{\partial t}(s,x_1(s)) + \left(\frac{\partial V_1}{\partial x}(s,x_1(s))\right)^T f(x_1(s),u(s))\Big)ds\\
&+\sum_{j\ge 1}\mathbb{1}_{t\le\tau^*_j<T}(V_1(\tau^{*+}_j,x_1(\tau^{*+}_j)) - V_1(\tau^{*-}_j,x_1(\tau^{*-}_j))),
\end{aligned}$$

where we defined $\tau^*_0 := t$. From (11a), we know that, for any given control $u(s)$, the following inequality holds:

$$\frac{\partial V_1}{\partial t}(s,x_1(s)) + \left(\frac{\partial V_1}{\partial x}(s,x_1(s))\right)^T f(x_1(s),u(s)) \ge -h_1(x_1(s),u(s)). \tag{17}$$

Therefore, we obtain

$$\begin{aligned}
V_1(T,x_1(T)) - V_1(t,x) \ge &-\sum_{j\ge 1}\int_{\tau^{*+}_{j-1}}^{\tau^{*-}_j} h_1(x_1(s),u(s))ds\\
&+\sum_{j\ge 1}\mathbb{1}_{t\le\tau^*_j<T}(V_1(\tau^{*+}_j,x_1(\tau^{*+}_j)) - V_1(\tau^{*-}_j,x_1(\tau^{*-}_j))).
\end{aligned}$$

From the terminal condition (11b) on $V_1(\cdot)$ and (11c), we obtain

$$\begin{aligned}
V_1(t,x) &\le s_1(x(T)) + \sum_{j\ge 1}\int_{\tau^{*+}_{j-1}}^{\tau^{*-}_j} h_1(x_1(s),u(s))ds + \sum_{j\ge 1}\mathbb{1}_{t\le\tau^*_j<T}\ b_1(x_1(\tau^{*-}_j),\xi^*_j)\\
&= J_1(t,x,\gamma_{[t,T]},\delta^*_{[t,T]}).
\end{aligned}$$

For a strategy γ^* of Player 1, (16a) holds for the equilibrium control $u^*(t)$ of Player 1, so we obtain

$$\begin{aligned}
V_1(t,x) &= s_1(x^*(T)) + \sum_{j\ge 1}\int_{\tau^{*+}_{j-1}}^{\tau^{*-}_j} h_1(x^*(s),u^*(s))ds + \sum_{j\ge 1}\mathbb{1}_{t\le\tau^*_j<T}\ b_1(x^*(\tau^{*-}_j),\xi^*_j)\\
&= J_1(t,x,{\gamma^*}_{[t,T]},\delta^*_{[t,T]}),
\end{aligned}$$

where x^* is the state trajectory generated by Player 1 choosing the strategy $\gamma^*_{[t,T]}$ and Player 2 choosing the strategy $\delta^*_{[t,T]}$. Therefore, $\gamma^*_{[t,T]}$ is the best response to Player 2's strategy $\delta^*_{[t,T]}$.

Next, we consider an arbitrary admissible strategy $\delta_{[t,T]}$ of Player 2 such that the intervention instants are given by τ_i, $i \ge 1$ and the corresponding impulse levels are given by ξ_i. Further, $x_2(\cdot)$ is the state trajectory generated by the strategy pairs $(\gamma^*_{[t,T]},\delta_{[t,T]})$. We obtain the following relation by taking the total derivative of $V_2(\cdot)$ between the impulse instants (τ_{j-1},τ_j), integrating over time from τ_{j-1} to τ_j, and taking the summation for all $j \ge 1$:

$$\begin{aligned}
V_2(T,x_2(T)) - V_2(t,x) = &\sum_{j\ge 1}\int_{\tau^{+}_{j-1}}^{\tau^{-}_j} \Big(\frac{\partial V_2(s,x_2(s))}{\partial s} + \left(\frac{\partial V_2(s,x_2(s))}{\partial x}\right)^T f(x_2(s),\gamma^*(s,x_2(s)))\Big)ds\\
&+\sum_{j\ge 1}\mathbb{1}_{t\le\tau_j<T}(V_2(\tau^{+}_j,x_2(\tau^{+}_j)) - V_2(\tau^{-}_j,x_2(\tau^{-}_j))).
\end{aligned} \tag{18}$$

The value function satisfies (12a) for all $(t, x) \in \Sigma$, so we have

$$\frac{\partial V_2}{\partial t}(s, x_2(s)) + \left(\frac{\partial V_2}{\partial x}(s, x_2(s))\right)^T f(x_2(s), \gamma^*(s, x_2(s))) \geq -h_2(x_2(s), \gamma^*(s, x_2(s)). \tag{19}$$

Given an impulse of size, ξ_j, and from the definition of an intervention operator given in (12f), we obtain

$$\mathcal{R}V_2(\tau_j^-, x_2(\tau_j^-)) \leq V_2(\tau_j^+, x_2(\tau_j^+)) + b_2(x_2(\tau_j^-), \xi_j).$$

Also, from (12b), we know that

$$\mathcal{R}V_2(\tau_j^-, x_2(\tau_j^-)) - V_2(\tau_j^-, x_2(\tau_j^-)) \geq 0.$$

Therefore, we obtain

$$V_2(\tau_j^+, x_2(\tau_j^+)) - V_2(\tau_j^-, x_2(\tau_j^-)) \geq \mathcal{R}V_2(\tau_j^-, x_2(\tau_j^-)) - V_2(\tau_j^-, x_2(\tau_j^-)) - b_2(x_2(\tau_j^-), \xi_j) \geq -b_2(x_2(\tau_j^-), \xi_j). \tag{20}$$

Substitute (19) and (20) in (18) to obtain

$$V_2(T, x_2(T)) - V_2(t, x) \geq \sum_{j \geq 1} \int_{\tau_{j-1}^+}^{\tau_j^-} -h_2(x_2(s), \gamma^*(s, x(s))ds - \sum_{j \geq 1} \mathbb{1}_{t \leq \tau_j < T}\ b_2(x(\tau_j^-), \xi_j).$$

Substituting the terminal condition $V_2(T, x_2(T)) = s_2(x_2(T))$, given in (12d), in the above inequality yields

$$\begin{aligned} V_2(t, x) \leq s_2(x_2(T)) &+ \sum_{j \geq 1} \int_{\tau_{j-1}^+}^{\tau_j^-} h_2(x_2(s), \gamma^*(s, x_2(s))ds \\ &+ \sum_{j \geq 1} \mathbb{1}_{t \leq \tau_j < T}\ b_2(x_2(\tau_j^-), \xi_j) = J_2(t, x, \gamma^*_{[t,T]}, \delta_{[t,T]}). \end{aligned}$$

The strategy δ^* of Player 2 entails giving impulses at $\tau_j^* = t$ where the pair $(t, x(t)) \in \Sigma$ is such that (12b) holds with equality, and the corresponding impulse levels ξ_j^* satisfy (16c). Therefore, for a strategy δ^*, we obtain

$$\begin{aligned} &V_2(\tau_j^{*+}, x^*(\tau_j^{*+})) - V_2(\tau_j^{*-}, x^*(\tau_j^{*-})) = -b_2(x^*(\tau_j^{*-}), \xi_j^*), \\ &\frac{\partial V_2}{\partial t}(s, x^*(s)) + \left(\frac{\partial V_2}{\partial x}(s, x^*(s))\right)^T f(x^*(s), \gamma^*(s, x^*(s))) = -h_2(x^*(s), \gamma^*(s, x^*(s)), \end{aligned}$$

and the cost-to-go function is given by

$$\begin{aligned} V_2(t, x) &= s_2(x^*(T)) + \sum_{j \geq 1} \int_{\tau_{j-1}^{*+}}^{\tau_j^{*-}} h_2(x^*(s), \gamma^*(s, x^*(s))ds + \sum_{j \geq 1} b_2(x^*(\tau_j^{*-}), \xi_j^*) \\ &= J_2(t, x, \gamma^*_{[t,T]}, \delta^*_{[t,T]}). \end{aligned}$$

Therefore, δ^* is the best response strategy to Player 1's strategy γ^*. □

Remark 9 *An important feature of the FNE solution introduced in Definition 2 is that if the strategy pair (γ^*, δ^*) provides a FNE to differential game described by (5–8) with duration $[0, T]$, then its restriction to the time interval $[t, T]$, denoted by $(\gamma^*_{[t,T]}, \delta^*_{[t,T]})$, provides a FNE to the same differential game defined on the shorter time interval $[t, T]$, with any initial state $x(t)$. Since, this property holds true for all $0 \leq t \leq T$ and for all state values $x(t)$, the FNE (γ^*, δ^*) is strongly time consistent.*

Next, we show that there can only be a finite number of impulses during the game.

Proposition 1 *Let Assumption 4 hold. Then, the value functions of Player 1 and Player 2 are bounded. The equilibrium number of impulses $K \in \mathbb{N}$ is bounded by*

$$K = \lceil \frac{2\,(T\|h_2\|_\infty + \|s_2\|_\infty)}{\mu} \rceil, \tag{21}$$

where $\mu = \inf_{\eta \in \Omega_2} b_2(x, \eta) > 0,\ \forall x \in \mathbb{R}^n$, and $\lceil y \rceil$ denotes the smallest integer that is greater than or equal to y.

Proof. See Appendix A.1. □

QVIs have been solved in the literature under some restrictive assumptions on the value functions, even for games with linear objective functions, see e.g., [12] and [30]. An additional difficulty in our case is that the QVIs are coupled with HJB equations associated with Player 1's best response. In the next section, we specialize our results to linear-quadratic differential games and provide a complete analytical characterization of the FNE strategies.

4 A scalar linear-quadratic differential game with targets

In this section, we consider a scalar linear-quadratic adaptation of the differential game (1-4), referred to as iLQDG hereafter. Player 1 and Player 2 aim to minimize the costs resulting from the deviation of the state away from their target state values ρ_1 and ρ_2, respectively. In our model, the structure of Player 2's problem (objective functions and state dynamics) is inspired by the impulse optimal control problem analyzed in [11].

(iLQDG):

$$J_1(0, x_0, u(\cdot), \tilde{v}) = \int_0^T \frac{1}{2}\left(w_1(x(t) - \rho_1)^2 + r_1 u(t)^2\right) dt + \sum_{i=1}^{k} z_1|\xi_i| + \frac{1}{2} s_1(x(T) - \rho_1)^2, \tag{22a}$$

$$J_2(0, x_0, u(\cdot), \tilde{v}) = \int_0^T \frac{1}{2} w_2(x(t) - \rho_2)^2 dt + \sum_{i=1}^{k} h(\xi_i) + \frac{1}{2} s_2(x(T) - \rho_2)^2, \tag{22b}$$

$$\dot{x}(t) = ax(t) + bu(t),\ x(0^-) = x_0,\ \forall t \neq \tau_i,\ i \geq 1, \tag{22c}$$

$$x(\tau_i^+) = x(\tau_i^-) + \xi_i, \tag{22d}$$

where

$$h(\xi_i) := \begin{cases} C + c\xi_i & \text{if } \xi_i > 0 \\ \min(C, D) & \text{if } \xi_i = 0 \\ D - d\xi_i & \text{if } \xi_i < 0, \end{cases} \tag{23}$$

and w_1, r_1, z_1, s_1, w_2, s_2, C, D, c, d are positive constants.

In the above iLQDG, the impulse can be positive, negative, or 0. Each intervention results in fixed costs, equal to C or D, for Player 2, even if the magnitude of the impulse at the intervention instant τ_i is 0. Player 1 incurs a positive cost $z_1|\xi_i|$ due to interventions by Player 2. We can view $z_1|\xi_i|$ as the cost associated with the disruption of Player 1's resources due to Player 2's actions.

The continuation set of Player 2 is described in the following assumption (see also [11] and [15]):

Assumption 5 *Player 2 gives an impulse if (t, x) does not lie in the continuation set $\mathcal{C}$ given by*

$$\mathcal{C} = \{(t, x) \in \Sigma \mid \ell_1(t) < x < \ell_2(t)\}. \tag{24}$$

Player 2 shifts the state to $\alpha(t)$ if $x \leq \ell_1(t)$, and to $\beta(t)$ if $x \geq \ell_2(t)$, so that the following relation holds:

$$\ell_1(t) < \alpha(t) < \beta(t) < \ell_2(t). \tag{25}$$

The threshold policy of Player 2 involves determining the boundaries $\ell_1(\cdot)$ and $\ell_2(\cdot)$ of the continuation region $\mathcal{C}$ as well as the values $\alpha(\cdot)$ and $\beta(\cdot)$, to which Player 2 shifts the state whenever the state reaches the boundaries $\ell_1(\cdot)$ or $\ell_2(\cdot)$, respectively. The functions $\ell_1(\cdot)$, $\alpha(\cdot)$, $\beta(\cdot)$, and $\ell_2(t)$ are obtained from the QVIs.

Assumption 6 *The state feedback strategy of Player 1 defined in the continuation set $\mathcal{C}$ is given by $\gamma(t,x) = p_1(t)x + q_1(t)$ where the real valued functions $p_1 : [0,T] \to \mathbb{R}$ and $q_1 : [0,T] \to \mathbb{R}$ are continuous.*

It is to be noted that the above assumption allows for discontinuities in the control of Player 1 at the impulse instants due to the corresponding jumps in the state. However, for a given state value in the continuation set $\mathcal{C}$, Player 1's strategy is continuous in t and x.

We make the following assumption on the admissible controls of Player 1 and Player 2:

Assumption 7 *The admissible control u of Player 1 and impulse size ξ for Player 2 lie in the interior of the bounded and open convex sets Ω_1 and Ω_2, respectively*

4.1 Optimal control problem of Player 1

Let the equilibrium strategy of Player 2 be given by δ^* such that Player 2 gives an impulse if the state leaves the continuation set $\mathcal{C}$ described in Assumption 5. Then, the equilibrium strategy of Player 1 can be determined by finding the value function that satisfies (11a)–(11c) for the iLQDG.

Player 1 solves a linear-quadratic optimal control problem in the continuation region $\mathcal{C}$, and at the impulse instant τ_i, Player 1's cost is given by $z_1|\xi_i|$. Therefore, we can make the following guess on the form of the value function of Player 1:

Assumption 8 *The value function of Player 1 is given by:*

$$V_1(t,x) = \begin{cases} \Phi_1(t,\alpha(t)) + z_1|\alpha(t) - x| & x \le \ell_1(t) \\ \Phi_1(t,x) & x \in (\ell_1(t), \ell_2(t)) \\ \Phi_1(t,\beta(t)) + z_1|\beta(t) - x| & x \ge \ell_1(t) \end{cases}. \tag{26}$$

Since the game is linear-quadratic, Φ_1 is quadratic in the state:

$$\Phi_1(t,x) = \frac{1}{2}p_1(t)x^2 + q_1(t)x + n_1(t). \tag{27}$$

The equilibrium control of Player 1 is obtained by substituting the value function in the HJB equation. From (11a), we have

$$-\frac{\partial \Phi_1(t,x)}{\partial t} = \min_{u\in\Omega_u} \Big(\frac{1}{2}w_1(x-\rho_1)^2 + \frac{1}{2}r_1 u(t)^2 + \left(\frac{\partial \Phi_1}{\partial x}\right)(ax + bu(t))\Big). \tag{28}$$

Differentiating the right-hand side of the above equation and equating the result to zero yields the equilibrium strategy of Player 1 (see Assumption 7):

$$\gamma^*(t,x) = u^*(t) = -\frac{b}{r_1}\left(\frac{\partial \Phi_1}{\partial x}\right) = -\frac{b}{r_1}(p_1(t)x + q_1(t)). \tag{29}$$

Substituting (29) in the state dynamics (22c), we obtain

$$\begin{aligned} \dot{x}(t) &= ax(t) + bu^*(t) = ax(t) - \frac{b^2}{r_1}\left(p_1(t)x(t) + q_1(t)\right) \\ &= \left(a - \frac{b^2}{r_1}p_1(t)\right)x(t) - \frac{b^2}{r_1}q_1(t) \\ &= a_x(t)x(t) + b_x q_1(t), \end{aligned} \tag{30}$$

where $a_x(t) = a - \frac{b^2}{r_1}p_1(t)$ and $b_x = -\frac{b^2}{r_1}$. On substituting (29) and (27) in (28), we obtain

$$
\begin{aligned}
&-\frac{1}{2}\dot{p}_1(t)x^2 - \dot{q}_1(t)x - \dot{n}_1(t) = \frac{1}{2}w_1(x-\rho_1)^2 - \frac{1}{2}b_x(p_1(t)x + q_1(t))^2 + (p_1(t)x+q_1(t))\,(a_x(t)x + b_x q_1(t)) \\
&\Rightarrow -\dot{p}_1(t)x^2 - 2\dot{q}_1(t)x - 2\dot{n}_1(t) = w_1x^2 + w_1\rho_1^2 - 2xw_1\rho_1 - b_x\left(p_1(t)^2x^2 - q_1(t)^2\right) + 2a_x(t)\,(p_1(t)x+q_1(t))\,x.
\end{aligned}
$$

Upon rearranging a few terms in the above equation, we get

$$\left(\dot{p}_1(t) + w_1 + b_x p_1(t)^2 + 2p_1(t)a\right)x^2 + w_1\rho_1^2 + 2\dot{n}_1(t) + b_x q_1(t)^2 + \left(2\dot{q}_1(t) + 2a_x(t)q_1(t) - 2w_1\rho_1\right)x = 0.$$

Since the above equation must hold for all x except at $(t,x) \notin \mathcal{C}$, $p_1(\cdot)$, $q_1(\cdot)$, and $n_1(\cdot)$ evolve as follows:

$$\dot{p}_1(t) = -w_1 - b_x p_1(t)^2 - 2p_1(t)a, \tag{31a}$$

$$\dot{q}_1(t) = -a_x(t)q_1(t) + w_1\rho_1, \tag{31b}$$

$$\dot{n}_1(t) = -\frac{1}{2}b_x q_1(t)^2 - \frac{w_1\rho_1^2}{2}, \tag{31c}$$

where $p_1(T) = s_1$, $q_1(T) = -s_1\rho_1$, and $n_1(T) = \frac{1}{2}s_1\rho_1^2$.

The solution of (31a) is given by the following equation (see Appendix A.3):

$$p_1(t) = \frac{1}{b_x}\left(-a + \frac{\theta}{2} - \frac{\theta}{C_1e^{\theta t}+1}\right). \tag{32}$$

Using the value of $p_1(t)$ given in (32), we obtain

$$a_x(t) = a + b_x p_1(t) = \frac{\theta}{2} - \frac{\theta}{C_1e^{\theta t}+1}.$$

Proposition 2 *Let Assumptions 5-8 hold. Then, the equilibrium state-feedback strategy of Player 1 is given by*

$$\gamma^*(t,x) = \left(\frac{\theta}{2} - \frac{\theta}{C_1e^{\theta t}+1}\right)x - \frac{b^2}{r_1}q_1(t),\ \forall (t,x) \in \mathcal{C}, \tag{33}$$

where

$$\theta = 2\sqrt{a^2 + w_1\frac{b^2}{r_1}}, \tag{34}$$

$$C_1 = \left(\frac{2\theta}{\theta + 2\frac{b^2}{r_1}s_1 - 2a} - 1\right)e^{-\theta T}. \tag{35}$$

When an impulse occurs, that is, $(\tau_i^*, x(\tau_i^*)) \in \mathcal{I}$, it follows from (11c) that V_1 satisfies

$$\frac{1}{2}p_1(\tau_i^{*-})x^2 + q_1(\tau_i^{*-})x + n_1(\tau_i^{*-}) = \frac{1}{2}p_1(\tau_i^{*+})(x+\xi_i^*)^2 + q_1(\tau_i^{*+})(x+\xi_i^*) + n_1(\tau_i^{*+}) + z_1|\xi_i^*|.$$

The equilibrium strategy of Player 2 is to bring the state to $\alpha(t)$ if $x \le \ell_1(t)$, and to $\beta(t)$ if $x \ge \ell_2(t)$, that is, $x + \xi_i^* = \alpha(\tau_i^{*-})$ if $x \le \ell_1(t)$ and $x + \xi_i^* = \beta(\tau_i^{*-})$ if $x \ge \ell_2(t)$. Therefore, we have

$$
\begin{aligned}
\frac{1}{2}p_1(\tau_i^{*-})x^2 + q_1(\tau_i^{*-})x + n_1(\tau_i^{*-}) &= \frac{1}{2}p_1(\tau_i^{*+})\alpha(\tau_i^{*-})^2 + q_1(\tau_i^{*+})\alpha(\tau_i^{*-}) + n_1(\tau_i^{*+}) \\
&\quad + z_1|\alpha(\tau_i^{*-}) - x|,\ x \le \ell_1(t), \\
\frac{1}{2}p_1(\tau_i^{*-})x^2 + q_1(\tau_i^{*-})x + n_1(\tau_i^{*-}) &= \frac{1}{2}p_1(\tau_i^{*+})\beta(\tau_i^{*-})^2 + q_1(\tau_i^{*+})\beta(\tau_i^{*-}) + n_1(\tau_i^{*+}) \\
&\quad + z_1|\beta(\tau_i^{*-}) - x|,\ x \ge \ell_2(t).
\end{aligned}
$$

Since $\xi_i^* = \alpha(\tau_i^{*-}) - x > 0$ and $\xi_i^* = \beta(\tau_i^{*-}) - x < 0$, we have

$$\begin{aligned}
\frac{1}{2}p_1(\tau_i^{*-})x^2 + q_1(\tau_i^{*-})x + n_1(\tau_i^{*-}) &= \frac{1}{2}p_1(\tau_i^{*+})\alpha(\tau_i^{*-})^2 + (z_1 + q_1(\tau_i^{*+}))\alpha(\tau_i^{*-}) + n_1(\tau_i^{*+}) \\
&\quad - z_1 x,\ x \le \ell_1(t), \\
\frac{1}{2}p_1(\tau_i^{*-})x^2 + q_1(\tau_i^{*-})x + n_1(\tau_i^{*-}) &= \frac{1}{2}p_1(\tau_i^{*+})\beta(\tau_i^{*-})^2 + (-z_1 + q_1(\tau_i^{*+}))\beta(\tau_i^{*-}) + n_1(\tau_i^{*+}) \\
&\quad + z_1 x,\ x \ge \ell_2(t).
\end{aligned}$$

The above equations and continuity of p_1 and q_1 (from Assumption 6) imply that, at the impulse instants, the following conditions are satisfied:

$$\begin{aligned}
n_1(\tau_i^{*-}) &= n_1(\tau_i^{*+}) + \frac{1}{2}p_1(\tau_i^*)\alpha(\tau_i^{*-})^2 - \frac{1}{2}p_1(\tau_i^*)x^2 + (z_1 + q_1(\tau_i^*))\alpha(\tau_i^{*-}) - (q_1(\tau_i^{*-}) + z_1)x, x \le \ell_1(t), \\
n_1(\tau_i^{*-}) &= n_1(\tau_i^{*+}) + \frac{1}{2}p_1(\tau_i^*)\beta(\tau_i^{*-})^2 - \frac{1}{2}p_1(\tau_i^*)x^2 - (z_1 - q_1(\tau_i^*))\beta(\tau_i^{*-}) - (q_1(\tau_i^{*-}) - z_1)x, x \ge \ell_2(t).
\end{aligned}$$

4.2 Impulse control problem of Player 2

Player 2 solves the QVIs associated to her impulse control problem for a given equilibrium strategy γ^* of Player 1.

In the continuation region, Player 2's running cost is quadratic in the state, and it is is linear in the state in the intervention region. Therefore, we can make the following conjecture on the form of the value function of Player 2:

Assumption 9 *The value function of Player 2 is given by*

$$V_2(t,x) = \begin{cases} \Phi_2(t,\alpha(t)) + C + c(\alpha(t) - x) & x \le \ell_1(t) \\ \Phi_2(t,x) & x \in (\ell_1(t), \ell_2(t)) \\ \Phi_2(t,\beta(t)) + D + d(x - \beta(t)) & x \ge \ell_2(t), \end{cases} \tag{36}$$

where

$$\Phi_2(t,x)b = \frac{1}{2}p_2(t)x^2 + q_2(t)x + n_2(t). \tag{37}$$

A similar assumption on the form of the value function was made in [11] to obtain analytical solutions for an impulse optimal control problem.

The value function V_2 coincides with continuous and continuously differentiable function Φ_2 in the continuation region $\mathcal{C}$. We conjecture that Φ_2 is quadratic in state because the cost functions are quadratic in state. In the intervention region, the value function is equal to the sum of the intervention cost incurred by the player to shift the state to the continuation region and the cost-to-go (which is equal to $\Phi_2(t,\alpha(t)$ or $\Phi_2(t,\beta(t)$ depending on the state value at the impulse time) from playing optimally afterwards.

When the state lies in the continuation region, that is, $x \in (\ell_1(t), \ell_2(t))$, the value function of Player 2 satisfies (12a) with equality

$$\frac{\partial \Phi_2(t,x)}{\partial t} + \left(\frac{\partial \Phi_2}{\partial x}\right)(ax + b\gamma^*(t,x)) + \frac{1}{2}w_2(x-\rho_2)^2 = 0.$$

Substituting the partial derivatives of $\Phi_2(t,x)$ and the equilibrium control of Player 1 from (29) in the above equation yields

$$\begin{aligned}
&\frac{1}{2}\dot{p}_2(t)x^2 + \dot{q}_2(t)x + \dot{n}_2(t) + (p_2(t)x + q_2(t))\, a_x(t)x - w_2 x\rho_2 \\
&\quad + b_x q_1(t)\,(p_2(t)x + q_2(t)) + \frac{1}{2}w_2x^2 + \frac{1}{2}w_2\rho_2^2 = 0.
\end{aligned}$$

On comparing the coefficients, we obtain

$$\dot{p}_2(t) = -w_2 - 2p_2(t)\left(\frac{\theta}{2} - \frac{\theta}{C_1e^{\theta t}+1}\right), \tag{38a}$$

$$\dot{q}_2(t) = -\left(\frac{\theta}{2} - \frac{\theta}{C_1e^{\theta t}+1}\right)q_2(t) - b_xp_2(t)q_1(t) + w_2\rho_2, \tag{38b}$$

$$\dot{n}_2(t) = -b_xq_1(t)q_2(t) - \frac{1}{2}w_2\rho_2^2, \tag{38c}$$

where $p_2(T) = s_2$, $q_2(T) = -s_2\rho_2$, and $n_2(T) = \frac{1}{2}s_2\rho_2^2$.

The solution of (38a) is given by

$$p_2(t) = \frac{-w_2e^{2t\theta}C_1^2 - 2t\theta w_2e^{t\theta}C_1 + w_2 + H\theta e^{t\theta}}{\theta(C_1e^{t\theta}+1)^2}, \tag{39}$$

where

$$H = 2C_1s_2 + s_2e^{-T\theta} - \frac{w_2e^{-T\theta} - C_1^2w_2e^{T\theta}}{\theta} + C_1^2s_2e^{T\theta} + 2C_1Tw_2. \tag{40}$$

We make the following assumption on the problem parameters so that $p_2 > 0$ for all $t \in [0, T]$, and consequently, the value function of Player 2 is strictly convex in the continuation region $\mathcal{C}$.

Assumption 10 *For $t \in [0, T]$, the problem parameters satisfy*

$$H\theta + w_2(1 - \mathrm{e}^{t\theta}C_1^2 - 2t\theta C_1) > 0. \tag{41}$$

4.2.1 Intervention set and continuation set

In the intervention region ($(t, x) \in \mathcal{I}$), (12b) holds with equality, that is,

$$V_2(t, x) = \mathcal{R}V_2(t, x) = \min_{\eta\in\Omega_2}(V_2(t, x+\eta) + h(\eta)). \tag{42}$$

For the problem parameters assumed in this section, V_2 is strictly convex in x (see Assumption 10) and continuously differentiable for $y = x + \eta \in \mathcal{C}$. Since $\alpha(t), \beta(t) \in \mathcal{C}$, and $x + \eta$ takes a value of $\alpha(t)$ or $\beta(t)$ at the intervention instants and the derivative of y with respect to η is equal to 1, we can use the first-order conditions to obtain

$$\frac{\partial\Phi_2(t,\alpha(t))}{\partial y} + \frac{\partial h(\eta)}{\partial\eta} = 0,\ x \le \ell_1(t), \tag{43}$$

$$\frac{\partial\Phi_2(t,\beta(t))}{\partial y} + \frac{\partial h(\eta)}{\partial\eta} = 0,\ x \ge \ell_2(t). \tag{44}$$

Using the quadratic form of the value function in (37) for the state value in the continuation region $(\ell_1(t), \ell_2(t))$, we get

$$\begin{aligned}\frac{\partial\Phi_2}{\partial y}(t,\alpha(t)) = p_2(t)\alpha(t) + q_2(t) = -c,\\ \frac{\partial\Phi_2}{\partial y}(t,\beta(t)) = p_2(t)\beta(t) + q_2(t) = d.\end{aligned} \tag{45}$$

Therefore, the following functions $\alpha(\cdot)$ and $\beta(\cdot)$ give the state values after an impulse occurs at equilibrium:

$$\alpha(t) = -\frac{q_2(t)+c}{p_2(t)}, \forall t \in [0, T], \tag{46a}$$

$$\beta(t) = \frac{d - q_2(t)}{p_2(t)},\ \forall t \in [0, T]. \tag{46b}$$

Since (12b) holds with equality in the intervention region, we have

$$V_2(t,x) = \begin{cases} V_2(t,\alpha(t)) + C + c(\alpha(t) - x) & x \le \ell_1(t) \\ V_2(t,\beta(t)) + D + d(x - \beta(t)) & x \ge \ell_2(t). \end{cases} \tag{47}$$

Also, $\alpha(t)$ and $\beta(t)$ lie in the continuation region $\mathcal{C}$, which implies $V_2(t,\alpha(t)) = \Phi_2(t,\alpha(t))$ and $V_2(t,\beta(t)) = \Phi_2(t,\beta(t))$. For $x = \ell_1(t)$ and $x = \ell_2(t)$, we substitute (37) in the above equations and simplify to obtain

$$\frac{1}{2}p_2(t)\ell_1(t)^2 + q_2(t)\ell_1(t) = \frac{1}{2}p_2(t)\alpha(t)^2 + q_2(t)\alpha(t) + C + c(\alpha(t) - \ell_1(t)), \tag{48a}$$

$$\frac{1}{2}p_2(t)\ell_2(t)^2 + q_2(t)\ell_2(t) = \frac{1}{2}p_2(t)\beta(t)^2 + q_2(t)\beta(t) + D + d(\ell_2(t) - \beta(t)). \tag{48b}$$

To characterize the left boundary of the continuation region, we substitute $\alpha(t)$ in (48a) to get

$$p_2(t)\ell_1(t)^2 + 2(q_2(t)+c)\ell_1(t) - p_2(t)\left(-\frac{q_2(t)+c}{p_2(t)}\right)^2 - 2(q_2(t)+c)\left(-\frac{q_2(t)+c}{p_2(t)}\right) - 2C = 0$$

$$\Rightarrow p_2(t)\ell_1(t)^2 + 2(q_2(t)+c)\ell_1(t) + \frac{(q_2(t)+c)^2}{p_2(t)} - 2C = 0.$$

Since $C > 0$, $p_2(t) > 0$, and $\ell_1(t) < \alpha(t)$, the left boundary of the continuation region is given by

$$\ell_1(t) = \frac{-c - q_2(t) - \sqrt{2Cp_2(t)}}{p_2(t)}. \tag{49a}$$

On substituting $\beta(t)$ in (48b), we obtain the right boundary of the continuation region

$$p_2(t)\ell_2(t)^2 + 2(q_2(t)-d)\ell_2(t) - p_2(t)\left(\frac{d-q_2(t)}{p_2(t)}\right)^2 - 2(q_2(t)-d)\left(\frac{d-q_2(t)}{p_2(t)}\right) - 2D = 0$$

$$\Rightarrow p_2(t)\ell_2(t)^2 + 2(q_2(t)-d)\ell_2(t) + \frac{(d-q_2(t))^2}{p_2(t)} - 2D = 0.$$

From $D > 0, p_2(t) > 0$ and $\ell_2(t) > \beta(t)$, we obtain

$$\ell_2(t) = \frac{-q_2(t) + d + \sqrt{2Dp_2(t)}}{p_2(t)}. \tag{49b}$$

By construction, $V_1(t,x)$ satisfies the sufficient conditions in (11), and therefore, V_1 is a value function of Player 1. In the next theorem, we give conditions under which $V_2(t,x)$ in (36) satisfies the QVIs (12).

Theorem 2 *Let Assumptions 5–10 hold.* $V_2(t,x)$ *in* (36) *is the value function of Player 2 if* $\ell_1(t) \le x_{11}(t)$ *and* $\ell_2(t) \ge x_{22}(t)$ *for each* $t \in [0,T]$ *where* $\ell_1(t)$ *and* $\ell_2(t)$ *are given in* (49a) *and* (49b), *respectively,*

$$x_{11}(t) = \frac{(ca + w_2\rho_2) - \sqrt{\theta_\alpha(t)}}{w_2}, \tag{50a}$$

$$x_{22}(t) = \frac{-(da - w_2\rho_2) + \sqrt{\theta_\beta(t)}}{w_2}, \tag{50b}$$

$$\theta_\alpha(t) = c^2a^2 + 2w_2\left(ca\rho_2 - \frac{\partial\Phi_2(t,\alpha(t))}{\partial t}\right), \tag{50c}$$

$$\theta_\beta(t) = d^2a^2 - 2w_2\left(da\rho_2 - \frac{\partial\Phi_2(t,\beta(t))}{\partial t}\right), \tag{50d}$$

and $x_{11}(t)$ *and* $x_{22}(t)$ *are well defined with* $\theta_\alpha(t) \ge 0$ *and* $\theta_\beta(t) \ge 0$ *for all* $t \in [0,T]$.

Proof. See Appendix A.2. □

5 Numerical examples

To illustrate our results, we consider an iLQDG with time horizon $T = 1$ and other problem parameters given in Table 1.

Table 1: Parameters for numerical example

a	b	w_1	s_1	r_1	z_1	w_2	s_2	c	C	D	d	ρ_1	ρ_2
0.1	−0.3	1	1	1	2	4	1	2	3	5	3	2.5	5

In Figure 1, we provide a complete characterization of the state feedback policy of Player 2 for the problem parameters in Table 1. Player 2 gives an impulse at any time t if the state reaches a level $\ell_1(t)$ or lower and brings the state to $\alpha(t)$. If the state reaches a level $\ell_2(t)$ or higher, then Player 2 gives an impulse to bring the state to $\beta(t)$. Since the cost coefficient s_2 of the salvage value for Player 2 is lower than the running cost coefficient w_2, the functions $\ell_2(t)$, $\beta_2(t)$, $\ell_1(t)$, and $\alpha(t)$ diverge over time away from the target state value $\rho_2 = 5$. Also, the fixed cost and the marginal cost of intervention are small if the state crosses the lower boundary compared to the case when the state crosses the upper boundary. As a result, $|\ell_1(t) - \alpha(t)| < |\ell_2(t) - \beta(t)|$ for all $t \in [0, T]$. For initial state values of 2, 5, and 8, the evolution of equilibrium state trajectories is given by $x_1^*(t)$, $x_2^*(t)$, and $x_3^*(t)$, respectively; see Figure 1. The equilibrium strategies are strongly time consistent which implies that if the state deviates from the equilibrium path such that the state value $x(t)$ is below $\ell_1(t)$ or above $\ell_2(t)$ at any $t \in (0, T)$, Player 2 brings the state to $\alpha(t)$ or $\beta(t)$, respectively; this observation is illustrated in Figure 1.

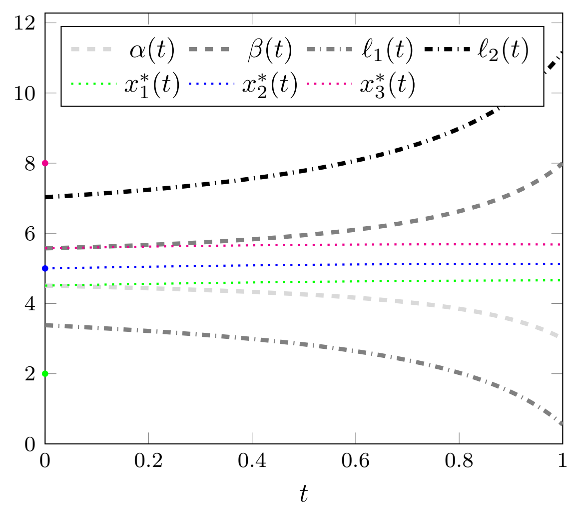

Figure 1: Evolution of the intervention region for the parameters in Table 1.

In Figure 2, we can see that the value functions of Player 1 and Player 2 at the initial time are quadratic in state when the state is in $(\ell_1(0), \ell_2(0))$, and that, outside this region, the value functions are linear in state. The value function of Player 1 jumps at $\ell_1(0)$ and $\ell_2(0)$ whereas Player 2's value function is continuous for all initial state values.

Next, we consider the case where the penalty associated with the state deviating from the target value at the terminal time is the same as the running cost. Therefore, in Figure 3, we can see that $\ell_1(\cdot)$, $\alpha(\cdot)$, $\beta(\cdot)$, and $\ell_2(\cdot)$ are a further away from the target state of Player 2 near the initial time, as compared to Figure 1. Here, $x_1^*(t)$, $x_2^*(t)$, and $x_3^*(t)$ denote the equilibrium evolution of the state trajectory for initial state values of 1, 6, and 10, respectively. The value functions of Player 1 and Player 2 at the initial time are given in Figure 4 for different values of the initial state.

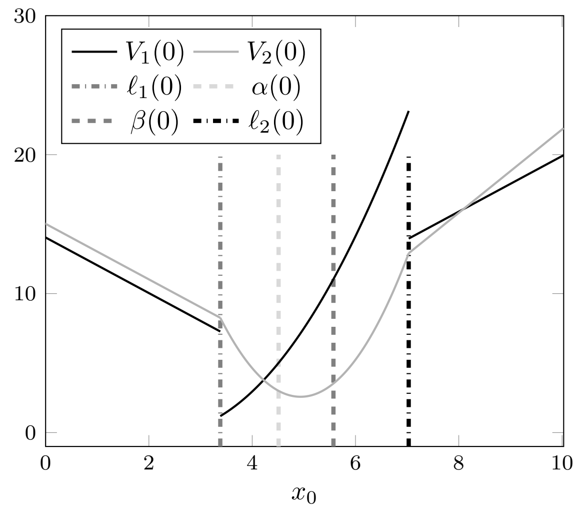

Figure 2: Value function for the parameters in Table 1.

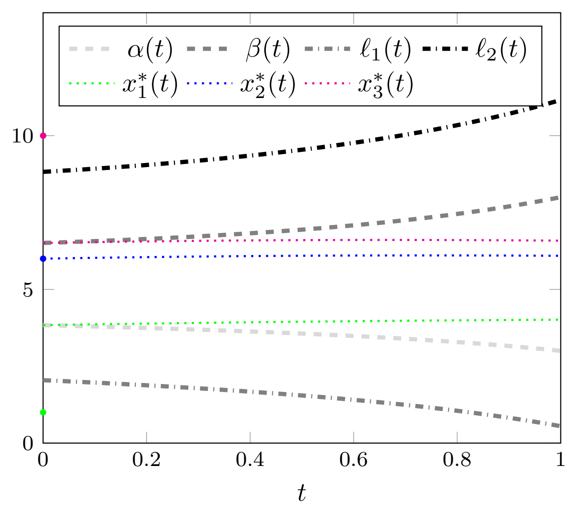

Figure 3: Evolution of the intervention region for the parameters in Table 1 with $w_2 = 1$.

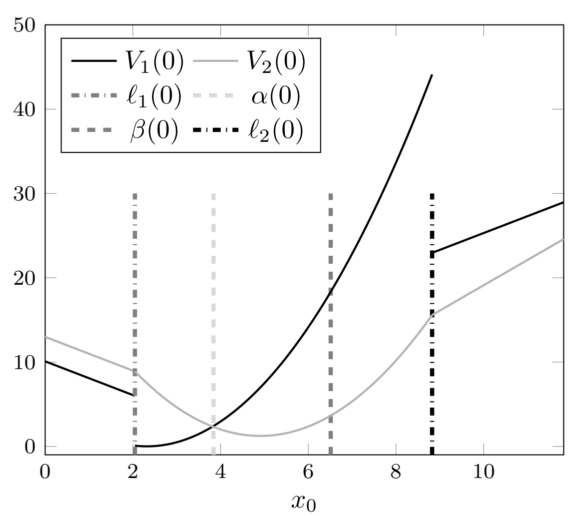

Figure 4: Value function for the parameters in Table 1 with $w_2 = 1$.

6 Conclusions

In this paper, we considered a two-player finite-horizon nonzero-sum differential game where Player 1 uses piecewise-continuous controls and Player 2 uses impulse controls. We determined an upper bound on the equilibrium number of impulses and provided sufficient conditions to characterize the feedback Nash equilibria for this general class of differential games with impulse controls. The sufficient conditions are given as a coupled system of Hamilton-Jacobi-Bellman equations with jumps and quasi-variational inequalities. To the best of our knowledge, this is the first characterization of feedback Nash equilibrium in differential games with impulse controls where at least one player uses piecewise-continuous controls. In this, our paper also differs from earlier papers on impulse games where equilibrium solutions were derived for problems in which both players use impulse controls only. Furthermore, we extended a well-studied linear-quadratic impulse control problem to a game setting where both players use their controls to minimize the cost associated with the state deviating from their target values.

We obtained closed-form solutions for the feedback Nash equilibrium in the scalar linear-quadratic differential game based on certain regularity assumptions on the value function that have been assumed in the literature (see e.g., [11] and [12]). In future work, we plan to relax these assumptions and develop policy iteration-type algorithms [33] that can solve the quasi-variational inequalities for the impulse player in the general class of differential games with impulse control.

A Appendix

A.1 Proof of Proposition 1

A feasible strategy of Player 2 is not to give any impulse in $[0, T]$ so that

$$\sum_{j\geq 1} \mathbb{1}_{t\leq \tau_j \leq T}\ b_2(x(\tau_j), \xi_j) = 0, \tag{51}$$

and it follows from the boundedness of h_2 and s_2 in Assumption 4 that

$$V_2(t,x) \leq \int_t^T h_2(x(s), \gamma^*(s, x(s)))ds + s_2(x(T)) \leq \|h_2\|_\infty (T-t) + \|s_2\|_\infty.$$

Next, for any $\epsilon > 0$, we choose a strategy $\delta_{[t,T]} \in \Delta_{[t,T]}$ so that

$$V_2(t,x) + \epsilon > J_2(x, \gamma^*_{[t,T]}, \delta_{[t,T]}) \geq -\|h_2\|_\infty (T-t) - \|s_2\|_\infty,$$

where the second inequality follows from Assumption 4. This proves that the value function is bounded such that

$$|V_2(t,x)| \leq \|h_2\|_\infty (T-t) + \|s_2\|_\infty,\ \forall (t,x) \in [0,T] \times \mathbb{R}^n. \tag{52}$$

For any $\epsilon > 0$, consider an ϵ-optimal strategy v^ϵ with $N(v^\epsilon)$ impulses. From the boundedness of h_2, we obtain

$$V_2(t,x) + \epsilon > J_2(x, \gamma_1, v^\epsilon) \geq -\|h_2\|_\infty (T-t) + \mu N(v^\epsilon) - \|s_2\|_\infty.$$

Using the above relation and (52), we obtain

$$-\|h_2\|_\infty (T-t) + \mu N(v^\epsilon) - \|s_2\|_\infty < \|h_2\|_\infty (T-t) + \|s_2\|_\infty + \epsilon.$$

Since $\mu > 0$, we can rewrite the above inequality as follows:

$$N(v^\epsilon) < \frac{2\left(\|h_2\|_\infty (T-t) + \|s_2\|_\infty\right) + \epsilon}{\mu}.$$

Since $\epsilon > 0$ is arbitrarily chosen for an ϵ-optimal strategy of Player 2, the upper bound K on the number of impulses is given by (21) as $\epsilon \to 0$.

For a feasible strategy of Player 1 given by $\gamma(t,x) = 0$ for all $(t,x) \in \Sigma$ and the upper bound K on the number of impulses, we have

$$\begin{aligned} V_1(t,x) &\leq \int_t^T h_1(x(s),0)ds + \sum_{j\geq 1} \mathbb{1}_{t\leq \tau_j < T}\ b_1(x(\tau_i),\xi_j) + s_1(x(T)) \\ &\leq \int_t^T h_1(x(s),0)ds + K\|b_1\|_\infty + s_1(x(T)) \\ &\leq \|h_1\|_\infty (T-t) + K\|b_1\|_\infty + \|s_1\|_\infty, \end{aligned}$$

where the last inequality follows from the boundedness of b_1 and s_1 in Assumption 4. For any $\epsilon > 0$, we take a strategy $\gamma_{[t,T]} \in \Gamma_{[t,T]}$ so that

$$V_1(t,x) + \epsilon > J_1(x, \gamma_{[t,T]}, \delta^*_{[t,T]}) \geq -\|h_1\|_\infty (T-t) - K\|b_1\|_\infty - \|s_1\|_\infty.$$

This proves that the value function of Player 1 is bounded.

A.2 Proof of Theorem 2

From (45), we have $\frac{\partial V_2(t,\alpha(t))}{\partial x} = -c$ and $\frac{\partial V_2(t,\beta(t))}{\partial x} = d$. Using the strict convexity of V_2 in x for $(t,x) \in \mathcal{C}$ (Assumption 10), we obtain

$$-c < \frac{\partial V_2(t,x)}{\partial x} < d,\ \forall (t,x) : x \in (\alpha(t), \beta(t)).$$

Therefore, $\mathcal{R}V_2(t,x) = \Phi_2(t,x) + \min(C,D)$ when the time and state pairs (t,x) are such that $x \in (\alpha(t), \beta(t))$.

When $x \in (\ell_1(t), \alpha(t))$, we have $\frac{\partial V_2(t,x)}{\partial x} \leq -c$ and, for $x \in (\beta(t), \ell_2(t))$, we obtain $\frac{\partial V_2(t,x)}{\partial x} \geq d$ from the strict convexity of $V_2(t,x)$ in $x \in (\ell_1(t), \ell_2(t))$. Therefore, the operator $\mathcal{R}$ satisfies the following system:

$$\mathcal{R}V_2(t,x) = \begin{cases} \Phi_2(t,\alpha(t)) + C + c(\alpha(t) - x) & x \leq \alpha(t) \\ \Phi_2(t,x) + \min(C,D) & x \in (\alpha(t), \beta(t)) \\ \Phi_2(t,\beta(t)) + D + d(x - \beta(t)) & x \geq \beta(t). \end{cases} \tag{53}$$

Clearly, $V_2 - \mathcal{R}V_2 < 0$ in the continuation region and $V_2(t,x) = \mathcal{R}V_2(t,x)$ in the intervention region.

Next, we derive the conditions under which the value function of Player 2 satisfies (12a). For $x < \ell_1(t)$, we have

$$V_2(t,x) = \Phi_2(t,\alpha(t)) + C + c(\alpha(t) - x). \tag{54}$$

When $x < \ell_1(t)$, we obtain

$$\begin{aligned} &\frac{\partial V_2(t,x)}{\partial t} + \mathcal{H}_2(x, \gamma^*(t,x), \frac{\partial V_2(t,x)}{\partial x}) \\ &\quad= \frac{\partial V_2(t,x)}{\partial t} + \frac{1}{2} w_2 (x-\rho_2)^2 + \frac{\partial V_2(t,x)}{\partial x}(ax + \mathbb{1}_{\ell_1(t)<x<\ell_2(t)}\ b\gamma^*(t,x)) \\ &\quad= \frac{\partial \Phi_2(t,\alpha(t))}{\partial t} + \left(\frac{\partial \Phi_2(t,\alpha(t))}{\partial x} + c\right)\frac{d\alpha(t)}{dt} - cax + \frac{1}{2}w_2 x^2 + \frac{1}{2} w_2 \rho_2^2 - w_2 x \rho_2. \end{aligned}$$

Substituting (45) in the above equation, we get the roots of the above equation as follows:

$$x_{11}(t), x_{12}(t) = \frac{(ca + w_2\rho_2) \pm \sqrt{\theta_\alpha(t)}}{w_2}, \tag{55}$$

where $x_{11}(t) < x_{12}(t)$, and $\theta_\alpha(t)$ is given by equation (50c). Therefore, (12a) holds if $\ell_1(t) \leq x_{11}(t)$ and $\theta_\alpha(t) \geq 0$ for all $t \in [0, T]$.

For $x > \ell_2(t)$, we obtain

$$\begin{aligned}
&\frac{\partial V_2(t,x)}{\partial t} + \mathcal{H}_2(x, \gamma^*(t,x), \frac{\partial V_2(t,x)}{\partial x}) \\
&\quad = \frac{\partial V_2(t,x)}{\partial t} + \frac{1}{2} w_2 (x-\rho_2)^2 + \frac{\partial V_2(t,x)}{\partial t}(ax + \mathbb{1}_{\ell_1(t)<x<\ell_2(t)} b\gamma^*(t,x)) \\
&\quad = \frac{\partial \Phi_2(t,\beta(t))}{\partial t} + \left(\frac{\partial \Phi_2(t,\beta(t))}{\partial x} - d\right)\frac{d\beta(t)}{dt} + dax + \frac{1}{2} w_2 x^2 + \frac{1}{2} w_2 \rho_2^2 - w_2 x \rho_2.
\end{aligned}$$

On substituting (45) in the above equation, we obtain the roots of the above equation as follows:

$$x_{21}(t), x_{22}(t) = \frac{-(da - w_2\rho_2) \pm \sqrt{\theta_\beta(t)}}{w_2}, \tag{56}$$

where $x_{21}(t) < x_{22}(t)$ and $\theta_\beta(t)$ is given by (50d). Therefore, (12a) holds if $\ell_2(t) \geq x_{22}(t)$.

A.3 Analytical solution of ODE

To solve the differential equation $\dot{p}_1(t) + b_x(p_1(t))^2 + 2ap_1(t) + w_1 = 0$ for $t \in (\tau_i, \tau_{i+1}), i \in \{0, 1, \cdots, k\}$, we substitute $p_1(t) = \frac{\dot{\mu}(t)}{b_x \mu(t)}$ to obtain a second-order ordinary differential equation $\ddot{\mu}(t) + 2a\dot{\mu}(t) + b_x w_1 \mu(t) = 0$. When $\theta = 2\sqrt{a^2 - w_1 b_x}$, the solution of this equation is

$$\mu(t) = e^{-at}(F_1 e^{\frac{1}{2}\theta t} + F_2 e^{-\frac{1}{2}\theta t}),$$

where F_1 and F_2 are constants. So, $p_1(t)$ is given by

$$\begin{aligned}
p_1(t) = \frac{\dot{\mu}(t)}{b_x \mu(t)} &= \frac{-a\mu(t) + \frac{\theta}{2} e^{-at}(F_1 e^{\frac{1}{2}\theta t} - F_2 e^{-\frac{1}{2}\theta t})}{b_x e^{-at}(F_1 e^{\frac{1}{2}\theta t} + F_2 e^{\frac{1}{2}\theta t})} \\
&= \frac{1}{b_x}\left(-a + \frac{\theta}{2} - \frac{\theta}{C_1 e^{\theta t} + 1}\right).
\end{aligned}$$

Substitute $p_1(T) = s_1$ in the above equation to obtain

$$C_1 = \left(\frac{2\theta}{\theta - 2b_x s_1 - 2a} - 1\right) e^{-\theta T}. \tag{57}$$

References

[1] T. Başar and G. Olsder, Dynamic Noncooperative Game Theory, 2nd edition. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1999.

[2] U. Sadana, P. V. Reddy, and G. Zaccour, Nash equilibria in non-zero sum differential games with impulse control, 2021. [Online]. Available: https://doi.org/10.1016/j.ejor.2021.03.025

[3] T. Başar, Time consistency and robustness of equilibria in noncooperative dynamic games. in Dynamic Policy Games in Economics, F. V. der Ploeg and A. de Zeeuw, Eds. North-Holland: Springer International Publishing, 2018, pp. 9–54.

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

[5] J. C. Engwerda, LQ Dynamic Optimization and Differential Games. New York: Wiley, 2005.

[6] T. Başar, A. Haurie, and G. Zaccour, Nonzero-Sum Differential Games, in Handbook of Dynamic Game Theory, T. Başar and G. Zaccour, Eds. Cham: Springer International Publishing, 2018, pp. 61–110.

[7] A. Bensoussan and J. L. Lions, Contrôle Impulsionnel et Inéquations Quasi-Variationelles. Paris, France: Bordas, 1982.

[8] ——, Impulse Control and Quasi-Variational Inequalities. Paris, France: Gauthier-Villars, 1984.

[9] J. P. Aubin, Mathematical Methods of Game and Economic Theory. Amsterdam, The Netherlands: North-Holland, 1982.

[10] A. Cadenillas and F. Zapatero, Optimal central bank intervention in the foreign exchange market, Journal of Economic Theory, 87(1):218–242, 1999.

[11] G. Bertola, W. J. Runggaldier, and K. Yasuda, On classical and restricted impulse stochastic control for the exchange rate, Applied Mathematics & Optimization, 74(2):423–454, Oct 2016.

[12] R. Aïd, M. Basei, G. Callegaro, L. Campi, and T. Vargiolu, Nonzero-sum stochastic differential games with impulse controls: A verification theorem with applications, Mathematics of Operations Research, 45(1):205–232, 2020.

[13] M. Basei, H. Cao, and X. Guo, Nonzero-sum stochastic games and mean-field games with impulse controls, Mathematics of Operations Research, 2021, (Forthcoming).

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

[15] W. J. Runggaldier and K. Yasuda, Classical and restricted impulse control for the exchange rate under a stochastic trend model, Journal of Economic Dynamics and Control, vol. 91, pp. 369–390, 2018, special Issue in Honour of Prof. Carl Chiarella.

[16] R. Korn, Portfolio optimisation with strictly positive transaction costs and impulse control, Finance and Stochastics, 2(2):85–114, Feb 1998.

[17] P. V. Reddy, S. Wrzaczek, and G. Zaccour, Quality effects in different advertising models - An impulse control approach, European Journal of Operational Research, 255(3):984–995, 2016.

[18] M. Chahim, R. Brekelmans, D. den Hertog, and P. Kort, An impulse control approach to dike height optimization, Optimization Methods and Software, 28(3):458–477, 2013.

[19] M. Chahim, D. Grass, R. F. Hartl, and P. M. Kort, Product innovation with lumpy investment, Central European Journal of Operations Research, 25(1):159–182, 2017.

[20] K. Erdlenbruch, A. Jean-Marie, M. Moreaux, and M. Tidball, Optimality of impulse harvesting policies, Economic Theory, 52(2):429–459, 2013.

[21] D. P. Berovic and R. B. Vinter, The application of dynamic programming to optimal inventory control, IEEE Transactions on Automatic Control, 49(5):676–685, 2004.

[22] V. Taynitskiy, E. Gubar, and Q. Zhu, Optimal Impulse Control of SIR Epidemics Over Scale-Free Networks, in Game Theory for Networking Applications, J. B. Song, H. Li, and M. Coupechoux, Eds. Cham: Springer International Publishing, 2019, pp. 119–129.

[23] A. A. Chikrii, I. I. Matychyn, and K. A. Chikrii, Differential games with impulse control, in Advances in Dynamic Game Theory. Annals of the International Society of Dynamic Games, S. Jørgensen, M. Quincampoix, and T. L. Vincent, Eds. Boston: Birkhäuser, 2007, 9:37–55.

[24] P. Azimzadeh, A zero-sum stochastic differential game with impulses, precommitment, and unrestricted cost functions, Applied Mathematics & Optimization, 79(2):483–514, 2019.

[25] U. Sadana, P. V. Reddy, T. Başar, and G. Zaccour, Sampled-data Nash equilibria in differential games with impulse control, Les Cahiers du GERAD G–2020–81 GERAD, HEC Montreal, Canada, GERAD, HEC Montreal, 2020. [Online]. Available: https://www.gerad.ca/en/papers/G-2020-81/view

[26] U. Sadana, P. V. Reddy, and G. Zaccour, Open-loop and Feedback Nash equilibria in scalar linear-state differential games with impulse control, Les Cahiers du GERAD G–2020–19 GERAD, HEC Montreal, Canada, GERAD, HEC Montreal, 2020. [Online]. Available: https://www.gerad.ca/en/papers/G-2020-19/view

[27] D. Chang, H. Wang, and Z. Wu, Maximum principle for non-zero sum differential games of BSDEs involving impulse controls, in Proceedings of the 32nd Chinese Control Conference, 2013, pp. 1564–1569.

[28] G. Ferrari and T. Koch, On a strategic model of pollution control, Annals of Operations Research, 275(2):297–319, 2019.

[29] R. Aïd, L. Campi, L. Li, and M. Ludkovski, An impulse-regime switching game model of vertical competition, Dynamic Games and Applications, 2021, ePub ahead of print, March 22.

[30] L. Campi and D. De Santis, Nonzero-sum stochastic differential games between an impulse controller and a stopper, Journal of Optimization Theory and Applications, 186(2):688–724, 2020.

[31] A. Cosso, Stochastic Differential Games Involving Impulse Controls and Double-Obstacle Quasi-variational Inequalities, SIAM Journal on Control and Optimization, 51(3):2102–2131, 2013.

[32] J. Elliott and N. J. Kalton, The existence of value in differential games, Memoirs of the American Mathematical Society, 1972.

[33] O. Bokanowski, S. Maroso, and H. Zidani, Some convergence results for howard's algorithm, SIAM Journal on Numerical Analysis, 47(4):3001–3026, 2009.