Feedback Nash Equilibria in Differential Games With Impulse Control

Utsav Sadana, Puduru Viswanadha Reddy, Member, IEEE, and Georges Zaccour

Abstract—In this article, 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, whereas 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 equations for Player 1 and the quasi-variational inequalities for Player 2. Furthermore, 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 toward 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.

Index Terms—Feedback Nash equilibrium (FNE), impulse controls, linear–quadratic differential games (LQDGs), nonzero-sum differential games, quasivariational inequalities (QVIs).

I. INTRODUCTION

ANY 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. The fixed costs associated with the impulses make it infeasible for the regulation agency to intervene at each time instant during the game. The firm, on the other hand, uses continuous controls as it incurs only marginal costs for its production. In [1], Ferrari and Koch consider a nonzero-sum impulse game between a regulation body and a firm where the firm’s production level is proportional to the emission of pollutants. The level of production is modeled as an Itô-diffusion process and both the firm and government affect the level of production (state variable) at discrete instants of time only. In particular, the drift and diffusion coefficients in their model that continuously affect the level of production are not controlled by either of the two players. This is in contrast to the zero-sum differential game models studied in deterministic [2], [3], [4] and stochastic settings [5] where one player continuously controls the state while the other player intervenes only occasionally in the game. In this article, we consider a deterministic nonzero-sum two-player differential game model with one player using continuous control and the other player using impulse control. This also differentiates our work from the literature on impulse games where all players in the game are assumed to make discrete-time interventions and do not have any continuous controls, see, e.g., [1], [6], [7], and [8]. Nonzero-sum differential games with impulse control have useful applications in counterterrorism, where a terrorist organization continuously builds up its resources to launch attacks while the government disrupts the resources at certain discrete time instants (see [9], [10], [11]).

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 [12]. In our previous paper [9], 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 [13]. 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 [14], [15], [16]. The resulting feedback strategies of the players are known to be strongly time consistent [13].

The objective of this article is to study the class of games that we have considered in [9] but here under a feedback information structure. In [9], 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 article 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. In this article, an FNE is obtained under the assumption that the impulse controls lie within the class of threshold policies (see e.g., [6], [8], [17], [18], [19]), 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) [20], [21], [22]. 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 [23] and [17]). 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 an FNE strategy. In particular, we show that the (value) functions that satisfy the Hamilton–Jacobi–Bellman (HJB) equations for Player 1, coupled with a system of QVIs for Player 2, coincide with the respective payoffs of the players in an 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 [6] and [8]), 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 an FNE strategy obtained by using the verification theorem is a function of the current time and state pairs, it is 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 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., [12], [14], [15], [16], and [24]. 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 an FNE in nonzero-sum LQDGs with impulse controls.

The specialized linear–quadratic game we study in this article 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 multiagent adaptation of the impulse optimal control problem (single player) studied in [17]. In particular, in our setting, Player 2’s impulse optimal control problem is a modified version of the impulse control problem analyzed in [17]. Our regularity assumptions on the value function and impulse controls of Player 2 also follow from the work in [17] where analytical solutions of the HJB equation are obtained in the continuation region by using a quadratic form on the value function; see also [25].

The rest of the article is organized as follows. In Section I-A, 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 II. In Section III, we provide a verification theorem for the existence of an FNE. In Section IV, we specialize our results to a scalar linear–quadratic game, and we solve this game in Section V for different problem parameters. Finally, Section VI concludes this article.

A. 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., [17] and [25]). For an elaborate description of (single-agent) impulse control problems, refer to the work in [26]. Recent theoretical and algorithmic advancements for solving QVIs associated with impulse control problems can be found in [27], [28], and [29]. 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 [30], [31]; management [32], [33], [34], [35], [36]; and epidemiology [37]. In contrast, the literature on differential games with impulse controls has been very limited and has predominantly dealt with zero-sum games (see, e.g., [5] and [38]). With the exception of our previous papers [9], [39], 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 [40].

In [9], we provided an algorithm for computing the OLNE in LQDGs with impulse control. In [39], we characterized the sampled-data Nash equilibrium for the class of games introduced.
In [9], in this article, we consider a general class of differential games with feedback information structure, and by the same token, push further the literature on nonzero-sum differential games.

Our work is closely related to the impulse games studied in [1, 6, 7, 8, and 19] with a feedback information structure where, however, all players are assumed to make discrete-time interventions in the continuous-time stochastic processes. To illustrate, Ferrari and Koch [1] 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 [19], a game problem between an impulse player and a stopper is solved using the QVIs. In [6], Aïd et al. studied infinite-horizon nonzero-sum game problem assuming threshold-type impulse controls and showed that a system of QVIs gives sufficient conditions for an 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. A policy iteration-type algorithm is given in [30] to solve the system of QVIs associated with 1-D nonzero impulse games introduced in [6] (see also [18]). However, the algorithm is derived under the assumption that both players have symmetric payoff functions and the convergence of the algorithm follows from the techniques developed in [28] and [29] for solving (one-player) stochastic impulse control problems. Basi et al. [8] extended the two-player model studied in [6] to an $N$-player setting and analyzed the corresponding mean-field game. In order to determine the optimal harvesting plan for natural resource management, Christensen et al. [42] solved a mean-field game and mean-field type impulse control problem, respectively, depending on whether the agents compete or cooperate with each other. 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, Aid et al. [6] and Basi et al. [8] determined semianalytic solutions for linear-state impulse stochastic games.

II. Model

We consider a deterministic two-player nonzero-sum differential game of finite duration $T$ 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 $\tau_i \in [0,T]$ denotes an intervention instant and $\xi_i \in \Omega_2 \subset \mathbb{R}^{m_2}$ denotes the parameter controlling the magnitude of impulse at time $\tau_i$. The impulse times $\tau_i$, $i \geq 1$ form an increasing sequence (see Definition 1) and the sets $\Omega_1$ and $\Omega_2$ are assumed to be bounded and convex.

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

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

And at the impulse instant $\tau_i$, Player 2 introduces jumps that are given by

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

where $f : \mathbb{R}^n \times \mathbb{R}^{m_1} \rightarrow \mathbb{R}^n$, $g : \mathbb{R}^n \times \mathbb{R}^{m_2} \rightarrow \mathbb{R}^n$, and $\tau_i^-$ and $\tau_i^+$ denote the time instants just before and after the intervention time $\tau_i$, $x(\tau_i^-) = \lim_{t \downarrow \tau_i} x(t)$, $x(\tau_i^+) = \lim_{t \uparrow \tau_i} x(t)$, and $0^-$ denotes the time instant just before $0$. The state variable is assumed to be left-continuous at points of discontinuity.

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

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

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

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

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

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

i) $f(x, u)$ is (uniformly) Lipschitz continuous in $x$, that is, there exists a constant $c_f > 0$, such that

$$|f(x, u) - f(y, u)| \leq c_f |x - y|, \quad \forall x, y \in \mathbb{R}^n, u \in \Omega_1.$$  (5)

ii) $g(x, \xi)$ is (uniformly) Lipschitz continuous in $x$, such that, for $c_g > 0$, we have

$$|g(x, \xi) - g(y, \xi)| \leq c_g |x - y|, \quad \forall x, y \in \mathbb{R}^n, \xi \in \Omega_2.$$  (6)

iii) $\forall x \in \mathbb{R}^n, \inf_{\xi \in \Omega_2} b_2(x, \xi) = \mu > 0.$

iv) The functions $f$, $g$, $h_i$, $b_i$, and $s_i$ are bounded for $i \in \{1, 2\}$.

Assumptions 1.i) and 1.ii) ensure that there exists a unique state trajectory $x(\cdot)$ for any measurable $u(\cdot)$ and impulse sequence $\{(\tau_i, \xi_i)\}_{i \geq 1}$. Assumption 1.iii) ensures that Player 2 intervenes only a finite number of times in the game due to the fixed cost associated with each impulse (see [17], where similar assumptions are provided in the context of an impulse optimal control problem). Assumption 1.iv) 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.
TABLE I
NOMENCLATURE

| Type            | Symbol | Definition                                                                 | Equations |
|-----------------|--------|-----------------------------------------------------------------------------|-----------|
| Controls        | $u(t)$ | Piecewise continuous control of Player 1                                    | (3)       |
|                 | $\dot{e}$ | Impulse control of Player 2                                                 | (4)       |
|                 | $\Omega_1$ | Set of admissible controls of Player 1                                     | (5)       |
|                 | $\tau_i$ | ith impulse instant                                                         | (2)       |
|                 | $\xi_i$ | Parameter controlling impulse level of $i$th impulse                       |           |
| Strategies      | $\gamma$ | State-feedback strategy of Player 1                                         | (8)       |
|                 | $\delta$ | State-feedback strategy of Player 2                                         | (9)       |
|                 | $\Delta$ | Set of admissible state-feedback strategies of Player 1                    | (11a)     |
|                 | $\zeta()$ | Measurable function that determines the size of impulse at current time and current state | (12)     |
|                 | $C$    | Collection of time and state pairs for which it is optimal not to give an impulse | (15)    |
|                 | $\mathcal{I}$ | Collection of time and state pairs for which it is optimal to give an impulse | (16)    |
| Values          | $V_j$ | Value function of Player $j$ in continuation region in $\text{lQDG}$         | (11a), (11b) |
|                 | $\Phi_j$ | Cost-to-go of Player $j$ in continuation region in $\text{lQDG}$           | (28), (30) |
| Threshold policy| $\ell_1()$ | Left boundary of the continuation interval $C$                            | (50a)     |
|                 | $\ell_2()$ | Right boundary of the continuation interval $C$                            | (50b)     |
|                 | $\alpha()$ | Equilibrium state after a positive impulse                                | (47a)     |
|                 | $\beta()$ | Equilibrium state after a negative impulse                                 | (47b)     |
|                 | $\mu$ | Lower bound on the impulse cost $b_2()$ of Player 2                        | (22)      |
|                 | $c_f$ | Lipschitz constant for the state derivative $f()$ with respect to time    | (48a)     |
|                 | $c_g$ | Lipschitz constant of $g()$                                                 | (48b)     |
| Miscellaneous   | $\mathcal{H}_j$ | Hamiltonian operator of Player $j$                                         | (13d), (14e) |
|                 | $K$ | Upper bound on the equilibrium number of impulses                           | (22)      |
|                 | $\theta_a(), \theta_b(\cdot)$ | Functions of time introduced to simplify notations                   | (51c), (51d) |
|                 | $x_{11}(\cdot), x_{22}(\cdot)$ | Functions of time introduced to simplify notations                  | (51a), (51b) |
|                 | $r_1, r_2, r_3, r_4$ | Parameters                                                              | (52b), (52c), (51a), (61a), (61b) |
|                 | $\theta$ | Parameter                                                                  | (32a)     |

We summarize the notations used in the article in Table I.

III. 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. In this information structure, players use strategies that are functions of the current time $t$ and current state vector $x(t)$.

A. Strategy of Player 1 and Player 2

The strategy spaces of the players are described as follows: Let $\Sigma := \{(t, x(t)) \in [0, T], x \in \mathbb{R}^n\}$ 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 $\gamma : [0, T] \times \mathbb{R}^n \rightarrow \Omega_1$, while Player 2 exercises discrete-time actions given by her state-feedback intervention policy $\delta$. Following the literature (see [17] and [6]) on impulse controls, the intervention policy $\delta$ involves determining a continuation set $C$ and a continuous function $\zeta()$ such that Player 2 gives an impulse if and only if $(t, x(t)) \notin C$, and when Player 2 gives an impulse, its magnitude depends on the function $\zeta() : \Sigma \rightarrow \Omega_2$. The intervention set $\mathcal{I}$ is given by $\mathcal{I} = \Sigma \setminus C$. For a given strategy pair $(\gamma, \delta)$, Player 1’s control is given by $u(t) = \gamma(t, x(t))$ and Player 2’s impulse control $\dot{e}$ is a sequence $\{(\tau_i, \xi_i)\}_{i \geq 1}$ where $\tau_i$ is the impulse instant and $\xi_i$ is the impulse level parameter such that $\xi_i = \zeta(t, x(t))$.

**Definition 1:** The sequence $\dot{e} = \{(\tau_i, v_i)\}_{i \geq 1}$, 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

$$\mathcal{T} = \{\tau_i, i \geq 1 \mid 0 \leq \tau_1 < \tau_2 < \cdots < T\}$$

$$\tau_n = \inf \{t > \tau_{n-1} : (t, x(t)) \not\in C\}, n \geq 1, \tau_0 := 0$$

where $\tau_n$ depends on the current state value for all $n \in \mathbb{N}$.

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

**Remark 1:** We emphasize that the timing of the interventions is given in feedback form as the continuation set $C$ depends on both the current time and the current state vector. In particular, the continuation and intervention sets will be characterized, in Section III-B2, 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., [5] and [43]) assuming nonanticipative strategies [44] 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 [6]. As mentioned in [6], 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 $(\gamma, \delta)$ can be described as follows: Player
1 continuously controls the state trajectory using state feedback $\gamma(t, x(t))$ during the time that the state lies in the continuation set $C$. When the state leaves set $C$, Player 2 intervenes and gives an impulse with size dependent on $\zeta(t, x(t))$ to bring the state into set $C$.

Next, we determine the cost-to-go functions for Player 1 and Player 2 for a given strategy pair $(\gamma, \delta)$ and for any starting position of the game $(t, x(t))$. Suppose $\gamma([t, T]) \in \Gamma([t, T])$ and $\delta([t, T]) \in \Delta([t, T])$ are restrictions of $\gamma$ and $\delta$, 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 $(t, x(t))$ is given by

$$
\dot{x}(t) = f(x(t), \gamma(t, x(t))), \quad x(t) = x, \text{ for } (t, x(t)) \in C \quad (6)
$$

and the cost-to-go functions are given by

$$
J_1(t, x(t), \gamma([t, T]), \delta([t, T])) = \int_t^T h_1(s, \gamma(t, x(t)))ds + \sum_{i \geq 1} \mathbb{I}_{[t, \tau_i]} b_1(x(\tau_i^+), \xi_i) + s_1(x(T)) \quad (8)
$$

$$
J_2(t, x(t), \gamma([t, T]), \delta([t, T])) = \int_t^T h_2(s, \gamma(t, x(s)))ds + \sum_{i \geq 1} \mathbb{I}_{[t, \tau_i]} b_2(x(\tau_i^-), \xi_i) + s_2(x(T)). \quad (9)
$$

The subgame starting at $(t, x(t))$ that is described by (6)–(9) 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 FNE is defined as follows.

**Definition 2:** For the differential game described by (6)–(9) with a memoryless perfect state information pattern, the strategy profile $(\gamma^*, \delta^*)$ in $\Gamma \times \Delta$ constitutes an FNE solution if, for any $(t, x(t)) \in \Sigma$; we have

$$
J_1(t, x(t), \gamma^*[t, T], \delta^*[t, T]) \leq J_1(t, x(t), \gamma([t, T]), \delta([t, T])) \quad \forall \gamma([t, T]) \in \Gamma([t, T]),
$$

$$
J_2(t, x(t), \gamma^*[t, T], \delta^*[t, T]) \leq J_2(t, x(t), \gamma([t, T]), \delta([t, T])) \quad \forall \delta([t, T]) \in \Delta([t, T]).
$$

**B. Verification Theorem**

In this section, we provide methods for identifying an FNE associated with the differential game described by (6)–(9). To this end, from (10a), we know that an FNE strategy $\gamma^*$ of Player 1 provides the best response to Player 2’s FNE strategy $\delta^*$. Similarly, from (10b), Player 2’s FNE strategy $\delta^*$ is the best response to Player 1’s FNE strategy $\gamma^*$. Using (10a) and (10b), the equilibrium cost-to-go of the players, also referred to as value functions, denoted by $V_j : [t, T] \times \mathbb{R}^n, j = 1, 2$ are defined as follows:

$$
V_1(t, x(t)) = \inf_{\gamma([t, T]) \in \Gamma([t, T])} J_1 \left( t, x(t), \gamma([t, T]), \delta([t, T]) \right) \quad (11a)
$$

$$
V_2(t, x(t)) = \inf_{\delta([t, T]) \in \Delta([t, T])} J_2 \left( t, x(t), \gamma([t, T]), \delta([t, T]) \right). \quad (11b)
$$

The following are the standing assumptions on the value functions, which will be used throughout the article.

**Assumption 2:** The value function of Player 1, $V_1(t, x(t))$, is differentiable in both $t$ and $x(t)$ when $(t, x(t)) \in C$.

**Assumption 3:** The value function of Player 2, $V_2(t, x(t))$, is differentiable in both $t$ and $x(t)$ for almost all values in $\Sigma$.

**Remark 4:** $V_2(t, x(t))$ can have kinks at those state values that lie at the boundary of the continuation set $C$. In general, in impulse control problems and impulse games using verification theorems, the value function is required to be differentiable for all state values in order to apply Ito’s lemma (see [6], and the references therein). As a result, certain smooth fit conditions are imposed at the boundaries of the continuation set $C$.

**Assumption 4:** There exists a unique measurable function $\zeta : \Sigma \rightarrow \Omega_2$ such that

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

**1) Optimal Control Problem of Player 1:** From (10a), the value function $V_1(t, x(t))$ associated with Player 1’s optimal control problem satisfies the following HJB equation for a given impulse control $\{(\tau_i^+, \xi_i^+)\}_{i \geq 1}$ corresponding to Player 2’s FNE strategy $\delta^*$:

$$
-\nabla_x V_1(t, x(t)) = \min_{\varphi \in \Omega_1(t, x(t))} \mathcal{H}_1(t, x(t), \varphi, \nabla_x V_1(t, x(t))),
$$

$$
(t, x(t)) \in C \quad (13a)
$$

$$
V_1(T, x(T)) = s_1(x(T)) \quad \forall (T, x(T)) \in \Sigma \quad (13b)
$$

$$
V_1(\tau_i^-) = V_1(\tau_i^- + g(x(\tau_i^-), \xi_i^+)) + b_1(x(\tau_i^-), \xi_i^+) \quad (13c)
$$

where

$$
\mathcal{H}_1(t, x(t), \varphi, \nabla_x V_1(t, x(t))) = h_1(x(t), \varphi) \quad \nabla_x V_1(t, x(t))^T f(x(t), \varphi). \quad (13d)
$$

The above conditions can be interpreted as follows. From Definition 1, an admissible impulse cannot occur at the terminal time, hence condition (13b) holds. In the continuation region $C$, Player 2 does not give any impulse, and therefore, the value function of Player 1 satisfies the HJB equation (13a). 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,

---

1This assumption is also made in [17] to define (weak) QVIs (see [14a] and references [20], [21]).

2For an LQDG, we show in Section V (see Figs. 2 and 4) that the value function of Player 2 given by $V_2(0, x_0)$ at initial time $0$ and initial state $x_0$ is nondifferentiable only for the initial state values $\ell_1(0)$ and $\ell_2(0)$, which are the left and right boundaries of the continuation set $C$. 

---

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
2) Impulse Control Problem of Player 2: Player 2 solves the impulse optimal control problem (10b) for a given equilibrium trajectory $\gamma^*(t, x(t))$ of Player 1.

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

$$
\forall t \in [0, T], \text{ a.a. } x \in \mathbb{R}^n, \text{ we have}
$$

$$
\nabla_t V_2(t, x(t)) + H_2(x(t), \gamma^*(t, x(t))), \nabla_x V_2(t, x(t)) \geq 0
$$

(14a)

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

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

(14b)

$$
(V_2(t, x(t)) - \mathcal{R}V_2(t, x(t))) \left(\nabla_t V_2(t, x(t)) + H_2(x(t), \gamma^*(t, x(t))), \nabla_x V_2(t, x(t))\right) = 0
$$

(14c)

and $V_2(T, x(T)) = s_2(x(T)) \forall (T, x(T)) \in \Sigma$

(14d)

where the Hamiltonian operator $H_2$ and intervention operator $\mathcal{R}$ are defined as follows:

$$
H_2(x(t), \gamma^*(t, x(t))), \nabla_x V_2(t, x(t)) = h_2(x(t), \gamma^*(t, x(t)))
$$

$$
+ (\nabla_x V_2(t, x(t)))^T f(x(t), \gamma^*(t, x(t)))
$$

(14e)

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

(14f)

Recall that in Assumption 4, (12) gives the unique optimal impulse level at any $(t, x(t))$ since it minimizes the sum of the immediate cost $b_2(x(t), \eta)$ incurred from giving an impulse of size $\eta$ and the cost-to-go from playing optimally afterward (see also [6], where a similar assumption is used to solve stochastic impulse games).

Remark 5: QVIs can be interpreted as follows.

1) Condition (14b) ensures that the value function $V_2(\cdot)$ evaluated at any $(t, x(t)) \in \Sigma$ is always above the obstacle $\mathcal{R}V_2(t, x)$, that is, $V_2$ is at most equal to the minimum cost that Player 2 incurs from intervening at time $t$ and playing optimally afterward.

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 $(t, x(t)) \in \Sigma$. Hence, when $V_2(t, x(t)) = \mathcal{R}V_2(t, x(t))$, Player 2 gives an impulse.

3) At any $(t, x(t)) \in \Sigma$, condition (14e) ensures that either Player 2 waits so that the HJB-like equation (14a) for Player 2 holds with equality or Player 2 gives an impulse.

Remark 6: The condition $V_2(\tau, x(\tau)) = \mathcal{R}V_2(\tau, x(\tau))$ results in the continuity of the value function of Player 2 at the impulse instant $\tau$ under the feedback information structure. For impulse control problems studied by using the Pontryagin maximum principle, the Hamiltonian continuity condition [45] gives the timing of interventions (see also [9], 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

$$
\mathcal{C} = \{(t, x(t)) \in [0, T] | V_2(t, x(t)) < R_2(t, x(t)) \}
$$

(15)

$$
\mathcal{I} = \{(t, x(t)) \in \Sigma | V_2(t, x(t)) = R_2(t, x(t)) \}
$$

(16)

Remark 7: In impulse games studied in [6] and [8], 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. Furthermore, in the infinite-horizon impulse game studied in [6], the continuation set depends only on the current state while the continuation set in our finite-horizon game problem depends on both the current time and the current state.

The sufficient conditions to characterize an FNE of the differential game described in (6–9) are given in the next theorem.

Theorem 1 (Verification Theorem): Let Assumptions 1–4 hold. Suppose there exist functions $H_j : \Sigma \rightarrow \mathbb{R} (j = 1, 2)$ such that $H_j(t, x(t))$ satisfies the HJB equations (13) and $V_j(t, x(t))$ satisfies the QVIs (14) for all $(t, x(t)) \in \Sigma$. Suppose there exist strategies $(\gamma^*, \delta^*)$ with the following properties. Player 1’s control $u^*(t) = \gamma^*(t, x(t))$ satisfies for all $t \in [0, T]$

$$
\gamma^*(t, x(t)) = \arg \min_{\varphi \in \Omega_1} H_1(x(t), \varphi, \nabla_x V_1(t, x(t))
$$

(17a)

and Player 2’s impulse control is a sequence $\{(\tau_i^*, \xi_i^*)\}_{i \geq 1}$ where interventions occur at $\tau_i^* = t$ if $(t, x(t)) \in \mathcal{I}$, that is, $(t, x(t))$ satisfy

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

(17b)

and the corresponding impulse level parameters $\xi_i^*$ are given by

$$
\xi_i^* = \zeta(t, x(t))
$$

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

(17c)

Then, $(\gamma^*, \delta^*)$ is an FNE of the differential game described by (6–9). Furthermore, $V_j(t, x(t))$ is the equilibrium cost-to-go of Player $j$, $(j = 1, 2)$ for the subgame starting at $(t, x(t)) \in \Sigma$ and defined over the horizon $[t, T]$.
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 $\delta_{[t,T]}^{\ast}$, such that Player 1’s control $u(t)$ is given by $u(s) = \gamma_{[t,T]}(s, x_1(s)), s \in [t, T]$. Using the total derivative of $V_1(\cdot)$ between the impulse instants $(\tau_{i-1}^\ast, \tau_i^\ast)$, integrating with respect to $t$ from $\tau_{i-1}^\ast$ to $\tau_i^\ast$, and taking the summation for all $i \geq 1$, we obtain

$$V_1(T, x_1(T)) - V_1(t, x(t)) = \sum_{i \geq 1} \int_{\tau_{i-1}^\ast}^{\tau_i^\ast} \left( \nabla_x V_1(s, x_1(s)) \right) ds + \sum_{i \geq 1} \int_{\tau_{i-1}^\ast}^{\tau_i^\ast} f(x(s), u(s)) ds$$

where we defined $\tau_0^\ast := t$. From (13a), we know that, for any given control $u(s)$, the following inequality holds:

$$\nabla_x V_1(s, x_1(s)) + \left( \nabla_x V_1(s, x_1(s)) \right)^\top f(x(s), u(s)) \geq - h_1(x_1(s), u(s)).$$

Therefore, we obtain

$$V_1(T, x_1(T)) - V_1(t, x(t)) \geq - \sum_{i \geq 1} \int_{\tau_{i-1}^\ast}^{\tau_i^\ast} h_1(x_1(s), u(s)) ds + \sum_{i \geq 1} \int_{\tau_{i-1}^\ast}^{\tau_i^\ast} f(x(s), u(s)) ds$$

From the terminal condition (13b) on $V_1(\cdot)$ and (13c), we obtain

$$V_1(t, x(t)) \leq s_1(x(T)) + \sum_{i \geq 1} \int_{\tau_{i-1}^\ast}^{\tau_i^\ast} h_1(x_1(s), u(s)) ds + \sum_{i \geq 1} \int_{\tau_{i-1}^\ast}^{\tau_i^\ast} f(x(s), u(s)) ds$$

where $x^*(\cdot)$ is the state trajectory generated by Player 1 choosing the strategy $\gamma_{[t,T]}^\ast$ and Player 2 choosing the strategy $\delta_{[t,T]}^\ast$. Therefore, $\gamma_{[t,T]}^\ast$ is the best response to Player 2’s strategy $\delta_{[t,T]}^\ast$.

Next, we consider an arbitrary admissible strategy $\delta_{[t,T]}$ of Player 2 such that the intervention instants are given by $\tau_i, i \geq 1$ and the corresponding impulse levels are given by $\xi_i$. Furthermore, $x_2(\cdot)$ is the state trajectory generated by the strategy pairs $(\gamma_{[t,T]}^\ast, \delta_{[t,T]}^\ast)$. We obtain the following relation by taking the total derivative of $V_2(\cdot)$ between the impulse instants $(\tau_{i-1}, \tau_i)$, integrating over time from $\tau_{i-1}$ to $\tau_i$, and taking the summation for all $i \geq 1$:

$$V_2(T, x_2(T)) - V_2(t, x(t)) = \sum_{i \geq 1} \int_{\tau_{i-1}}^{\tau_i} \left( \nabla_s V_2(s, x_2(s)) \right) ds + \sum_{i \geq 1} \int_{\tau_{i-1}}^{\tau_i} f(x_2(s), \gamma^\ast(s, x_2(s))) ds$$

Therefore, we obtain

$$\nabla V_2(s, x_2(s)) + \left( \nabla V_2(s, x_2(s)) \right)^\top f(x_2(s), \gamma^\ast(s, x_2(s))) \geq - h_2(x_2(s), \gamma^\ast(s, x_2(s))).$$

Given an impulse level parameter $\xi_i$ and the definition of an intervention operator given in (14d), we obtain

$$\mathcal{R} V_2(\tau_i^\ast, x_2(\tau_i^\ast)) \leq V_2(\tau_i^\ast, x_2(\tau_i^\ast)) + b_2(x_2(\tau_i^\ast), \xi_i).$$

Also, from (14b), we know that

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

Therefore, we obtain

$$V_2(\tau_i^\ast, x_2(\tau_i^\ast)) - V_2(\tau_i^\ast, x_2(\tau_i^\ast)) \geq - b_2(x_2(\tau_i^\ast), \xi_i).$$

Substitute (20) and (21) into (19) to obtain

$$V_2(T, x_2(T)) - V_2(t, x(t)) \geq \sum_{i \geq 1} \int_{\tau_{i-1}}^{\tau_i} \left( \nabla_s V_2(s, x_2(s)) \right) ds + \sum_{i \geq 1} \int_{\tau_{i-1}}^{\tau_i} f(x_2(s), \gamma^\ast(s, x_2(s))) ds$$

Substituting the terminal condition $V_2(T, x_2(T)) = s_2(x_2(T))$, given in (14d), in the above inequality yields

$$V_2(T, x_2(T)) \leq s_2(x_2(T)) + \sum_{i \geq 1} \int_{\tau_{i-1}}^{\tau_i} h_2(x_2(s), \gamma^\ast(s, x_2(s))) ds + \sum_{i \geq 1} \int_{\tau_{i-1}}^{\tau_i} f(x_2(s), \gamma^\ast(s, x_2(s))) ds$$

where $x^\ast(\cdot)$ is the state trajectory generated by Player 1 choosing the strategy $\gamma_{[t,T]}^\ast$ and Player 2 choosing the strategy $\delta_{[t,T]}^\ast$. Therefore, $\gamma_{[t,T]}^\ast$ is the best response to Player 2’s strategy $\delta_{[t,T]}^\ast$.
(17c). Therefore, for a strategy $$\delta^*$$, we obtain
$$V_2(\tau^+, x^*(\tau^+)) - V_2(\tau^-, x^*(\tau^-)) = -b_2(x^*(\tau^-), \xi_i^*)$$
$$\nabla_t V_2(s, x^*(s)) + (\nabla_x V_2(s, x^*(s)))^T f(x^*(s), \gamma^*(s, x^*(s))) = -h_2(x^*(s), \gamma^*(s, x^*(s)))$$
and the cost-to-go function is given by
$$V_2(t, x(t) = s_2(x^*(T)) + \sum_{i \geq 1} \int_{\tau_i^-}^{\tau_i^+} h_2(x^*(s), \gamma^*(s, x^*(s))) ds$$
$$+ \sum_{i \geq 1} b_2(x^*(\tau_i^-), \xi_i^*)$$
$$= J_2(t, x(t), \gamma^*[t,T], \delta^*[t,T]).$$

Therefore, $$\delta^*$$ is the best response strategy to Player 1’s strategy $$\gamma^*$$.

**Remark 8:** An important feature of an FNE solution introduced in Definition 2 is that if the strategy pair $$(\gamma^*, \delta^*)$$ provides an FNE to the differential game described by (6–9) with duration $[0, T]$, then its restriction to the time interval $[t, T]$, denoted by $$(\gamma^*[t,T], \delta^*[t,T])$$, provides an 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 $$(\gamma^*, \delta^*)$$ is strongly time consistent.

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

**Theorem 2:** Let Assumption 1 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
$$K = \left\lceil \frac{2(T\|h_2\|_\infty + \|s_2\|_\infty)}{\mu} \right\rceil$$
(22)
where $\mu = \inf_{x \in \mathbb{R}^n} b_2(x, \eta) > 0 \forall x \in \mathbb{R}^n$, and $[y]$ denotes the smallest integer that is greater than or equal to $y$.

**Proof:** See Appendix A. \(\square\)

The difficulty in solving the QVIs arises from the fact that the obstacle $RV_2$ depends on the value function $V_2$ [20] that we seek to characterize for Player 2. As a result, QVIs have been solved in the literature under some restrictive assumptions on the value functions, even for infinite-horizon games with linear objective functions, see e.g., [6] and [19]. An additional difficulty in our case is that the QVIs are coupled with HJB equations associated with Player 1’s best response and the value functions of players depend on both the current time and state pairs since the game is of finite duration. Even for a single-agent finite-horizon linear–quadratic impulse problem analyzed in [17], the value function is assumed to be quadratic in the state in the continuation region. In the next section, we specialize our results to LQDGs making the standard assumption of quadratic-in-the-state value functions for the players [14], [16], [17] and provide a complete analytical characterization of an FNE strategy.

**IV. SCALAR LQDG WITH TARGETS**

In this section, we illustrate the theory developed in the previous two sections by considering a scalar linear–quadratic adaptation of the differential game (1–4), referred to as iLQDG hereafter. Player 1’s problem structure follows from the classical linear–quadratic regulator problem. To motivate the problem of Player 2, one can consider a government that subsidizes the use of environment friendly technology if the pollution levels exceed a given level and revokes the subsidy once the pollution reaches a minimum threshold. Suppose Player 1 and Player 2 aim to minimize the costs resulting from the deviation of the state away from their target state values $\rho_1$ and $\rho_2$, 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 [17].

**iLQDG:**
$$J_1(0, x_0, u(\cdot), \bar{v}) = \int_0^T \frac{1}{2}(w_1(x(t) - \rho_1)^2 + r_1 u(t)^2) \, dt$$
$$+ \sum_{i=1}^k z_i |\xi_i| + \frac{1}{2} s_1(x(T) - \rho_1)^2 \quad (23a)$$
$$J_2(0, x_0, u(\cdot), \bar{v}) = \int_0^T \frac{1}{2}w_2(x(t) - \rho_2)^2 \, dt$$
$$+ \sum_{i=1}^k h(\xi_i) + \frac{1}{2} s_2(x(T) - \rho_2)^2 \quad (23b)$$
$$\dot{x}(t) = ax(t) + bu(t), \quad x(0^-) = x_0 \forall t \neq \tau_i, \ i \geq 1$$
(23c)
$$x(\tau_i^+) = x(\tau_i^-) + \xi_i$$
(23d)

where
$$h(\xi_i) := \begin{cases} 
C + c \xi_i & \text{if } \xi_i > 0 \\
\min(C, D) & \text{if } \xi_i = 0 \\
D - d \xi_i & \text{if } \xi_i < 0
\end{cases} \quad (24)$$
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 $\tau_i$ is 0. Player $j$ incurs an instantaneous cost $\frac{1}{2}w_j(x(t) - \rho_j)^2$ if the state deviates from $\rho_j$ while the terminal cost is $\frac{1}{2}s_j(x(T) - \rho_j)^2$ for Player $j$. Also, 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.

Given that Player 2 aims to keep the state value close to her target, we assume that Player 2’s control policy is of threshold-type that involves determining the boundaries of the interval $((\ell_1(\cdot), \ell_2(\cdot))$ where Player 2 does not give an impulse as well as characterizing the levels of impulses when state leaves the interval $((\ell_1(\cdot), \ell_2(\cdot))$. Suppose that Player 2 shifts the state to $\alpha(\cdot)$ or $\beta(\cdot)$ depending on whether the state reached the boundary $\ell_1(\cdot)$ or $\ell_2(\cdot)$, respectively. The threshold policy of Player 2 naturally leads to a form of continuation set of Player 2 that is described ahead (see also [17] and [25]).
Assumption 5: Player 2 gives an impulse if \( (t, x(t)) \) does not lie in the continuation set \( C \) given by

\[
C = \{(t, x(t)) \in \Sigma \mid \ell_1(t) < x(t) < \ell_2(t)\}. \tag{25}
\]

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:

\[
\ell_1(t) < \alpha(t) < \beta(t) < \ell_2(t). \tag{26}
\]

Player 2 endogenously determines the functions \( \ell_1(\cdot), \alpha(\cdot), \beta(\cdot) \), and \( \ell_2(\cdot) \) by solving an associated system of QVIs for her impulse optimal control problem.

Next, we state the assumptions on the controllers of Player 1 and Player 2 that we have used to derive closed-form expressions for FNE strategies.

Assumption 6: The admissible control \( u \) of Player 1 and impulse size \( \xi \) for Player 2 lie in the interior of the bounded and open convex sets \( \Omega_1 \) and \( \Omega_2 \), respectively.

Assumption 7: The value function of Player 1 is given by

\[
V_1(t, x(t)) = \begin{cases} 
\Phi_1(t, \alpha(t)) + z_1|\alpha(t) - x(t)|, & x(t) \leq \ell_1(t) \\
\Phi_1(t, x(t)), & x(t) \in (\ell_1(t), \ell_2(t)) \\
\Phi_1(t, \beta(t)) + z_1|\beta(t) - x(t)|, & x(t) \geq \ell_2(t).
\end{cases}
\]

(27)

Since the game is linear–quadratic, \( \Phi_1 \) is quadratic in the state

\[
\Phi_1(t, x(t)) = \frac{1}{2} p_1(t)x(t)^2 + q_1(t)x(t) + n_1(t)
\]

where \( p_1, q_1 \in C^1([0, T] \times \mathbb{R}) \).

Assumption 8: The value function of Player 2 is given by

\[
V_2(t, x(t)) = \begin{cases} 
\Phi_2(t, \alpha(t)) + C + c(\alpha(t) - x(t)), & x(t) \leq \ell_1(t) \\
\Phi_2(t, x(t)), & x(t) \in (\ell_1(t), \ell_2(t)) \\
\Phi_2(t, \beta(t)) + D + d(x(t) - \beta(t)), & x(t) \geq \ell_2(t)
\end{cases}
\]

(28)

where

\[
\Phi_2(t, x(t)) = \frac{1}{2} p_2(t)x(t)^2 + q_2(t)x(t) + n_2(t)
\]

and \( p_2, q_2, n_2 \in C^1([0, T] \times \mathbb{R}) \).

Assumption 9: For \( t \in [0, T] \), the problem parameters satisfy

\[
\nu_2 \theta + w_2(1 - e^{\theta \nu_1} - 2t\theta \nu_1) > 0
\]

(31)

where

\[
\theta = 2 \sqrt{a^2 + \frac{b^2}{r_1}}
\]

(32a)

\[
\nu_1 = \left( \frac{2 \theta}{\theta + 2 \frac{r_1}{\nu_1} s_1 - 2a} - 1 \right) e^{-\theta T}
\]

(32b)

\[
\nu_2 = 2 \nu_1 s_2 + s_2 e^{-\theta T} - \frac{w_2 e^{-\theta T} - \nu_1^2 w_2 e^{T \theta} + \nu_1^2 s_2 e^{T \theta} + 2 \nu_1 T w_2}{\theta}
\]

(32c)

A. Optimal Control Problem of Player 1

Let the equilibrium strategy of Player 2 be given by \( \delta^* \) such that Player 2 gives an impulse if the state leaves the continuation set \( C \) described in Assumption 5. Then, the equilibrium strategy of Player 1 can be determined by finding the value function that satisfies (13a)–(13c) for the iLQDG.

Player 1 solves a linear–quadratic optimal control problem in the continuation region \( C \), and at the impulse instant \( \tau_i \), Player 1’s cost is given by \( z_1 \xi_i \). Therefore, in Assumption 7, we can make a guess on the form of the value function of Player 1.

In the following theorem, we give an analytical characterization of the state-feedback strategy of Player 1 using the HJB equation in the continuation region \( C \):

Theorem 3: Let Assumptions 5–7 hold. Then, the equilibrium state-feedback strategy of Player 1 is given by

\[
\gamma^*(t, x(t)) = \frac{1}{b} \left( \frac{\theta}{2} - \frac{\theta}{\nu_1 e^{\theta t} + 1} - \alpha \right) x(t) - \frac{b}{r_1} q_1(t)
\]

(33)

\[\forall (t, x(t)) \in C\]

where

\[
q_1(t) = \frac{2 \rho_1 w_1}{\theta} - \rho_1 \left( s_1 \theta - 2 w_1 (1 - 2 e^{\frac{\theta (s_1 - 1)}{2}}) + \nu_1 e^{T \theta} (2 w_1 + s_1 \theta) \right) e^{\frac{\theta (s_1 - 1)}{2} \theta} (\nu_1 e^{\theta t} + 1).
\]

Proof: The equilibrium control of Player 1 is obtained by substituting the value function in the HJB equation. From (13a),

\[
- \nabla_x \Phi_1(t, x(t)) = \min_{u \in \Omega_1} \left( \frac{1}{2} w_1 (x(t) - \rho_1)^2 + \frac{1}{2} r_1 u(t)^2 + \nabla_x \Phi_1(t, x(t))(ax(t) + bu(t)) \right).
\]

(34)

Differentiating the right-hand side of the above equation and equating the result to zero yields the equilibrium strategy of Player 1 (see Assumption 6)

\[
\gamma^*(t, x(t)) = u^*(t) = - \frac{b}{r_1} \nabla_x \Phi_1(t, x(t))
\]

(35)

\[= - \frac{b}{r_1} (p_1(t)x(t) + q_1(t)).\]

Substituting (35) in the state dynamics (23c), we obtain

\[
\dot{x}(t) = ax(t) + bu^*(t) = ax(t) - \frac{b^2}{r_1} (p_1(t)x(t) + q_1(t))
\]

(36)

\[= ax(t)x(t) + bx q_1(t)\]

where

\[
a_x(t) = a - \frac{b^2}{r_1} p_1(t), \quad b_x = - \frac{b^2}{r_1}.
\]

On substituting (35) and (28) into (34), we obtain

\[
- \dot{p}_1(t)x(t)^2 - 2q_1(t)x(t) - 2a_1(t)
\]

\[= w_1 x(t)^2 + w_1 p_1^2 - 2x(t)w_1 p_1 - b_x (p_1(t)x(t)^2) - q_1(t)^2 \]

\[+ 2a_x(t) (p_1(t)x(t) + q_1(t)) x(t).\]
Since the above equation must hold for all \( x \) except at \( (t, x(t)) \not\in C, p_1(\cdot), q_1(\cdot), \) and \( n_1(\cdot) \) evolve as follows:

\[
\begin{align*}
p_1(t) &= -w_1 - b_1 p_1(t)^2 - 2p_1(t) a \\
q_1(t) &= -a_2(t)q_1(t) + w_1 p_1 \\
n_1(t) &= -\frac{1}{b_2} q_1(t)^2 - \frac{w_1 p_1^2}{2}
\end{align*}
\]  

(37a, 37b, 37c)

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

The solution of (37a) is given by the following equation (see Appendix C):

\[
p_1(t) = \frac{1}{b_2} \left( -a - \frac{\theta}{2 + \frac{\theta}{\nu_1 e^{\nu_1 t} + 1}} \right) \ .
\]  

(38)

Using the value of \( p_1(t) \) given in (35), we obtain

\[
\gamma^\ast(t, x(t)) = \frac{1}{b} \left( -a + \frac{\theta}{2 + \frac{\theta}{\nu_1 e^{\nu_1 t} + 1}} \right) x(t) - \frac{b}{\tau_1} q_1(t) 
\]  

\[
\square
\]

It is to be noted that there can be 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 \( C, \) Player 1’s strategy is continuous in \( t \) and \( x(t). \)

Next, we determine \( n_1 \) for any \( (t, x(t)) \), and thereby characterize the value function of Player 1 for given impulse instants \( \tau_i^\ast \) and impulse levels \( \alpha(\tau_i^\ast) \) or \( \beta(\tau_i^\ast) \) depending on the state values at the impulse instants.

When an impulse occurs, that is, \( (\tau_i^\ast, x(\tau_i^\ast)) \in \mathbb{I}, \) it follows from (13c) that \( V_1 \) satisfies

\[
\begin{align*}
\frac{1}{2} p_1(\tau_i^-) x(\tau_i^-)^2 + q_1(\tau_i^-) x(\tau_i^-) + n_1(\tau_i^-) \\
= \frac{1}{2} p_1(\tau_i^\ast)^2 (x(\tau_i^\ast) - \xi_i^\ast)^2 + q_1(\tau_i^\ast) (x(\tau_i^\ast) + \xi_i^\ast) \\
+ n_1(\tau_i^\ast) + z_1(\xi_i^\ast).
\end{align*}
\]

The equilibrium strategy of Player 2 is to bring the state to \( \alpha(t) \) if \( x(t) \leq \ell_1(t) \), and to \( \beta(t) \) if \( x(t) \geq \ell_2(t) \), that is, \( x(\tau_i^\ast) + \xi_i^\ast = \alpha(\tau_i^\ast) \) if \( x(\tau_i^-) \leq \ell_1(\tau_i^-) \) and \( x(\tau_i^\ast) + \xi_i^\ast = \beta(\tau_i^\ast) \) if \( x(\tau_i^-) \geq \ell_2(\tau_i^-). \) Since \( \xi_i^\ast = \alpha(\tau_i^\ast) - x(\tau_i^\ast) > 0 \) and \( \xi_i^\ast = \beta(\tau_i^\ast) - x(\tau_i^\ast) < 0, \) we have

\[
\begin{align*}
\frac{1}{2} p_1(\tau_i^-) x(\tau_i^-)^2 + q_1(\tau_i^-) x(\tau_i^-) + n_1(\tau_i^-) \\
= \frac{1}{2} p_1(\tau_i^\ast)^2 \alpha(\tau_i^\ast)^2 + (z_1 + q_1(\tau_i^\ast)) \alpha(\tau_i^\ast) + n_1(\tau_i^\ast) + \frac{1}{2} p_1(\tau_i^\ast) x(\tau_i^\ast)^2 + q_1(\tau_i^\ast) x(\tau_i^\ast) + n_1(\tau_i^\ast) \\
= \frac{1}{2} p_1(\tau_i^\ast) \beta(\tau_i^\ast)^2 + (-z_1 + q_1(\tau_i^\ast)) \beta(\tau_i^\ast) + n_1(\tau_i^\ast) + \frac{1}{2} p_1(\tau_i^\ast) x(\tau_i^\ast)^2 + q_1(\tau_i^\ast) x(\tau_i^\ast) + n_1(\tau_i^\ast) \\
+ z_1(\tau_i^\ast), x(\tau_i^-) \leq \ell_1(\tau_i^-) \\
+ \frac{1}{2} p_1(\tau_i^\ast) x(\tau_i^-)^2 + q_1(\tau_i^-) x(\tau_i^-) + n_1(\tau_i^-) \\
= \frac{1}{2} p_1(\tau_i^\ast) \beta(\tau_i^\ast)^2 + (-z_1 + q_1(\tau_i^\ast)) \beta(\tau_i^\ast) + n_1(\tau_i^\ast) + \frac{1}{2} p_1(\tau_i^\ast) x(\tau_i^\ast)^2 + q_1(\tau_i^\ast) x(\tau_i^\ast) + n_1(\tau_i^\ast) \\
+ z_1(\tau_i^\ast), x(\tau_i^-) \geq \ell_2(\tau_i^-) .
\end{align*}
\]

The above equations and continuity of \( p_1 \) and \( q_1 \) (from Assumption 7) imply that, at the impulse instants, the following conditions are satisfied:

\[
\begin{align*}
n_1(\tau_i^-) &= n_1(\tau_i^+) + \frac{1}{2} p_1(\tau_i^+) \alpha(\tau_i^-)^2 - \frac{1}{2} p_1(\tau_i^-) x(\tau_i^-)^2 \\
&- (q_1(\tau_i^-) + z_1) x(\tau_i^-) \\
&+ (z_1 + q_1(\tau_i^+)) \alpha(\tau_i^-), x(\tau_i^-) \leq \ell_1(\tau_i^-) \\
n_1(\tau_i^-) &= n_1(\tau_i^+) + \frac{1}{2} p_1(\tau_i^+) \beta(\tau_i^-)^2 - \frac{1}{2} p_1(\tau_i^-) x(\tau_i^-)^2 \\
&- (q_1(\tau_i^-) - z_1) x(\tau_i^-) \\
&- (z_1 - q_1(\tau_i^+)) \beta(\tau_i^-), x(\tau_i^-) \geq \ell_2(\tau_i^-).
\end{align*}
\]  

(39a, 39b)

At the initial time, \( t = 0 \), the term \( n_1(0) \) in the expression of the value function of Player 1 captures the cost of impulse accumulated over the entire time horizon of the game for Player 1.

\[
\square
\]

B. Impulse Control Problem of Player 2

Player 2 solves the QVIs associated with her impulse control problem for a given equilibrium strategy \( \gamma^\ast \) of Player 1.

In the continuation region, Player 2’s running cost is quadratic in the state, and it is linear in the state in the intervention region. Therefore, we can make the conjecture in Assumption 8 on the form of the value function of Player 2. A similar assumption on the form of the value function was made in [17] to obtain analytical solutions for an impulse optimal control problem. The value function \( V_2 \) coincides with continuous and continuously differentiable function \( \Phi_2 \) in the continuation region \( C. \) We conjecture that \( \Phi_2 \) is quadratic in the 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 afterward.

When the state lies in the continuation region, that is, \( x(t) \in (\ell_1(t), \ell_2(t)), \) the value function of Player 2 satisfies (14a) with equality

\[
\nabla \Phi_2(t, x(t)) + (\nabla \cdot \Phi_2(t, x(t))) (ax(t) + b \gamma^\ast(t, x(t))) \\
+ \frac{1}{2} w_2 (x(t) - \rho_2)^2 = 0.
\]

Substituting the partial derivatives of \( \Phi_2(t, x(t)) \) and the equilibrium control of Player 1 from (35) in the above equation, and comparing the coefficients, we obtain

\[
\begin{align*}
p_2(t) &= -w_2 - 2p_2(t) \left( \frac{\theta}{2} - \frac{\theta}{\nu_1 e^{\nu_1 t} + 1} \right) \\
q_2(t) &= \left( \frac{\theta}{2} - \frac{\theta}{\nu_1 e^{\nu_1 t} + 1} \right) q_2(t) - b_2 p_2(t) q_1(t) + w_2 \rho_2 \\
n_2(t) &= -b_2 q_1(t) q_2(t) - \frac{1}{2} w_2 \rho_2^2
\end{align*}
\]

(40a, 40b, 40c)

where \( p_2(T) = s_2, q_2(T) = -s_2 p_2, \) and \( n_2(T) = \frac{1}{2} s_2 \rho_2^2. \)
The solutions of (40a) and (40b) are given by
\[
p_2(t) = \frac{-w_2 e^{2t \nu_2} \nu_1^2 - 2t \theta w_2 e^{t \nu_1} + w_2 + \nu_2 \theta e^{t \nu_1}}{\theta (\nu_1 e^{t \theta} + 1)^2}
\]
\[
q_2(t) = \frac{e^{\nu_1 t} \nu_1^2 \nu_3 b_2 \left( -\nu_2 + \nu_1 w_2 (1 - \nu_1 e^{t \theta}) \right)}{\nu_1 r_1 \theta^3 (\nu_1 e^{t \theta} + 1)^2} + \frac{2 \rho_2 w_2 (\nu_1 e^{t \theta} - 1)}{\theta (\nu_1 e^{t \theta} + 1)^2} + \frac{\nu_2 e^{t \theta}}{\nu_1 e^{t \theta} + 1}
\]
\[
+ \frac{4 b_2 w_2 \rho_1 e^{t \theta} \left( w_2 e^{-t \theta} - \nu_1^2 w_2 e^{-t \theta} - \nu_2 \theta + 2 \nu_1 t w_2 \right)}{\nu_1 \theta^3 (\nu_1 e^{t \theta} + 1)^2}
\]
(41)

where constants \( \theta, \nu_1, \nu_2, \nu_3, \nu_4 \) are given in (32a), (32b), (32c), (61a), and (61b), respectively.

Assumption 9 ensures 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 \( \mathcal{C} \).

1) Characterization of Intervention set and Continuation set: In the intervention region \( ((t, x(t)) \in \mathcal{I}) \), (14b) holds with equality, that is,
\[
V_2(t, x(t)) = R V_2(t, x(t)) = \min_{\eta \in \Omega_2} V_2(t, x(t) + \eta) + h(\eta).
\]
(43)

For the problem parameters assumed in this section, \( V_2 \) is strictly convex in \( x(t) \) (see Assumption 9) and continuously differentiable for \( y = x(t) + \eta \in \mathcal{C} \). Since \( \alpha(t), \beta(t) \in \mathcal{C} \), \( x(t) + \eta \) takes a value of \( \alpha(t) \) or \( \beta(t) \) at the intervention instants and the derivative of \( y \) with respect to \( \eta \) is equal to 1, we can use the first-order conditions to obtain
\[
\nabla_y \Phi_2(t, \alpha(t)) + \frac{\partial h(\eta)}{\partial \eta} = 0, \quad x(t) \leq \ell_1(t)
\]
\[
\nabla_y \Phi_2(t, \beta(t)) + \frac{\partial h(\eta)}{\partial \eta} = 0, \quad x(t) \geq \ell_2(t).
\]
(44)
(45)

Using the quadratic form of the value function in (30) for the state value in the continuation region \( (\ell_1(t), \ell_2(t)) \), we get
\[
\nabla_y \Phi_2(t, \alpha(t)) = p_2(t) \alpha(t) + q_2(t) = -c
\]
\[
\nabla_y \Phi_2(t, \beta(t)) = p_2(t) \beta(t) + q_2(t) = d.
\]
(46)

Therefore, the following functions \( \alpha(\cdot) \) and \( \beta(\cdot) \) give the state values after an impulse occurs at equilibrium:
\[
\alpha(t) = \frac{-q_2(t) + c}{p_2(t)} \quad \forall t \in [0, T]
\]
\[
\beta(t) = \frac{d - q_2(t)}{p_2(t)} \quad \forall t \in [0, T].
\]
(47a)
(47b)

Since (14b) holds with equality in the intervention region, we have
\[
V_2(t, x(t)) = \begin{cases} 
V_2(t, \alpha(t)) + C + c(\alpha(t) - x(t)) & x(t) \leq \ell_1(t) \\
V_2(t, \beta(t)) + D + d(x(t) - \beta(t)) & x(t) \geq \ell_2(t).
\end{cases}
\]
(48)

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 \( x(t) = \ell_1(t) \) and \( x(t) = \ell_2(t) \), we substitute (30) in the above equations and simplify to obtain
\[
\frac{1}{2} p_2(t) \ell_1(t)^2 + q_2(t) \ell_1(t) = \frac{1}{2} p_2(t) \alpha(t)^2 + q_2(t) \alpha(t) + C + c(\alpha(t) - \ell_1(t))
\]
\[
\frac{1}{2} p_2(t) \ell_2(t)^2 + q_2(t) \ell_2(t) = \frac{1}{2} p_2(t) \beta(t)^2 + q_2(t) \beta(t) + D + d(\ell_2(t) - \beta(t)).
\]
(49a)
(49b)

To characterize the left boundary of the continuation region, we substitute \( \alpha(t) \) in (49a) to get
\[
p_2(t) \ell_1(t)^2 + 2(q_2(t) + c) \ell_1(t)
\]
\[
+ \frac{(q_2(t) + c)^2}{p_2(t)} - 2C = 0.
\]

Since \( C > 0, p_2(t) > 0, \) and \( \ell_1(t) < \alpha(t) \), the left boundary of the continuation region is given by
\[
\ell_1(t) = -\frac{c - q_2(t) - \sqrt{2C p_2(t)}}{p_2(t)}.
\]
(50a)

On substituting \( \beta(t) \) in (49b), we obtain the right boundary of the continuation region
\[
p_2(t) \ell_2(t)^2 + 2(2q_2(t)) + d \ell_2(t)
\]
\[
+ \frac{(d - q_2(t))^2}{p_2(t)} - 2D = 0.
\]

From \( D > 0, p_2(t) > 0 \) and \( \ell_2(t) > \beta(t) \), we obtain
\[
\ell_2(t) = -\frac{q_2(t) + d + \sqrt{2D p_2(t)}}{p_2(t)}.
\]
(50b)

By construction, \( V_1(t, x(t)) \) satisfies the sufficient conditions in (13), and therefore, \( V_1 \) is a value function of Player 1. In the next theorem, we give conditions under which \( V_2(t, x(t)) \) in (29) satisfies the QVIs (14).

Theorem 4: Let Assumptions 5–9 hold. \( V_2(t, x(t)) \) in (29) is the value function of Player 2 if \( \ell_1(t) \leq x_{12}(t) \) and \( \ell_2(t) \geq x_{22}(t) \) for each \( t \in [0, T] \) where \( \ell_1(t) \) and \( \ell_2(t) \) are given in (50a) and (50b), respectively
\[
x_{11}(t) = \frac{(c_1 + w_2 + p_2)}{\nu_2 w_2}, \quad w_2 \neq 0
\]
\[
x_{22}(t) = \frac{(-d + w_2 p_2)}{\nu_2 w_2}, \quad w_2 \neq 0
\]
\[
\theta_\alpha(t) = c^2 a^2 + 2w_2 (c_2 a p_2 - \nabla y \Phi_2(t, \alpha(t)))
\]
\[
\theta_\beta(t) = d^2 a^2 - 2w_2 (d a p_2 - \nabla y \Phi_2(t, \beta(t)))
\]

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

Proof: See Appendix B. 


TABLE II
PARAMETERS FOR NUMERICAL EXAMPLE (NOMINAL OR BASELINE SCENARIOS)

| α  | b  | w₁ | s₁ | r₁ | z₁ | w₂ |
|----|----|----|----|----|----|----|
| 0.1| -0.3| 1  | 1  | 0.1| 2  | 2  |

Table II

Fig. 1. Evolution of the intervention region for the parameters in Table II.

Fig. 2. Value function for the parameters in Table II.

Fig. 3. Evolution of the intervention region for the parameters in Table II with \( w_2 = 3 \).

Since the gradient of \( \Phi_2 \) with respect to \( t \) depends on the equilibrium control of Player 1, there is a strong coupling between the verification conditions for Player 2 and the optimal control problem of Player 1. Also, the conditions \( \ell_1(t) \leq x_{11}(t) \) and \( \ell_2(t) \geq x_{22}(t) \) have to be satisfied during the whole duration of the game and as a result, in the next section, we numerically verify these conditions for the analytically obtained expressions for the equilibrium impulses.

V. NUMERICAL EXAMPLES

We consider an iLQDG with time horizon \( T = 1 \) and other problem parameters given in Table II for which the verification conditions in Theorem 4 hold. The parameters are chosen to numerically illustrate the theoretical results.

In Fig. 1, we illustrate the state feedback policy of Player 2 for the problem parameters in Table II. 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 higher than the running cost coefficient \( w_2 \), the functions \( \ell_2(t), \beta_2(t), \ell_1(t), \) and \( \alpha(t) \) converge over time toward the target state value \( \rho_2 = 4 \). Also, the fixed cost of intervention is small if the state crosses the upper boundary compared to the case when the state crosses the lower 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 15, 11, 0, and 2, the evolution of equilibrium state trajectories is given by \( x_{11}^*(t), x_{21}^*(t), x_{22}^*(t) \), and \( x_{23}^*(t) \), respectively; see Fig. 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 Fig. 1.

In Fig. 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 for Player 2. Therefore, in Fig. 3, we can see that \( \ell_1(\cdot), \alpha(\cdot), \beta(\cdot) \), and \( \ell_2(\cdot) \) are closer to the initial state of Player 2 near the initial time, as compared to Fig. 1. Here, \( x_{11}^*(t), x_{21}^*(t), x_{22}^*(t), \) and \( x_{23}^*(t) \) denote the equilibrium evolution of the state trajectory for initial state values of 15, 11, 0, and 2, respectively. The value functions of Player 1 and Player 2 at the initial time are given in Fig. 4 for different values of the initial state.

Next, we consider a game problem where both players incur state-dependent costs only at the terminal time so that \( w_1 = w_2 = 0 \). In Fig. 5, we can see that if the initial state is \( -15 \) \((x_3^*(t))\), there is an impulse at \( t = 0 \) and there are two interior impulses in the interval \((0, T)\). For the initial state value of 2 \((x_4^*(t))\) and 8 \((x_7^*(t))\), there are no impulses, while for an initial state value of 19 \((x_9^*(t))\), there is one interior impulse in the interval \((0, T)\).
A. Sensitivity Analysis

In this section, we show the impact of a change in the problem parameters on the boundaries of the continuation set and impulse levels characterized by $\ell_1(\cdot)$, $\ell_2(\cdot)$, $\alpha(\cdot)$, and $\beta(\cdot)$. For the problem parameters considered in this section, the verification conditions given in Theorem 4 are satisfied.

In Fig. 6, we see that narrowing of $\alpha(\cdot)$ and $\beta(\cdot)$ as well as the boundaries of the continuation set toward the target state value of Player 1 and Player 2 when there is an increase in cost coefficient ($r$) of exercising the piecewise continuous control $u$ while other problem parameters are fixed at their nominal values given in Table II. From Fig. 7, it is clear that as the fixed cost of positive impulse decreases, Player 2 shifts the lower boundary toward her own target set.

VI. CONCLUSION

In this article, 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 HJB equations with jumps and QVIs. To the best of our knowledge, this is the first characterization of FNE in differential games with impulse controls where at least one player uses piecewise-continuous controls. In this, our article 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 FNE in the scalar
LQDG based on certain regularity assumptions on the value
function that have been assumed in the literature (see, e.g., [17]
and [6]). In future work, we plan to relax these assumptions and
develop policy iteration-type algorithms [46] that can solve the
QVIs for the impulse player in the general class of differential
games with impulse control. Another future research direction
could be to derive the sufficient conditions for the existence of an
FNE in a stochastic differential game with impulse control using
the HJB equations associated with the optimal control problem
of Player 1 and QVIs related to the stochastic impulse control
problem of Player 2.

APPENDIX

A. Proof of Proposition 2

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

$$\sum_{i \geq 1} 1_{t \leq \tau_i \leq T} b_2(x(\tau_i), \xi_i) = 0$$

(52)

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

$$V_2(t, x(t)) \leq \int_t^T h_2(x(s), \gamma^*(s, x(s))) ds + s_2(x(T))$$

$$\leq \|h_2\|_\infty (T - t) + \|s_2\|_\infty.$$  

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

$$V_2(t, x(t)) + \epsilon > J_2(x(t), \gamma^*_{[t, T]}, \delta_{[t, T]}),$$

where the second inequality follows from Assumption 1. This
proves that the value function is bounded such that

$$|V_2(t, x(t))| \leq \|h_2\|_\infty (T - t) + \|s_2\|_\infty,$$

$$\forall (t, x(t)) \in [0, T] \times \mathbb{R}^n.$$  

(53)

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

$$V_2(t, x(t)) + \epsilon > J_2(x(t), \gamma_1, v^\epsilon),$$

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

Using the above relation and (53), 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:

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

Since $\epsilon > 0$ is arbitrarily chosen for an $\epsilon$-optimal strategy of
Player 2, the upper bound $K$ on the number of impulses is given
by (22) as $\epsilon \to 0$.

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

$$V_1(t, x(t))$$

$$\leq \int_t^T h_1(x(s), 0) ds + \sum_{i \geq 1} 1_{t \leq \tau_i < T} b_1(x(\tau_i), \xi_i) + s_1(x(T))$$

$$\leq \|h_1\|_\infty (T - t) + K \|b_1\|_\infty + \|s_1\|_\infty,$$

where the last inequality follows from the boundedness of $b_1$ and
$s_1$ in Assumption 1. For any $\epsilon > 0$, we take a strategy $\gamma_{[t, T]} \in \Gamma_{[t, T]}$ so that

$$V_1(t, x(t)) + \epsilon > J_1(x(t), \gamma_{[t, T]}, \delta_{[t, T]}),$$

$$\geq -\|h_1\|_\infty (T - t) - K \|b_1\|_\infty - \|s_1\|_\infty.$$  

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

B. Proof of Theorem 4

From (46), we have $\nabla_x V_2(t, \alpha(t)) = -c$ and $\nabla_x V_2$
$(t, \beta(t)) = d$. Using the strict convexity of $V_2$ in $x(t)$ for
$(t, x(t)) \in C$ (see Assumption 9), we obtain

$$-c < \nabla_x V_2(t, x(t)) < d \quad \forall (t, x(t)) : (x(t) \in (\alpha(t), \beta(t)).$$

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

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

$$R V_2(t, x(t))$$

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

(54)

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

Next, we derive the conditions under which the value function
of Player 2 satisfies (14a). For $x(t) < \ell_1(t)$, we have

$$V_2(t, x(t)) = \Phi_2(t, \alpha(t)) + C + c(\alpha(t) - x(t)),$$

$$\forall (t, x(t)) \in [0, T] \times \mathbb{R}^n.$$  

(55)

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

$$\nabla_1 V_2(t, x(t)) + \nabla x V_2(t, x(t)), \nabla x V_2(t, x(t))$$

$$= \nabla_1 V_2(t, x(t)) + \frac{1}{2} w_2 x(t) - \rho_2^2$$

$$+ \nabla x V_2(t, x(t))(ax(t) + y_{\ell_1(t)}(t) < x(t) < \ell_1(t) \beta^* (t, x(t))$$

$$= \nabla_1 \Phi_2(t, \alpha(t)) + (\nabla_2 \Phi_2(t, \alpha(t)) + c) \frac{d\alpha(t)}{dt} - \rho_2 x(t)$$

$$+ \frac{1}{2} w_2 x(t)^2 + \frac{1}{2} w_2 \rho_2^2 - w_2 x(t) \rho_2.$$  

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Substituting (46) in the above equation, we get the roots of the above equation for \( w_2 \neq 0 \) as follows:

\[
x_{12}(t), x_{12}(t) = \frac{ca + w_2p_2 \pm \sqrt{\theta_0(t)}}{w_2} \quad (56)
\]

where \( x_{12}(t) = x_{12}(t) \), and \( \theta_0(t) \) is given by (51c). Therefore, (14a) holds if \( w_2 \neq 0 \) and \( \ell_1(t) \leq x_{11}(t) \) and \( \theta_0(t) \geq 0 \) for all \( t \in [0,T] \), or if \( w_2 = 0 \) and

\[
\ell_1(t) \leq \frac{\nabla_t \phi_2(t, \alpha(t))}{ca}. \quad (57)
\]

Similarly, for \( x(t) > \ell_2(t) \), we obtain

\[
x_{21}(t), x_{22}(t) = \frac{-(da - w_2p_2) \pm \sqrt{\theta_3(t)}}{w_2} \quad (58)
\]

where \( x_{21}(t) = x_{21}(t) \) and \( \theta_3(t) \) is given by (51d). Therefore, (14a) holds if \( w_2 \neq 0 \) and \( \ell_2(t) \geq x_{22}(t) \) or if \( w_2 = 0 \) and

\[
\ell_2(t) \geq - \frac{\nabla_t \phi_2(t, \beta(t))}{da}. \quad (59)
\]

**C. Analytical Solution of ODE**

To solve the differential equation \( p_1(t) + b_x(p_1(t))^2 + 2ap_1(t) + w_1 = 0 \) for \( t \in (r_i, r_{i+1}), i \in \{0,1,\ldots,k\} \), we substitute \( p_1(t) = \frac{\zeta(t)}{b_x(t)} \) to obtain a second-order ordinary differential equation \( \zeta(t) + 2a\zeta(t) + b_xw_1\zeta(t) = 0 \). When \( \theta = 2\sqrt{a^2 - w_1b_x} \), the solution of this equation is

\[
\zeta(t) = e^{-at}(F_1e^{\frac{a}{2}t} + F_2e^{-\frac{a}{2}t})
\]

where \( F_1 \) and \( F_2 \) are constants.

So, \( p_1(t) \) is given by

\[
p_1(t) = \frac{\zeta(t)}{b_x(t)} = \frac{-a\zeta(t) + \theta}{b_x}e^{-at}(F_1e^{\frac{a}{2}t} - F_2e^{-\frac{a}{2}t})
\]

\[
= \frac{1}{b_x} \left(-a + \frac{\theta}{2} - \frac{\theta}{v_1e^{\theta t} + 1}\right).
\]

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

\[
\nu_1 = \left(\frac{2\theta}{\theta - 2b_x^2s_1 - 2a - 1}\right)e^{-\theta T}. \quad (60)
\]

The values of constants used in the expression of \( q_2(t) \) are given as

\[
\nu_3 = \left(\frac{2\rho_1 w_1 - \rho_2 s_1 \theta}{e^{\frac{\theta}{2}T} - \theta} - \nu_1(2\rho_1 w_1 + \rho_2 s_1 \theta)\right) \quad (61a)
\]

\[
\nu_4 = -s_2\rho_2 e^{-\frac{\theta}{2}T} \left(v_1 e^{\theta T} + 1\right) + \frac{\nu_2b^2(-v_1w_2(1 - v_1e^{\theta T})}{} + \frac{2\rho_2w_2(\nu_1 e^{\theta T} - 1)}{} + \frac{\theta}{e^{\frac{\theta}{2}T} + 1} \quad (61b)
\]

\[
- \frac{4b^2w_1p_1}{v_1 e^{\theta T} + 1} - \nu_1(2w_2 e^{-\theta T} - \nu_2\theta + 2\nu_1 T \theta w_2) \frac{r_1^3(1 + e^{-\frac{\theta}{2}T})}{r_1^3(1 + e^{-\frac{\theta}{2}T})}.
\]

**ACKNOWLEDGMENT**

The authors would like to thank Herbert Dawid for providing valuable comments and suggestions that helped them improve the article. Also, the authors would like to thank the four anonymous reviewers for their constructive and very helpful comments.

**REFERENCES**

[1] G. Ferrari and T. Koch, “On a strategic model of pollution control,” *Ann. Operations Res.*, vol. 275, no. 2, pp. 297–319, 2019.

[2] J. Yong, “Zero-sum differential games involving impulse controls,” *Appl. Math. Optim.*, vol. 29, no. 3, pp. 243–261, 1994.

[3] N. El Farouq, G. Barles, and P. Bernhard, “Deterministic minmax impulse control,” *Appl. Math. Optim.*, vol. 61, no. 3, pp. 353–378, Jun. 2010.

[4] B. Ekeland, “Deterministic minmax impulse control in finite horizon: The viscosity solution approach,” *ESAIM, Control, Optim., Calculus Variations*, vol. 19, no. 1, pp. 63–77, 2013.

[5] P. Azimzadeh, “A zero-sum stochastic differential game with impulses, precommitment, and unrestricted cost functions,” *Appl. Math. Optim.*, vol. 79, no. 2, pp. 483–514, 2019.

[6] R. Aïd, M. Basie, G. Callegaro, L. Campi, and T. Vargiolu, “Non-zero-sum stochastic differential games with impulse controls: A verification theorem with applications,” *Math. Operations Res.*, vol. 45, no. 1, pp. 205–232, 2020.

[7] R. Aïd, L. Campi, L. Li, and M. Ludkovski, “An impulse-regime switching game model of vertical competition,” *Dyn. Games Appl.*, vol. 11, pp. 662–691, 2021.

[8] M. Basie, H. Cao, and X. Guo, “Nonzero-sum stochastic games and mean-field games with impulse controls,” *Math. Operations Res.*, vol. 47, no. 1, pp. 341–366, 2022.

[9] U. Sadaña, P. V. Reddy, and G. Zaccour, “Nash equilibrium in nonzero-sum differential games with impulse control,” *Eur. J. Oper. Res.*, vol. 295, no. 2, pp. 792–805, 2021.

[10] B. Crettez and N. Hayek, “Terrorists’ eradication versus perpetual terror war,” *J. Optim. Theory Appl.*, vol. 160, pp. 679–702, 2014.

[11] A. Novak, G. Feichtinger, and G. Leitmann, “A differential game related to terrorism: Nash and Stackelberg strategies,” *J. Optim. Theory Appl.*, vol. 144, pp. 533–555, 2010.

[12] T. Başar and G. Olisder, *Dynamic Noncooperative Game Theory*, 2nd ed. Philadelphia, PA, USA: SIAM, 1999.

[13] 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, The Netherlands: Springer, 1989, pp. 9–54.

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

[15] J. C. Engwerda, *LQ Dynamic Optimization and Differential Games*. New York, NY, USA: Wiley, 2005.

[16] T. Başar, A. Haurie, and G. Zaccour, “Nonzero-sum differential games,” in *Handbook of Dynamic Game Theory*, T. Başar and G. Zaccour, Eds. Cham, Switzerland: Springer, 2018, pp. 61–110.

[17] D. Ben-Yaakov, W. J. Runggaldier, and K. Yasuda, “On classical and restricted impulse stochastic control for the exchange rate,” *Appl. Math. Optim.*, vol. 74, no. 2, pp. 423–454, Oct. 2016.

[18] R. Aïd, F. Bernal, M. Minif, D. Zalbaljaegui, and J. Zubelli, “A policy iteration algorithm for nonzero-sum stochastic impulse games,” *ESAIM: Proc. Surv.*, vol. 65, pp. 27–45, 2019.

[19] L. Campi and D. De Santis, “Nonzero-sum stochastic differential games between an impulse controller and a stopper,” *J. Optim. Theory Appl.*, vol. 186, no. 2, pp. 688–724, 2020.

[20] A. Bensoussan and J. L. Lions, *Contrôle Impulsionnel Et Inéquations Quasi-Variationnelles*. Paris, France: Gauthier-Villars, 1984.

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

[22] A. Cadenillas and F. Zapatero, “Optimal central bank intervention in the foreign exchange market,” *J. Econ. Theory*, vol. 87, no. 1, pp. 218–242, 1999.
Dr. Zaccour is the Editor-in-Chief of Dynamic Games and Applications. He is currently a Postdoctoral Fellow with the Desautels Faculty of Management, McGill University, Montreal. His research interests include game theory and the control of multiagent systems.

Utsav Sadana received a dual degree with the Bachelor of Technology degree in materials science and engineering and the Master of Science degree in economics from the Indian Institute of Technology Kanpur, Kanpur, India, in 2017, and the Ph.D. degree in management science from HEC Montréal, Montreal, QC, Canada, in 2021.

He is currently a Postdoctoral Fellow with the Desautels Faculty of Management, McGill University, Montreal. His research interests include game theory, deep learning, and robust optimization.

Puduru Viswanadha Reddy (Member, IEEE) received the Ph.D. degree in operations research from Tilburg University, Tilburg, The Netherlands, in 2011.

He is currently an Assistant Professor with the Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai, India. His current research interests include game theory and the control of multiagent systems.

Georges Zaccour received the M.Sc. degree in international business and the Ph.D. degree in management science from HEC Montréal, Montréal, QC, Canada, in 1983 and 1987, respectively, and a license in mathematics and economics from Université Paris-Dauphine, Paris, France, in 1981.

He is currently the Chair in Game Theory and Management and a Full Professor of Management Science with HEC Montréal. He was the Director of GERAD, an interuniversity research center. He coauthored the books Differential Games in Marketing and Games and Dynamic Games. His research interests include differential games, optimal control and operations research applied to marketing, energy sector, and environmental management, in which he has authored or coauthored more than 280 papers and co-edited 15 volumes.

Dr. Zaccour is the Editor-in-Chief of Dynamic Games and Applications and a Fellow of The Royal Society of Canada. He was the President of the International Society of Dynamic Games from 2002 to 2006.