https://doi.org/10.1109/CDC.2017.8264504

Inverse Noncooperative Differential Games

Timothy L. Molloy, Jason J. Ford, and Tristan Perez

***Abstract*— Differential games are an important mathematical tool for studying conflict in applications across engineering, economics, and ecology. In noncooperative differential games, the players (or decision makers) are concerned with achieving their individual objectives under the dynamic constraints of a system governed by differential equations without relying on the cooperation of others. In this paper, we propose three methods of inverse noncooperative differential games that enable the recovery of player cost-functional parameters from open-loop Nash equilibria. We first propose a nested optimisation approach that provides a direct method for solving inverse differential game problems. We then develop connections between our inverse noncooperative differential game problem and the existing problem of inverse optimal control to propose two alternative methods with modest implementation and computational demands. We illustrate and compare our three methods in simulation.**

I. Introduction

Differential games provide a rich framework for analysing conflicts between multiple players (or decision makers) in continuous-time dynamical systems across fields including automatic control [1], [2], economics [3], management [3], and ecology [4]. In these fields, prior work in the area of differential games has primarily been concerned with developing techniques for identifying and analysing game solutions under a variety of optimality concepts, information structures, and cooperation assumptions. Noncooperative differential games have arguably attracted the most attention in the field of automatic control due to their utility in solving robust control problems [1], and their applicability to real-world problems such as vehicle collision avoidance [2]. A considerable body of literature therefore exists for solving noncooperative differential games to identify player controls (or decisions) that achieve the objectives of individual players despite conflicts of interest. In contrast, the inverse problem of recovering the underlying objectives of individual players from solutions to noncooperative differential games has received limited attention. In this paper, we consider the inverse differential game problem of computing the underlying objectives of individual players in an N-player noncooperative differential game from example game solutions.

A noncooperative differential game involves multiple players individually selecting controls for a system governed by differential equations with the aim of minimising their individual cost functionals [1]. Noncooperative differential games are solvable under various information structures governing the information players have access to about the system state and the player controls. Open-loop Nash equilibrium solutions are of particular interest since they provide a benchmark for the value of the game [5]. These open-loop Nash equilibrium solutions describe the controls that players should select given only knowledge of the initial game state so that they can do no better when every other player selects their open-loop Nash equilibrium controls [1]. When solved under this open-loop information structure, N-player noncooperative differential games are therefore analogous to solving N coupled optimal control problems.

The majority of recent work on differential games has focused on establishing properties of equilibrium solutions. However, in the related fields of optimal control, static game theory, and optimisation, there has been interest in inverse problems involving the recovery of underlying objectives from observed optimal or Nash equilibrium solutions [6]–[13]. For example, [14] used necessary optimality conditions to propose an approach for recovering the parameters of optimisation problems from optimal solutions. Similar approaches based on analogous necessary optimality conditions have since been developed to solve inverse static game [7], [8], [13] and inverse optimal control problems [9], [10], [15].

The literature concerned with dynamic games (i.e., games with state dynamics) has also started to examine the inverse problem of estimating the underlying objectives of players from Nash equilibrium solutions [16]–[21]. Most of these inverse dynamic game problems have been posed and solved in the discrete-time setting with the state evolving according to either a Markov decision process with finite-state and control space [16] or a set of difference equations [18]–[20]. Although some progress has been made in posing and solving inverse dynamic game problems in the continuous-time setting where the states evolve according to finite-state Markov processes [17], limited progress has been made in posing and solving inverse differential game problems (i.e., dynamic games with state dynamics governed by differential equations). For example, a two-player inverse differential game problem was posed and solved in [21] under the restrictive assumption that the control law of one player is known so that the problem reduces to solving an inverse optimal control problem.

In this paper, we consider inverse N-player noncooperative differential games under an open-loop information structure – including their reduction to inverse optimal control problems without the restrictive assumption of [21].

The authors are with the School of Electrical Engineering and Computer Science, Queensland University of Technology (QUT), Brisbane, QLD 4000, Australia. Tristan Perez is also with the Institute for Future Environments at QUT and the ARC Centre of Excellence in Mathematical and Statistical Frontiers (ACEMS). {t.molloy, j2.ford, tristan.perez}@qut.edu.au

This work was conducted under the Australian Research Council Linkage Grant LP130100483 with the support of Boeing Research and Technology Australia.

Although open-loop differential games can also be reformulated as (static) infinite-dimensional continuous games (cf. [22]), our reduction of inverse differential game problems to inverse optimal control problems rather than inverse continuous games is fruitful due to the range of control-theoretic techniques for inverse optimal control with continuous-time systems governed by differential equations [10]–[12]. For example, [12] proposes a bilevel optimisation approach for recovering parameters of an optimal control cost-functional from optimal state and control trajectories, and [10] proposes an elegant alternative approach that exploits the minimum principle of optimal control.

The key contribution of this paper is the proposal of three approaches for solving open-loop inverse N-player noncooperative differential game problems. We first propose a direct method for solving our inverse differential game problem without exploiting connections to inverse optimal control. We then exploit connections between our inverse differential game problem and the problem of inverse optimal control to propose two simplified methods of inverse differential games that are similar to the bilevel optimisation [12] and minimum principle [10] methods of inverse optimal control.

This paper is structured as follows. In Section II, we pose our inverse noncooperative differential game problem. In Section III, we propose our three methods of inverse noncooperative differential games. In Section IV, we present an illustrative simulation example before conducting a thorough simulation study of our three proposed methods. Finally, we present conclusions in Section V.

II. Problem Statement

We will consider a continuous-time N-player (potentially nonzero-sum) noncooperative differential game played with an open-loop information structure. Let us define the continuous-time (potentially nonlinear) state process of the game over the time interval $t \in [0, T]$ as

$$\dot{x}(t) = f\left(t, x(t), u^1(t), \dots, u^N(t)\right), \quad x(0) = x_0 \tag{1}$$

where $x_0 \in \mathbb{R}^n$ is the initial state vector, $f(\cdot, \cdot, \dots, \cdot)$ are (potentially nonlinear) functions that satisfy the usual conditions for ensuring the uniqueness of solutions (cf. [1, p. 226]), and $u^i(t) \in U^i$ are control inputs belonging to the control sets $U^i \subset \mathbb{R}^{m_i}$ for $i \in \mathbb{N} \triangleq \{1, \dots, N\}$. Let x denote the entire state trajectory, namely $x(t)$ for $t \in [0, T]$, and let u^i denote the entire ith control trajectory, namely $u^i(t)$ for $t \in [0, T]$. We assume that the ith control trajectory u^i is selected by a player who is attempting to minimise the cost functional

$$\begin{aligned} &J^i\left(u^1, \dots, u^N, \theta^i\right) \\ &\quad \triangleq \int_0^T \theta^{i\prime} g^i\left(t, x(t), u^1(t), \dots, u^N(t)\right) dt \end{aligned} \tag{2}$$

subject to the constraints of the dynamics (1) and controls $u^i(t) \in U^i$ where $g^i(\cdot, \cdot, \cdot, \dots, \cdot) : [0, T] \times \mathbb{R}^n \times \mathbb{R}^{m_1} \times \ldots \times \mathbb{R}^{m_N} \mapsto \mathbb{R}^{M_i}$ are known basis functions, and $\theta^i \in \Theta^i \subset \mathbb{R}^{M_i}$ is a vector of parameters.

The solutions to the N-player noncooperative differential game (1) and (2) played with an open-loop information structure are open-loop Nash equilibria. The control trajectories u_θ^i for $i \in \mathbb{N}$ constitute an open-loop Nash equilibrium solution of the game (1) and (2) with parameters $\theta^i \in \Theta$ for $i \in \mathbb{N}$ if and only if the inequalities [1, p. 266]

$$\begin{aligned} J^1\left(u_\theta^1, u_\theta^2, \dots, u_\theta^N, \theta^1\right) &\leq J^1\left(u^1, u_\theta^2, \dots, u_\theta^N, \theta^1\right) \\ J^2\left(u_\theta^1, u_\theta^2, \dots, u_\theta^N, \theta^2\right) &\leq J^2\left(u_\theta^1, u^2, \dots, u_\theta^N, \theta^2\right) \\ &\vdots \\ J^N\left(u_\theta^1, u_\theta^2, \dots, u_\theta^N, \theta^N\right) &\leq J^N\left(u_\theta^1, u_\theta^2, \dots, u^N, \theta^N\right) \end{aligned} \tag{3}$$

hold for all admissible player control trajectories u^i. We shall denote the states associated with open-loop Nash equilibrium control trajectories u_θ^i for $i \in \mathbb{N}$ as $x_\theta(t)$ for $t \in [0, T]$, and we will let x_θ denote the open-loop Nash equilibrium state trajectory. Finally, let $\mathcal{U}^i\left(\theta^1, \dots, \theta^N\right)$ denote the set of open-loop Nash equilibrium control trajectories u_θ^i for player i in the N-player nonzero-sum noncooperative game (1) and (2) with parameters $\theta^i \in \Theta^i$ for $i \in \mathbb{N}$. Similarly, let $\mathcal{X}\left(\theta^1, \dots, \theta^N\right)$ denote the set of open-loop Nash equilibrium state trajectories x_θ for the N-player noncooperative game (1) and (2) with parameters $\theta^i \in \Theta^i$ for $i \in \mathbb{N}$.

In the inverse N-player noncooperative differential game problem, we are given the state trajectory x^* and the control trajectories u^{i*} for $i \in \mathbb{N}$ that satisfy the conditions (3) of an open-loop Nash equilibrium solution for the parameters $\theta^i = \theta^{i*} \in \Theta$ for $i \in \mathbb{N}$. We assume that some (or all) of the player cost-functional parameters θ^{i*} are unknown. Our aim is to recover the unknown parameters θ^{i*} from open-loop Nash equilibrium trajectories given knowledge of the dynamics $f(\cdot, \cdot, \dots, \cdot)$, the constraint sets U^i, and the basis functions $g^i(\cdot, \cdot, \dots, \cdot)$ for $i \in \mathbb{N}$. We shall make no assumptions about the linearity of the state dynamics (1) or the convexity of the cost functionals (2).

By inspecting (3), we note that the unknown parameters θ^{i*} will only be recoverable up to an unknown scale since if x^* and u^{i*} constitute an open-loop Nash equilibrium of the differential game (1) and (2) with $\theta^i = \theta^{i*}$, then they are also an open-loop Nash equilibrium of the game with $\theta^i = r^i \theta^{i*}$ for any $0 < r^i < \infty$ and $i \in \mathbb{N}$. We also note that $\theta^i = 0$ is a trivial solution to the inverse differential game problem. Without loss of generality, in this paper we enforce unambiguous scaling and avoid trivial solutions by considering the sets $\Theta^i \triangleq \{\theta^i \in \mathbb{R}^{M_i} : \theta_1^i = 1\}$ for all $i \in \mathbb{N}$. Similar approaches to avoiding non-unique scaling are used widely in inverse optimal control literature (cf. [10], [12], [23]), along with approaches that constraint the sum of the elements of θ^i to 1 (cf. [24]).

III. Methods of Inverse Differential Games

In this section, we propose three methods for solving the inverse differential game problem. Our first method solves the inverse differential game problem by finding player cost-functional parameters that lead to matching open-loop Nash equilibrium trajectories. We then exploit connections between optimal control and differential games to propose our other two methods of inverse differential games.

A. Proposed Nested Optimisation (NOP) Method

To present our first method of inverse differential games, let us define the objective functional

$$\mathcal{J}_T\left(x, u^1, \ldots, u^N\right) = \int_0^T \left[\|x(t) - x^*(t)\|^2 + \sum_{i=1}^N \|u^i(t) - u^{i*}(t)\|^2\right] dt. \tag{4}$$

This objective functional provides a natural squared-error metric between candidate state x and control trajectories u^i, and the open-loop Nash equilibrium state x^* and control trajectories u^{i*} where $i \in \mathbb{N}$. In our nested optimisation (NOP) method, we therefore propose finding player cost-functional parameters θ^i for $i \in \mathbb{N}$ by solving the optimisation problem

$$\begin{aligned} \inf_{\theta^1,\ldots,\theta^N} \quad & \mathcal{J}_T\left(x_\theta, u_\theta^1, \ldots, u_\theta^N\right) \\ \text{s.t.} \quad & x_\theta \in \mathcal{X}\left(\theta^1, \ldots, \theta^N\right) \\ & u_\theta^i \in \mathcal{U}^i\left(\theta^1, \ldots, \theta^N\right),\ i = 1, \ldots, N \\ & \theta^i \in \Theta^i,\ i = 1, \ldots, N. \end{aligned} \tag{5}$$

Our proposed NOP method (5) intuitively aims to find player cost-functional parameters θ^i that give rise to open-loop Nash equilibrium trajectories that closely match the open-loop Nash equilibrium trajectories associated with the unknown parameters θ^{i*} for $i \in \mathbb{N}$. The NOP method involves nested optimisation since the constraints in (5) are met by solving the N-player noncooperative differential game (1) and (2) with candidate parameters θ^i for open-loop Nash equilibrium trajectories x_θ and u_θ^i for $i \in \mathbb{N}$. Solving N-player noncooperative differential games for open-loop Nash equilibria is nontrivial in general, and so the implementation of our proposed NOP method is typically demanding and complex. For example, we later implement our NOP method using a derivative-free numeric optimisation technique, and a numeric two-point boundary value problem solver to solve the N-player noncooperative differential games. We shall now propose two simplified methods of inverse differential games that do not require the nested solution of differential games for candidate cost-functional parameters.

B. Proposed Bilevel Optimisation (BOP) Method

To present our second method of inverse differential games, recall that in our inverse noncooperative differential game problem, we are given the open-loop Nash equilibrium control trajectories u^{i*} for all $i \in \mathbb{N}$. By inspecting the conditions for open-loop Nash equilibria (3), we see that the ith player control trajectory u^{i*} is the solution to the optimal control problem

$$\begin{aligned} \inf_{u^i} \quad & J^i\left(u^{1*}, \ldots, u^i, \ldots, u^{N*}, \theta^{i*}\right) \\ \text{s.t.} \quad & \dot{x}(t) = f\left(t, x(t), u^{1*}(t), \ldots, u^i(t), \ldots, u^{N*}(t)\right) \\ & x(0) = x_0 \\ & u^i(t) \in U^i,\ t \in [0, T] \end{aligned} \tag{6}$$

given the $N-1$ open-loop Nash equilibrium control trajectories of the other players (u^{j*} with $j \in \mathbb{N}$, $j \neq i$). Hence, we may recover the parameters θ^{i*} for any player $i \in \mathbb{N}$ by solving the inverse optimal control problem corresponding to the optimal control problem (6). Our second proposed method of inverse differential games is therefore to solve an inverse optimal control problem for each player $i \in \mathbb{N}$ using the bilevel optimisation method of [12]. Specifically, our proposed bilevel optimisation (BOP) method is to solve the optimisation problem

$$\begin{aligned} \inf_{\theta^i} \quad & \bar{\mathcal{J}}_T\left(x_\theta, u_\theta^i\right) \\ \text{s.t.} \quad & \dot{x}_\theta(t) = f\left(t, x_\theta(t), u^{1*}(t), \ldots, u_\theta^i(t), \ldots, u^{N*}(t)\right) \\ & J^i\left(u^{1*}, \ldots, u_\theta^i, \ldots, u^{N*}, \theta^i\right) \\ & \quad \leq J^i\left(u^{1*}, \ldots, u^i, \ldots, u^{N*}, \theta^i\right)\ \forall\, u^i \in U^i \\ & x_\theta(0) = x_0 \\ & u_\theta^i(t) \in U^i,\ t \in [0, T] \\ & \theta^i \in \Theta^i \end{aligned} \tag{7}$$

for all $i \in \mathbb{N}$ where we define the squared error objective functional

$$\bar{\mathcal{J}}_T\left(x, u^i\right) = \int_0^T \left[\|x(t) - x^*(t)\|^2 + \|u^i(t) - u^{i*}(t)\|^2\right] dt.$$

In contrast to our NOP method, our BOP method recovers the parameters of a single player without needing to compute those of the other players. Furthermore, our BOP method avoids solving nested differential games by instead solving optimal control problems. Since the solution of optimal control problems is usually easier than the solution of differential games, implementations of our proposed BOP method will usually be less complex and less computationally expensive than implementations of our NOP method. We shall now propose our final method of inverse differential games that avoids the need to solve nested optimal control and differential game problems for candidate parameters.

C. Proposed Minimum Principle (MP) Method

Our third proposed method of inverse differential games is inspired by the minimum principle method of inverse optimal control proposed in [10]. In contrast to [10], we shall propose our method by exploiting necessary conditions for open-loop Nash equilibria.

1) Necessary Conditions for Open-Loop Nash Equilibria: Let us define the Hamiltonian of the ith player, $i \in \mathbb{N}$, as

$$\begin{aligned} & H^i\left(t, \lambda^i(t), x(t), u^1(t), \ldots, u^N(t), \theta^i\right) \\ & \quad \triangleq \theta^{i\prime} g^i\left(t, x(t), u^1(t), \ldots, u^N(t)\right) \\ & \qquad + \lambda^{i\prime}(t) f\left(t, x(t), u^1(t), \ldots, u^N(t)\right) \end{aligned}$$

for $t \in [0, T]$ where $\lambda^i(\cdot) : [0, T] \mapsto \mathbb{R}^n$ for $t \in [0, T]$ are costate (or adjoint) functions. Let us also introduce the following assumptions.

Assumption 1: The function $f\left(t, \cdot, u^1(t), \ldots, u^N(t)\right)$ is continuously differentiable on $\mathbb{R}^n$ for all $t \in [0, T]$.

Assumption 2: The player cost-functional basis functions $g^i\left(t, \cdot, u^1(t), \ldots, u^N(t)\right)$ are continuously differentiable on $\mathbb{R}^n$ for all $i \in \mathbb{N}$ and all $t \in [0, T]$.

Assumption 3: The open-loop Nash equilibrium controls $u^{i*}(t)$ of the N-player noncooperative dynamic game (1) and (2) with $\theta^i = \theta^{i*}$ are in the interior (i.e., not the boundaries) of the control constraint sets U^i for all for $t \in [0, T]$ and all $i \in \mathbb{N}$.

Assumptions 1 and 2 are conditions on the dynamics and cost functional of the game that are trivially satisfied by linear dynamics with quadratic cost functional. Both of these assumptions are standard in the study of N-player noncooperative differential games (cf. [1, Chapter 6]). In contrast, Assumption 3 is potentially restrictive since it excludes the use of open-loop Nash equilibrium trajectories with active control constraints. It is however easily checked given knowledge of the player control trajectories u^{i*} and constraint sets U^i.

Let us now define the shorthand notation $f_t \triangleq f\left(t, x^*(t), u^{1*}(t), \ldots, u^{N*}(t)\right)$, and $g_t^i \triangleq g^i\left(t, x^*(t), u^{1*}(t), \ldots, u^{N*}(t)\right)$, for $i \in \mathbb{N}$. We shall denote the matrix of partial derivatives of $f(\cdot, \cdot, \ldots, \cdot)$ with respect to $x(t)$ evaluated at t, $x^*(t)$ and $u^{i*}(t)$ for $i \in \mathbb{N}$ as

$$\nabla_x f_t \triangleq \begin{bmatrix} \frac{\partial f_t^1}{\partial x_1(t)} & \cdots & \frac{\partial f_t^n}{\partial x_1(t)} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_t^1}{\partial x_n(t)} & \cdots & \frac{\partial f_t^n}{\partial x_n(t)} \end{bmatrix}.$$

We shall similarly use $\nabla_{u^i} f_t$ to denote the matrix of partial derivatives of $f(\cdot, \cdot, \ldots, \cdot)$ with respect to $u^i(t)$ evaluated at t, $x^*(t)$ and $u^{i*}(t)$. Finally, let $\nabla_x g_t^i$ and $\nabla_{u^i} g_t^i$ denote the matrices of partial derivatives of $g^i(\cdot, \cdot, \ldots, \cdot, \cdot)$ by $x(t)$ and $u^i(t)$, respectively (evaluated at $x^*(t)$, $u^{i*}(t)$ for $i \in \mathbb{N}$).

Under Assumptions 1 and 2, Theorem 6.11 of [1] gives that if x^* and u^{i*} for $i \in \mathbb{N}$ are open-loop Nash equilibrium trajectories of the N-player noncooperative differential game (1) and (2) with parameters $\theta^i = \theta^{i*} \in \Theta$ for $i \in \mathbb{N}$, then the player costate functions solve the differential equations

$$\dot{\lambda}^i(t) = -\nabla_x H^i\left(t, \lambda^i(t), x^*(t), u^{1*}(t), \ldots, u^{N*}(t), \theta^{i*}\right) \tag{8}$$

with terminal boundary condition $\lambda^{i\prime}(T) = 0$ for $i \in \mathbb{N}$. Furthermore, the player controls satisfy

$$u^{i*}(t) = \arg\min_{\bar{u} \in U^i} H^i\left(t, \lambda^i(t), x^*(t), u^{1*}(t), \ldots, u^{i-1*}(t), \bar{u}, u^{i+1*}(t), \ldots, u^{N*}(t), \theta^{i*}\right)$$

for all $t \in [0, T]$ and all $i \in \mathbb{N}$, which, under Assumption 3, reduces to the Hamiltonian gradient conditions that

$$\nabla_{u^i} H^i\left(t, \lambda^i(t), x^*(t), u^{1*}(t), \ldots, u^{N*}(t), \theta^{i*}\right) = 0 \tag{9}$$

for all $t \in [0, T]$ and all $i \in \mathbb{N}$.

By recalling the parameterisation of the player cost functionals (2) and our shorthand notation, we may rewrite the necessary conditions expressed in the differential equations (8) and (9) for $i \in \mathbb{N}$ as the system of equations

$$F^i(t) z^{i*}(t) + G^i(t) v^{i*}(t) = 0 \tag{10}$$

where $v^{i*}(t) \triangleq \dot{\lambda}^i(t)$, $G^i(t) \triangleq [I_n, 0]' \in \mathbb{R}^{(m_i+n)\times n}$,

$$F^i(t) \triangleq \begin{bmatrix} \nabla_x g_t^i & \nabla_x f_t \\ \nabla_{u^i} g_t^i & \nabla_{u^i} f_t \end{bmatrix}, \text{ and } z^{i*}(t) \triangleq \begin{bmatrix} \theta^{i*} \\ \lambda^i(t) \end{bmatrix}.$$

It follows that for $i \in \mathbb{N}$,

$$\dot{z}^{i*}(t) = B^i v^{i*}(t)$$

where $B^i \triangleq [0, I_n]' \in \mathbb{R}^{(M_i+n)\times n}$, and the terminal boundary condition implies that

$$z^{i*}(T) = \begin{bmatrix} \theta^{i*} \\ 0 \end{bmatrix}. \tag{11}$$

Here, we have used I_n to denote the $n \times n$ identity matrix. We now exploit the system of equations (10) to propose our minimum principle (MP) method of inverse differential games.

2) Minimum Principle (MP) Method: Our proposed MP method of inverse differential games is to find player cost-functional parameters and costate variables that minimise the violation of the necessary conditions of open-loop Nash equilibria (10) that hold exactly under Assumptions 1 – 3. Specifically, our proposed MP method is to identify functions $z^i(\cdot) : [0, T] \mapsto \mathbb{R}^{M_i+n}$ and $v^i(\cdot) : [0, T] \mapsto \mathbb{R}^n$ for each player $i \in \mathbb{N}$ that solve the optimal control problem

$$\begin{aligned} \inf_{z^i(\cdot), v^i(\cdot)} \quad & \int_0^T \|F^i(t) z^i(t) + G^i(t) v^i(t)\|^2 \, dt \\ \text{s.t.} \quad & \dot{z}^i(t) = B^i v^i(t). \end{aligned} \tag{12}$$

We may then simply extract the parameters of player's cost-functional $\theta^{i*} \in \Theta$ from the corresponding components of $z^i(t)$ for any $t \in [0, T]$ (under the constraints imposed by the set Θ).

Although our method appears to require the solution of N optimal control problems of the form (12) with free initial states, we note that the system dynamics are linear, and the objective functional is quadratic in the sense that

$$\begin{aligned} &\|F^i(t) z^i(t) + G^i(t) v^i(t)\|^2 \\ &\quad = z^{i\prime}(t) Q^i(t) z^i(t) + v^{i\prime}(t) R^i(t) v^i(t) + 2 z^{i\prime}(t) S^i(t) v^i(t) \end{aligned}$$

for $i \in \mathbb{N}$ where $Q^i(t) \triangleq F^{i\prime}(t) F^i(t)$, $R^i(t) \triangleq G^{i\prime}(t) G^i(t)$, and $S^i(t) \triangleq F^{i\prime}(t) G^i(t)$. We may therefore apply the standard tools of linear quadratic (LQ) optimal control to solve our inverse differential game problem (regardless of the linearity of the system dynamics (1)). The following theorem establishes a useful approach for solving our proposed MP problem (12).

Theorem 1: Consider any player $i \in \mathbb{N}$. If $R^i(t)$ is positive definite for all $t \in [0, T]$, and

$$Q^i(t) - S^i(t)(R^i(t))^{-1} S^{i\prime}(t)$$

is nonnegative definite for all $t \in [0, T]$, then the function $z^i(\cdot)$ solving (12) is the solution to the differential equation

$$\dot{z}^i(t) = B^i K^i(t) z^i(t) \tag{13}$$

for $t \in [0,T]$ with initial conditions $z^i(0) = z_0^i$ given by the solution to the optimisation problem

$$\inf_{z_0^i \in \mathbb{R}^{M_i+n}} z_0^{i\prime} P^i(0) z_0^i. \tag{14}$$

Here,

$$K^i(t) \triangleq -(R^i(t))^{-1} \left[B^{i\prime} P^i(t) + S^{i\prime}(t) \right]$$

and $P^i(\cdot) : [0,T] \mapsto \mathbb{R}^{(M_i+n)\times(M_i+n)}$ is the solution to the Riccati differential equation

$$0 = \dot{P}^i - (P^i B^i + S^i)(R^i)^{-1}(B^{i\prime} P^i + S^{i\prime}) + Q^i \tag{15}$$

with $P^i(T) = 0$ where we have omitted the time argument for brevity.

Proof: Under the theorem conditions on $R^i(t)$ and $Q^i(t) - S^i(t)(R^i(t))^{-1} S^{i\prime}(t)$, [25, Section 3.4] gives that the control function $v^i(t)$ solving the LQ optimal control problem (12) for any initial state $z^i(0) = z_0^i$ is given by

$$v^i(t) = K^i(t) z^i(t)$$

for all $t \in [0,T]$. The first theorem result (13) follows by substituting the optimal control $v^i(t) = K^i(t) z^i(t)$ into the dynamics $\dot{z}^i(t) = B^i v^i(t)$.

Now, under the conditions of the theorem, [25, Section 3.4] also gives that the minimum value of the optimal control problem (12) with solutions $\dot{z}^i(t) = B^i v^i(t)$ and $v^i(t) = K^i(t) z^i(t)$ is given by

$$z_0^{i\prime} P^i(0) z_0^i \tag{16}$$

for any initial state $z^i(0) = z_0^i$ where $P^i(0)$ is obtained from the solution to the Riccati differential equation (15). Hence, minimising (12) with an unknown initial state z_0^i is equivalent to minimising (16) over z_0^i. The second theorem result (14) follows and the proof is complete. ■

Theorem 1 establishes that our MP method of inverse differential games (12) reduces to solving the Riccati differential equation (15) followed by solving the quadratic program (14) for each player $i \in \mathbb{N}$. We are free to impose the set constraint $\theta^i \in \Theta^i$ during the solution of the quadratic program (14). We note that $R^i(t)$ will always be positive definite due to the definition of $G^i(t)$, and the nonnegative definite condition on $Q^i(t) - S^i(t)(R^i(t))^{-1} S^{i\prime}(t)$ is testable without knowledge of the unknown cost-functional parameters. Furthermore, as in our BOP method, our MP method allows for the recovery of a single player's cost-functional parameters without needing to recover the cost-functional parameters of the other players (and the true player cost-functional parameters and costate functions are clearly one solution to our proposed MP method (12)).

Finally, it is fruitful to note that an analogous method of discrete-time inverse dynamic games is developed in [19] using necessary conditions for discrete-time open-loop Nash equilibria. In contrast to our proposed MP method (12) involving optimisation over both the unknown player cost-functional parameters and player costate functions, the discrete-time approach of [19] only involves optimisation over the player cost-functional parameters since the player costates can be expressed as linear functions of the parameters via a backwards recursion. Although this discrete-time optimisation problem is simple, the backwards recursive relationship between the player costates and parameters prohibits the use of truncated state and control trajectories (which are handled by our proposed MP method).

We now illustrate and compare the practical performance of our three proposed methods of inverse differential games.

IV. Simulation Example and Study

In this section, we examine the performance of our proposed NOP, BOP, and MP methods of inverse noncooperative differential games in simulations of a two-player noncooperative differential game. To introduce the game, consider the integrator with two (unconstrained) control inputs

$$\dot{x}(t) = u^1(t) + u^2(t), \ x(0) = 20$$

for $t \in [0, 10]$. We assume that the ith player with $i \in \mathbb{N} = \{1, 2\}$ is attempting to drive the value of the integrator to a desired value $c^i \in \mathbb{R}$ by minimising the cost-functional

$$\begin{aligned} &\tilde{J}^i \left(u^1, u^2, \theta^i\right) \\ &\quad = \int_0^{10} \theta_1^i \left(x(t) - c^i\right)^2 + \theta_2^i \left(u^i(t)\right)^2 \, dt. \end{aligned} \tag{17}$$

The inverse noncooperative differential game problem therefore corresponds to finding the parameters θ^i and the value c^i for both players.

The cost-functional (17) is nonlinear in the unknown parameters θ^i and value c^i. However, by expanding the parentheses in (17), defining the additional parameter $\theta_3^i \triangleq -2c^i\theta_1^i$ and disregarding the constant term $\theta_1^i (c^i)^2$, we see that a differential game played with the cost-functionals (17) will share the same open-loop Nash equilibrium solutions as a differential game played with the cost-functionals

$$\begin{aligned} &J^i \left(u^1, u^2, \theta^i\right) \\ &\quad = \int_0^{10} \theta_1^i \left(x(t)\right)^2 + \theta_2^i \left(u^i(t)\right)^2 + \theta_3^i x(t) \, dt \end{aligned} \tag{18}$$

for $i \in \mathbb{N}$. With these linearly parameterised cost functionals, we may now apply our three methods of inverse differential games with the basis functions

$$g^i\left(t, x(t), u^1(t), u^2(t)\right) = \begin{bmatrix} (x(t))^2 & \left(u^i(t)\right)^2 & x(t) \end{bmatrix}'$$

to recover the unknown parameters $\theta^{i*} \in \Theta^i \subset \mathbb{R}^3$, and hence the values c^i for $i \in \mathbb{N}$.

A. Illustrative Example and Computational Effort

To illustrate our proposed methods of inverse differential games, we first simulated open-loop Nash equilibrium state x^* and control u^{i*} trajectories for our two-player differential game with cost functionals (18), and parameters $\theta^{1*} = [1, 2, 5]'$ and $\theta^{2*} = [1, 2, -5]'$. With these parameters, Player 1 attempts to drive the integrator to $c^1 = -2.5$, and Player 2 attempts to drive it to $c^2 = 2.5$. The simulated open-loop Nash equilibrium state and control trajectories are shown in Fig. 1. We used our methods of inverse differential games

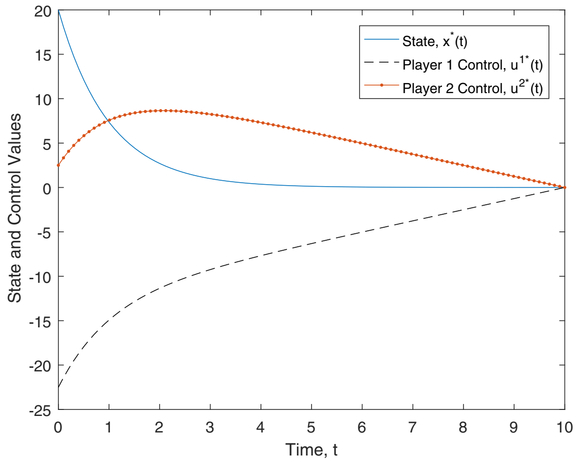

Fig. 1. Illustrative Example: Open-loop Nash equilibrium state and control trajectories. The state is initially $x(0) = 20$ before Player 1 attempts to drive it to $x(t) = -2.5$ and Player 2 attempts to drive it to $x(t) = 2.5$. The open-loop Nash equilibrium solution corresponds to driving $x^*(t) \to 0$ as $t \to 10$.

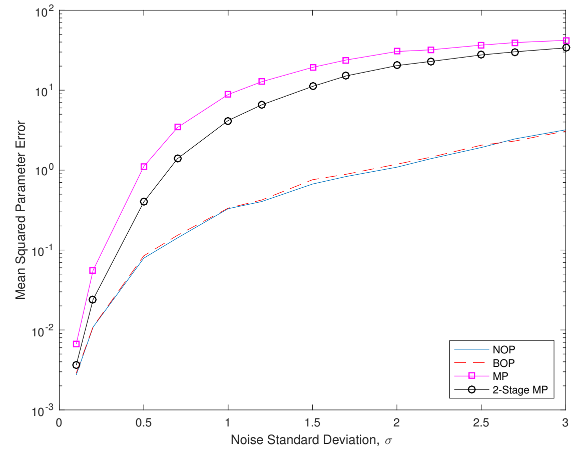

Fig. 2. Simulation Study With Noise: Mean squared error of the calculated value of the player cost-functional parameter θ_2^1 against the standard deviation of additive Gaussian noise corrupting the open-loop Nash equilibrium state and control trajectories. The mean squared error was calculated over 100 independent noise realisations.

to compute the parameters θ^i from these open-loop Nash equilibrium state and control trajectories. The performance of the methods is reported in Table I.

From Table I, we see that all three methods recovered the unknown parameters θ^{1*} and θ^{2*} with errors comparable to the precision of the numeric optimisation and boundary value problem solvers. However, the compute time required by our MP method is significantly less than that of the BOP and NOP methods which require the solution of nested optimal control and differential game problems, respectively. These compute times were calculated on a Dell Latitude E5550 i7 5600U notebook with MATLAB 2016a using the `bvp4c` boundary value problem solver (in all three methods) and the simplex optimisation algorithm of `fminsearch` (for the NOP and BOP methods).

B. Simulation Study With Noise

We now study the performance of our proposed methods when the open-loop Nash equilibrium state and control trajectories are sampled and corrupted by additive Gaussian noise. To conduct our study, we sampled the open-loop Nash equilibrium state and control trajectories simulated in Section IV-A at a rate of 0.1 seconds and added zero-mean Gaussian noise with standard deviation $\sigma \in [0.1, 3]$. We linearly interpolated the noise-corrupted sampled state and control trajectories to enable the application of our continuous-time NOP, BOP and MP methods. Similar to the two-stage approach adopted in [15], we also applied our MP method to noise-corrupted sampled trajectories after they had been preprocessed with cubic smoothing splines – we refer to this method as the 2-Stage MP method. The mean squared error in the calculated value of Player 1's parameter θ_2^1 is plotted in Fig. 2 for all four methods. The mean value was calculated over 200 independent realisations of the noise. Here, we do not report the mean squared errors of the other Player 1 parameters and those of Player 2 since they are all less than those shown in Fig. 2.

The results in Fig. 2 suggest that our BOP and NOP methods are able to recover the parameters of the player cost functionals with the lowest mean squared error over the full range of noise standard deviations considered. In contrast, our proposed MP method performs poorly when the state and control trajectories are corrupted by noise (despite offering the best error performance in the illustrative example of Section IV-A with no noise). Although preprocessing the trajectories with cubic smoothing splines appears to improve the performance of our proposed MP method (in the form of the 2-Stage MP method), our BOP and NOP methods still appear to offer lower parameter estimation error.

V. CONCLUSION

This paper considers the problem of inverse N-player noncooperative differential games in the case of an open-loop information structure. To solve this problem, we propose a direct method involving the nested solution of noncooperative differential games. We then exploit connections between our inverse differential game problem and the problem of inverse optimal control to propose two simplified methods of inverse differential games based on the bilevel optimisation and minimum principle methods of inverse optimal control. The utility of our proposed methods is demonstrated through simulations of a two-player noncooperative differential game.

Future work will investigate the use of these inverse differential game techniques to model behaviours of biological agents such as birds in mid-air collision encounters.

ACKNOWLEDGEMENT

Computational resources and services used in this paper were provided by the HPC and Research Support Group, Queensland University of Technology, Brisbane, Australia.

TABLE I

ILLUSTRATIVE EXAMPLE: PERFORMANCE SUMMARY OF INVERSE DIFFERENTIAL GAME METHODS.

Method	Absolute Parameter Error		No. Differential Games Solved	No. Optimal Control Problems Solved	Compute Time (seconds)
	Mean	Maximum			
NOP	4.08e-5	8.83e-5	388	0	369
BOP	5.15e-4	1.10e-3	0	187	48.0
MP	1.18e-5	2.16e-5	0	0*	8.56

*Excludes the solution of two Riccati equations.

REFERENCES

[1] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*, 2nd ed. New York, NY: Academic Press, 1999, vol. 23.

[2] R. Isaacs, *Differential Games: Mathematical Theory with Application to Warfare and Pursuit Control and Optimisation*. New York: Dover Publications, 1965.

[3] E. Dockner, *Differential games in economics and management science*. Cambridge University Press, 2000.

[4] J. W. Weibull, *Evolutionary game theory*. MIT press, 1997.

[5] J. Engwerda, "Uniqueness conditions for the affine open-loop linear quadratic differential game," *Automatica*, vol. 44, no. 2, pp. 504 – 511, 2008.

[6] V. Kuleshov and O. Schrijvers, "Inverse game theory: Learning utilities in succinct games," in *Int. Conf. on Web and Internet Economics (WINE) 2015*, December 2015, pp. 413–427.

[7] D. Bertsimas, V. Gupta, and I. Paschalidis, "Data-driven estimation in equilibrium using inverse optimization," *Mathematical Programming*, vol. 153, no. 2, pp. 595–633, 2015.

[8] I. C. Konstantakopoulos, L. J. Ratliff, M. Jin, C. J. Spanos, and S. S. Sastry, "Inverse modeling of non-cooperative agents via mixture of utilities," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE, 2016, pp. 6327–6334.

[9] T. Molloy, D. Tsai, J. Ford, and T. Perez, "Discrete-time inverse optimal control with partial-state information: A soft-optimality approach with constrained state estimation," in *Decision and Control (CDC), 2016 IEEE 55th Annual Conference on*, Las Vegas, NV, Dec 2016.

[10] M. Johnson, N. Aghasadeghi, and T. Bretl, "Inverse optimal control for deterministic continuous-time nonlinear systems," in *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, Dec 2013, pp. 2906–2913.

[11] K. Hatz, J. P. Schlöder, and H. G. Bock, "Estimating parameters in optimal control problems," *SIAM Journal on Scientific Computing*, vol. 34, no. 3, pp. A1707–A1728, 2012.

[12] K. Mombaur, A. Truong, and J.-P. Laumond, "From human to humanoid locomotion—an inverse optimal control approach," *Autonomous robots*, vol. 28, no. 3, pp. 369–383, 2010.

[13] L. J. Ratliff, M. Jin, I. C. Konstantakopoulos, C. Spanos, and S. S. Sastry, "Social game for building energy efficiency: Incentive design," in *2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sept 2014, pp. 1011–1018.

[14] A. Keshavarz, Y. Wang, and S. Boyd, "Imputing a convex objective function," in *Intelligent Control (ISIC), 2011 IEEE International Symposium on*. IEEE, 2011, pp. 613–619.

[15] N. Aghasadeghi and T. Bretl, "Inverse optimal control for differentially flat systems with application to locomotion modeling," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, May 2014, pp. 6018–6025.

[16] P. Bajari, C. L. Benkard, and J. Levin, "Estimating dynamic models of imperfect competition," *Econometrica*, vol. 75, no. 5, pp. 1331–1370, 2007.

[17] P. Arcidiacono, P. Bayer, J. R. Blevins, and P. B. Ellickson, "Estimation of Dynamic Discrete Choice Models in Continuous Time with an Application to Retail Competition," *The Review of Economic Studies*, vol. 83, no. 3, pp. 889–931, 2016.

[18] D. Tsai, T. Molloy, and T. Perez, "Inverse two-player zero-sum dynamic games," in *Australian Control Conference (AUCC), 2016*, Newcastle, Australia, Nov 2016.

[19] T. Molloy, J. Ford, and T. Perez, "Inverse noncooperative dynamic games," in *IFAC 2017 World Congress*, Toulouse, France, July 2017.

[20] F. Köpf, J. Inga, S. Rothfuß, M. Flad, and S. Hohmann, "Inverse Reinforcement Learning for Identification in Linear-Quadratic Dynamic Games," in *IFAC 2017 World Congress*, Toulouse, France, July 2017.

[21] S. Rothfuß, J. Inga, F. Köpf, M. Flad, and S. Hohmann, "Inverse Optimal Control for Identification in Non-Cooperative Differential Games," in *IFAC 2017 World Congress*, Toulouse, France, July 2017.

[22] L. J. Ratliff, S. A. Burden, and S. S. Sastry, "On the Characterization of Local Nash Equilibria in Continuous Games," *IEEE Transactions on Automatic Control*, vol. 61, no. 8, pp. 2301–2307, Aug 2016.

[23] J. F.-S. Lin, V. Bonnet, A. M. Panchea, N. Ramdani, G. Venture, and D. Kulić, "Human motion segmentation using cost weights recovered from inverse optimal control," in *Humanoid Robots (Humanoids), 2016 IEEE-RAS 16th International Conference on*. IEEE, 2016, pp. 1107–1113.

[24] S. Albrecht, K. Ramirez-Amaro, F. Ruiz-Ugalde, D. Weikersdorfer, M. Leibold, M. Ulbrich, and M. Beetz, "Imitating human reaching motions using physically inspired optimization principles," in *Humanoid Robots (Humanoids), 2011 11th IEEE-RAS International Conference on*. IEEE, 2011, pp. 602–607.

[25] B. Anderson and J. Moore, *Optimal control: linear quadratic methods*. Englewood Cliffs, NJ: Prentice Hall, 1990.