Solution Sets for Inverse Non-Cooperative Linear-Quadratic Differential Games

Jairo Inga, Esther Bischoff, Timothy L. Molloy, Michael Flad, Sören Hohmann

Abstract—This paper addresses the inverse problem of differential games, where the aim is to compute cost functions which lead to observed Nash equilibrium trajectories. The solution of this problem allows the generation of a model for inferring the intent of several agents interacting with each other. We present a method to find all cost functions which lead to the same Nash equilibrium in an infinite-horizon LQ differential game. The approach relies on a reformulation of the coupled matrix Riccati equations which arise out of necessary and sufficient conditions for Nash equilibria. Furthermore, based on our findings, we present an approach to compute a solution given a set of observed state and control trajectories. Our results highlight properties of feedback Nash equilibria in LQ differential games and provide an efficient approach for the estimation of cost function matrices in such a scenario.

Index Terms—Game Theory, Optimal Control, Cooperative Control, Identification.

I. INTRODUCTION

NON-cooperative differential games, first presented by Isaacs [1], are a mathematical approach to model optimal decision making of several agents interacting with each other. This theory has received considerable attention from the automatic control community. In particular, linear-quadratic (LQ) differential games have been an object of broad research, especially due to the possibility to calculate a linear control law. This class of games has been employed in various applications including driver assistance systems [2], collision avoidance [3], control of mobile robots [4] and of energy grids [5]. Literature has grown around the development of methods to determine the outcome of a game given the objectives of each player, especially the Nash equilibrium which represents the result which potentially arises when all players minimize their individual cost function. However, in the last years, some effort has been spent to develop methods to solve inverse problems in a game theoretical setting, e.g. for human behavior identification during interaction with an automatic controller [6] or identification of biological systems behavior [7]. The inverse problem in differential games consists of determining the objectives of each player, described by cost functions, based on their observed behavior.

The problem of identifying cost functions has been thoroughly analyzed in the one-player-case, also known as inverse optimal control [8], [9]. It has received considerable attention due to its possible application in robotics and to describe biological systems, e.g. human motor behavior [10], [11]

J. Inga, E. Bischoff, M. Flad, and S. Hohmann are with the Institute of Control Systems, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany (e-mail: {jairo.inga,esther.bischoff,flad,soeren.hohmann}@kit.edu)

T. L. Molloy is with the School of Electrical Engineering and Computer Science, Queensland University of Technology (QUT), Brisbane, Australia (e-mail:t.molloy@qut.edu.au)

and insect flying trajectories [12]. Most of the applications of inverse optimal control with real data include the use of direct methods which demand the solution of several optimal control problems to iteratively find the cost function which best explains observed data [11], [13]. Solving an inverse differential game with a similar approach leads to large computation times as shown in [14]. Therefore, more efficient methods are required for estimating cost functions in a game scenario.

Previous work in inverse games includes methods for identifying the objectives or utility functions of agents in static games [15], [16]. Results for dynamic or differential games include a zero-sum two-player dynamic game [17] and a general non zero-sum N-player differential game [14] for an open-loop information structure. Concerning the closed-loop case, [18] and [6] presented results for a two-player inverse LQ dynamic and differential game, respectively. An inverse LQ dynamic game was also considered in [19], yet restricting the players' cost function matrices to only penalize their own controls and to only have diagonal entries.

The ill-posedness of inverse differential games is also an issue which is acknowledged in all mentioned previous work. However, it is generally reduced to the fact that the solution of a differential game (or optimal control problem) is unaffected by a multiplication of the cost functions with any constant $c \in \mathbb{R}^+$. In an inverse differential game, the ill-posedness transcends this fact, meaning that there are potentially several or an infinite number of different cost functions leading to the same Nash equilibrium which differ not only in a scaling factor, a fact which also holds for the one-player case [20].

In this paper, we present results for a general infinite-horizon N-player LQ feedback differential game. Our approach provides a means to find solution sets describing all possible cost functions which lead to the same Nash equilibrium represented by feedback matrices. To achieve this, we exploit the coupled algebraic Riccati equations which arise from the necessary and sufficient conditions for a Nash equilibrium. This allows us to give sufficient conditions for the existence of a solution of the inverse problem. Furthermore, for cases in which no exact feedback matrix is available or in which the sufficient conditions are not fulfilled, we present an approach based on a quadratic program such that cost function parameters can be found which best explain observed state and control trajectories.

The remainder of the paper is structured as follows. In Section II, we formulate the inverse problem of differential games in an LQ setting. Afterwards, we present in Section III a reformulation of the coupled matrix Riccati equations which allows us to give solution sets for inverse LQ differential games. We show simulation results to illustrate our method in Section IV before drawing conclusions in Section V.

II. PROBLEM FORMULATION

Consider a differential game with system dynamics

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \sum_{i=1}^{N} \boldsymbol{B}_i \boldsymbol{u}_i(t) \tag{1a}$$

$$\boldsymbol{x}(t_0) = \boldsymbol{x}_0 \tag{1b}$$

where $(\boldsymbol{A}, \begin{bmatrix} \boldsymbol{B}_1 & \cdots & \boldsymbol{B}_N \end{bmatrix})$ is stabilizable. We denote the system trajectories by $\boldsymbol{x}(t) \in \mathbb{R}^n$ and the control trajectories of all players $i \in \{1, \ldots, N\} =: \mathbb{P} \subset \mathbb{N}^+$ by $\boldsymbol{u}_i(t) \in \mathbb{R}^{p_i}$.

In the course of this paper, we consider players to be selecting their controls with linear feedback laws of the form

$$\boldsymbol{u}_i(t) = -\boldsymbol{K}_i \boldsymbol{x}(t), \quad \forall i \in \mathbb{P}, \tag{2}$$

and we define

$$\boldsymbol{F} \triangleq \boldsymbol{A} - \sum_{i \in \mathbb{P}} \boldsymbol{B}_i \boldsymbol{K}_i \tag{3}$$

as the closed loop system matrix. We further restrict the control laws $\boldsymbol{K}_i$ to belong to the set

$$\mathcal{F} = \left\{ (\boldsymbol{K}_1, ..., \boldsymbol{K}_N) \big| \boldsymbol{F} \text{ is stable} \right\}. \tag{4}$$

The non-emptiness of $\mathcal{F}$ follows from the stabilizability of $(\boldsymbol{A}, \begin{bmatrix} \boldsymbol{B}_1 & \cdots & \boldsymbol{B}_N \end{bmatrix})$ [21].

Within the differential game, each player i aims to minimize an individual quadratic performance index

$$J_i(\boldsymbol{x}_0, \boldsymbol{K}, \boldsymbol{Q}_i, \boldsymbol{R}_{ij}) = \frac{1}{2} \int_0^\infty \boldsymbol{x}^\top \boldsymbol{Q}_i \boldsymbol{x} + \sum_{j=1}^{N} \boldsymbol{u}_j^\top \boldsymbol{R}_{ij} \boldsymbol{u}_j \; \mathrm{d}t, \tag{5}$$

with respect to (1), where $\boldsymbol{Q}_i, \boldsymbol{R}_{ij}$ are symmetric for all $i, j \in \mathbb{P}$ and $\boldsymbol{R}_{ii} > \boldsymbol{0}$ for all $i \in \mathbb{P}$. We write J_i as a function of the N-tuple of feedback laws $\boldsymbol{K} = (\boldsymbol{K}_1, ..., \boldsymbol{K}_N)$ and the initial state $\boldsymbol{x}_0$ since together these generate the state and control trajectories $\boldsymbol{x}(t)$ and $\boldsymbol{u}_i(t)$ via (1) and (2). The restriction to $\mathcal{F}$ ensures finite cost function values.

We consider a non-cooperative differential game and assume that the players' strategies result in a feedback Nash equilibrium defined as follows:

Definition 2.1 (Definition 1 from [21]): An N-tuple $\boldsymbol{K}^* = (\boldsymbol{K}_1^*, ..., \boldsymbol{K}_N^*) \in \mathcal{F}$ is called a *feedback Nash equilibrium* if

$$J_i(\boldsymbol{x}_0, \boldsymbol{K}^*, \boldsymbol{Q}_i, \boldsymbol{R}_{ij}) \leq J_i(\boldsymbol{x}_0, \boldsymbol{K}_{\neg i}^*(\boldsymbol{\alpha}), \boldsymbol{Q}_i, \boldsymbol{R}_{ij}), \quad (6)$$

holds for all $i \in \mathbb{P}$, all $\boldsymbol{x}_0 \in \mathbb{R}^n$, and all $\boldsymbol{\alpha}$ such that $\boldsymbol{K}_{\neg i}^*(\boldsymbol{\alpha}) \in \mathcal{F}$, where $\boldsymbol{K}_{\neg i}^*(\boldsymbol{\alpha}) = (\boldsymbol{K}_1^*, ..., \boldsymbol{K}_{i-1}^*, \boldsymbol{\alpha}, \boldsymbol{K}_{i+1}^*, ..., \boldsymbol{K}_N^*)$.

We further introduce the following definition before formalizing the inverse LQ differential game problem.

Definition 2.2: The *canonical parameter set* of the LQ differential game is the set $\boldsymbol{\Theta}$ which contains all possible cost function parameters of (5), i.e. all possible matrices $\boldsymbol{Q}_i$ and $\boldsymbol{R}_{ij}, \forall i \in \mathbb{P}$, which lead to the Nash equilibrium given by $\boldsymbol{K}^* = (\boldsymbol{K}_1^*, ..., \boldsymbol{K}_N^*)$.

Problem 1: Assume that system dynamics matrices $\boldsymbol{A}$, $\boldsymbol{B}_i, \forall i \in \mathbb{P}$ are known and that the cost function structure is given by (5). Furthermore, assume that the Nash equilibrium feedback matrices $(\boldsymbol{K}_1^*, ..., \boldsymbol{K}_N^*)$ are known. Determine the canonical parameter set $\boldsymbol{\Theta}$.

Note 1: By solving Problem 1 we can also solve the related problem of finding $\boldsymbol{\Theta}$, if instead of $(\boldsymbol{K}_1^*, ..., \boldsymbol{K}_N^*)$, a corresponding set of trajectories $\left\{ \boldsymbol{x}^*(t), \{\boldsymbol{u}_i^*(t)\}_{i \in \mathbb{P}} \right\}$ is given. We will show further details in Section III-F.

III. INVERSE LINEAR-QUADRATIC DIFFERENTIAL GAME

In this section we first present general solution sets for inverse LQ differential games based on a reformulation of the Riccati equations. Afterwards, we consider the case in which only measurements of the state and control trajectories are available, i.e. $\boldsymbol{K}^*$ has to be estimated beforehand.

A. Coupled Algebraic Riccati Equations

We start by introducing the following theorem.

Theorem 1 (Theorem 4 of [21]): Let there exist an N-tuple of symmetric matrices $\boldsymbol{P}_i$, $i \in \mathbb{P}$ satisfying the N matrix algebraic Riccati equations

$$\boldsymbol{P}_i \boldsymbol{F} + \boldsymbol{F}^\top \boldsymbol{P}_i + \sum_{j \in \mathbb{P}} \boldsymbol{P}_j \boldsymbol{B}_j \boldsymbol{R}_{jj}^{-1} \boldsymbol{R}_{ij} \boldsymbol{R}_{jj}^{-1} \boldsymbol{B}_j^\top \boldsymbol{P}_j + \boldsymbol{Q}_i = \boldsymbol{0} \tag{7}$$

such that $\boldsymbol{F}$ is stabilized. Furthermore, let $\boldsymbol{K}_i^*$ be defined as

$$\boldsymbol{K}_i^* = \boldsymbol{R}_{ii}^{-1} \boldsymbol{B}_i^\top \boldsymbol{P}_i. \tag{8}$$

Then, $\boldsymbol{K}^* = (\boldsymbol{K}_1^*, ..., \boldsymbol{K}_N^*)$ is a feedback Nash equilibrium as in Definition 2.1 and $J_i(\boldsymbol{x}_0, \boldsymbol{K}^*, \boldsymbol{Q}_i, \boldsymbol{R}_{ij}) = \boldsymbol{x}_0^\top \boldsymbol{P}_i \boldsymbol{x}_0$. Conversely, if $\boldsymbol{K}^*$ is a feedback Nash equilibrium then the set of algebraic Riccati equations (7) has a stabilizing solution. ∎

Proof: See [21]. ∎

Theorem 1 represents a necessary and sufficient condition for feedback Nash equilibria. Hence, if the feedback matrices are given, the cost function parameters must fulfill (7). We exploit this fact in order to develop a method to solve the inverse LQ differential game. The approach is based on a reformulation of (7) which is presented in the following.

B. Reformulation of the Algebraic Riccati Equations

Before we present the reformulations, we first define a Kronecker sum [22] as

$$\boldsymbol{X} \oplus \boldsymbol{Y} = (\boldsymbol{X} \otimes \boldsymbol{I}_q) + (\boldsymbol{I}_r \otimes \boldsymbol{Y}), \tag{9}$$

for squared matrices $\boldsymbol{X} \in \mathbb{R}^{r \times r}$ and $\boldsymbol{Y} \in \mathbb{R}^{q \times q}$, where $\boldsymbol{I}_q$ denotes a q-dimensional identity matrix and $\otimes$ is the Kronecker product. In order to develop our reformulation of (7), we require the following result.

Lemma 1: Define $\boldsymbol{F}_\oplus \triangleq \boldsymbol{F}^\top \oplus \boldsymbol{F}^\top$ (cf. (3)). The inverse $\boldsymbol{F}_\oplus^{-1}$ exists.

Proof: $\boldsymbol{F}_\oplus^{-1}$ exists if all eigenvalues $\lambda_i \in \sigma(\boldsymbol{F}_\oplus)$, $i \in \{1, ..., n^2\}$ are different from zero. By using Theorem 4.8 of [23], we discern that $\lambda_i = \mu_j + \mu_k$, where $\mu_j, \mu_k \in \sigma(\boldsymbol{F})$, for $j, k \in \{1, ..., n\}$ such that i is associated to a particular combination of j and k, i.e. $j = \lceil \frac{i}{n} \rceil$ and $k = i - n(j-1)$. Since we only consider feedback matrices $\boldsymbol{K}_i, \forall i \in \mathbb{P}$ from the set (4), $\boldsymbol{F}$ is a stable matrix and thus $\lambda_i < 0, \forall i \in \{1, ..., n^2\}$. The lemma assertion follows. ∎

Unless otherwise stated, the following calculations are with respect to a particular player $i \in \mathbb{P}$. With the results of Lemma 1 we define the matrices

$$\boldsymbol{S}_i \triangleq (\boldsymbol{I}_n \otimes \boldsymbol{B}_i^\top)\boldsymbol{F}_\oplus^{-1} \in \mathbb{R}^{np_i \times n^2} \tag{10}$$

and

$$\boldsymbol{K}_i^\otimes \triangleq \boldsymbol{K}_i^\top \otimes \boldsymbol{K}_i^\top \in \mathbb{R}^{n^2 \times p_i^2}. \tag{11}$$

We write $\boldsymbol{K}_i^*$ as $\boldsymbol{K}_i$ in (11) and in the following Lemma for brevity.

Lemma 2: Let the parameter $\boldsymbol{\theta}_i \in \mathbb{R}^{L \times 1}$ denote the vectorized matrices of the cost function (5) as

$$\boldsymbol{\theta}_i = \left[\text{vec}(\boldsymbol{Q}_i)^\top \, \text{vec}(\boldsymbol{R}_{i1})^\top \cdots \text{vec}(\boldsymbol{R}_{ii})^\top \cdots \text{vec}(\boldsymbol{R}_{iN})^\top\right]^\top, \tag{12}$$

where $\text{vec}(\boldsymbol{X})$ represents a column vectorization of a matrix $\boldsymbol{X}$. Then, the matrices $\boldsymbol{Q}_i, \boldsymbol{R}_{ij}$ corresponding to $\boldsymbol{\theta}_i$ satisfy (7) if $\boldsymbol{\theta}_i$ fulfills

$$\boldsymbol{M}_i \boldsymbol{\theta}_i = \boldsymbol{0} \tag{13}$$

where $\boldsymbol{M}_i \in \mathbb{R}^{np_i \times L}$ is given by

$$\boldsymbol{M}_i \triangleq [\boldsymbol{S}_i \quad \boldsymbol{S}_i \boldsymbol{K}_1^\otimes \quad \cdots \quad \boldsymbol{S}_i \boldsymbol{K}_{i-1}^\otimes \ldots$$
$$(\boldsymbol{S}_i \boldsymbol{K}_i^\otimes + \boldsymbol{K}_i \otimes \boldsymbol{I}_p) \quad \boldsymbol{S}_i \boldsymbol{K}_{i+1}^\otimes \quad \cdots \quad \boldsymbol{S}_i \boldsymbol{K}_N^\otimes]. \tag{14}$$

Proof: We rewrite (7) as

$$\boldsymbol{0} = \text{vec}(\boldsymbol{P}_i \boldsymbol{F}) + \text{vec}(\boldsymbol{F}^\top \boldsymbol{P}_i) + \ldots$$
$$\sum_{j \in \mathbb{P}} \text{vec}(\boldsymbol{P}_j \boldsymbol{B}_j \boldsymbol{R}_{jj}^{-1} \boldsymbol{R}_{ij} \boldsymbol{R}_{jj}^{-1} \boldsymbol{B}_j^\top \boldsymbol{P}_j) + \text{vec}(\boldsymbol{Q}_i)$$
$$\boldsymbol{0} = \left[(\boldsymbol{F}^\top \otimes \boldsymbol{I}_n) + (\boldsymbol{I}_n \otimes \boldsymbol{F}^\top)\right] \text{vec}(\boldsymbol{P}_i) + \ldots$$
$$\sum_{j \in \mathbb{P}} (\boldsymbol{K}_j^\top \otimes \boldsymbol{K}_j^\top) \text{vec}(\boldsymbol{R}_{ij}) + \text{vec}(\boldsymbol{Q}_i)$$

and thus

$$\text{vec}(\boldsymbol{P}_i) = -\boldsymbol{F}_\oplus^{-1}\text{vec}(\boldsymbol{Q}_i) - \sum_{j \in \mathbb{P}} \boldsymbol{F}_\oplus^{-1} \boldsymbol{K}_j^\otimes \text{vec}(\boldsymbol{R}_{ij}). \tag{15}$$

The first equality follows from vectorizing (7), while for the second equality we used (8) and applied the equivalence

$$\text{vec}(\boldsymbol{X}\boldsymbol{Y}\boldsymbol{Z}) = (\boldsymbol{Z}^\top \otimes \boldsymbol{X}) \text{vec}(\boldsymbol{Y}) \tag{16}$$

which holds for any matrices $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$ with suitable dimensions [22]. The third equality (15) follows with the results of Lemma 1 and the definitions given in (11) and (9). Now we rewrite (8) as

$$(\boldsymbol{I}_n \otimes \boldsymbol{B}_i^\top)^{-1} (\boldsymbol{K}_i^\top \otimes \boldsymbol{I}_p) \text{vec}(\boldsymbol{R}_{ii}) = \text{vec}(\boldsymbol{P}_i) \tag{17}$$

using (16). Inserting (17) in (15) results in

$$\boldsymbol{S}_i \text{vec}(\boldsymbol{Q}_i) + (\boldsymbol{K}_i^\top \otimes \boldsymbol{I}_p) \text{vec}(\boldsymbol{R}_{ii}) + \sum_{j \in \mathbb{P}} \boldsymbol{S}_i \boldsymbol{K}_j^\otimes \text{vec}(\boldsymbol{R}_{ij}) = \boldsymbol{0} \tag{18}$$

and thus (13) follows immediately with (14) and (12). ∎

C. Solution Sets for Inverse LQ Differential Games

The parameters $\boldsymbol{\theta}_i$ for which (13) holds are valid solutions of (7) for a given $\boldsymbol{K}_i^*$. Note that the feedback matrices $(\boldsymbol{K}_1^*, \ldots, \boldsymbol{K}_N^*)$ completely characterize the Nash equilibrium trajectories $\boldsymbol{x}^*(t)$ and $\boldsymbol{u}_i^*(t), i \in \mathbb{P}$. This follows from (1) fulfilling all conditions for admitting a unique solution for any N-tuple of continuous controls (2) [24]. Thus, the parameters $\boldsymbol{\theta}_i$ are associated to a Nash equilibrium represented by either the feedback matrices or the state and control trajectories.

Remark 1: If different parameter sets exist which represent the same Nash equilibrium, then by equation (8) we discern that different matrices $\boldsymbol{P}_i, i \in \mathbb{P}$ exist which correspond to the parameters $\boldsymbol{\theta}_i$. Both $\boldsymbol{\theta}_i$ (by the results of Lemma 2) and $\boldsymbol{P}_i$ fulfill the N equations (7) and are found by means of the set of feedback matrices $\boldsymbol{K}_i^*, i \in \mathbb{P}$, thus representing the same Nash equilibrium trajectories.

The matrix Riccati equations (7) have multiple solutions which potentially represent different Nash equilibria [25]. However, note that we are only interested in all parameters $\boldsymbol{\theta}_i$ which represent a specific Nash equilibrium. We now present the following theorem as our main result.

Theorem 2: Let a differential game be given by (1) and (5). With assumed knowledge of the control laws $\boldsymbol{K}_i^*$, the canonical parameter set of the inverse LQ differential game is given by

$$\boldsymbol{\Theta} = \bigcup_{i \in \mathbb{P}} \ker(\boldsymbol{M}_i), \tag{19}$$

with convex boundaries such that $\boldsymbol{R}_{ii} > \boldsymbol{0}, \forall i \in \mathbb{P}$.

Proof: By inspecting (13) from Lemma 2 we can recognize that all parameters which satisfy the matrix Riccati equations lie within the kernel of $\boldsymbol{M}_i$. Therefore, all possible cost function parameters of player i which lead to the known Nash equilibrium are given by $\text{span}(\boldsymbol{v}_i^{(1)}, \ldots, \boldsymbol{v}_i^{(d_i)})$, where d_i represents the dimension of the kernel of $\boldsymbol{M}_i$ with basis vectors $\boldsymbol{v}_i$. Furthermore, $\boldsymbol{M}_i$ depends on $\boldsymbol{K}^*$. The set including the cost function parameters of all players corresponding to the Nash equilibrium represented by $\boldsymbol{K}^*$ or the trajectories $\boldsymbol{x}^*(t)$ and $\boldsymbol{u}_i^*(t), i \in \mathbb{P}$ is thus given by (19). ∎

Corollary 1: The trajectories constituting a Nash equilibrium under N cost functions $J_i(\boldsymbol{\theta}_i^*), i \in \mathbb{P}$ will constitute the same Nash equilibrium for $J_i(\boldsymbol{\theta}_i)$ with $\boldsymbol{\theta}_i = \kappa_i \boldsymbol{\theta}_i^*, \forall \kappa_i > 0$.

Proof: This can be easily be seen from $\boldsymbol{M}_i \kappa_i \boldsymbol{\theta}_i^* = \kappa_i \boldsymbol{M}_i \boldsymbol{\theta}_i^* = \boldsymbol{0}$. ∎

Note that the results of Lemma 2 as well as Theorem 2 are derived with respect to the parameter definition in (12) which considers the most general case where there are no restrictions on the cost function matrices. We now present properties of inverse LQ differential games based on the possible structures of the cost function matrices.

D. Sufficient Condition for Inverse Linear-Quadratic Differential Game Solutions

The kernel of $\boldsymbol{M}_i$ characterizes all non-trivial solutions of (13). Its dimension will depend on the number of linearly independent equations generated by the np_i rows of $\boldsymbol{M}_i$ compared to the number of unknown parameters L. Since $\text{rank}(\boldsymbol{M}_i) \leq \min(L, np_i)$, the number of players, states and

controls of each player as well as the assumed properties of the cost function matrices are important for evaluating the existence of inverse differential game solutions. A sufficient condition for the existence of a solution is given by $np_i < L$ since this implies $\text{rank}(\boldsymbol{M}_i) < L$, leading to $\dim(\ker(\boldsymbol{M}_i)) > 0$. In the following, we discuss two main cases of the structure of the cost function matrices:

1) Symmetric Cost Function Matrices: If we assume symmetry of all cost function matrices, then $L = 0.5(n^2 + n + \sum_{i\in\mathbb{P}}(p_i^2 + p_i))$. Since $np_i \leq 0.5(n^2 + p_i^2) < 0.5\left(n(n+1) + \sum_{i\in\mathbb{P}} p_i(p_i+1)\right) = L$ for any choice of $n, p_i, N \in \mathbb{N}_{>1}^+$, $\dim(\ker(\boldsymbol{M}_i)) > 0$ holds. The sufficient condition is fulfilled and an infinite number of parameters exists that solve the inverse differential game problem.

2) Diagonal Cost Function Matrices: Only in the case of diagonal matrices, where $L = n + \sum_{i\in\mathbb{P}} p_i$, combinations of n, p_i, N exist such that $np_i \geq L$. Here we note that, if $\text{rank}(\boldsymbol{M}_i) = L-1$, a unique algebraic solution for any player's parameters may be found by setting $\boldsymbol{\theta}_{i,(j)} = 1, j \in \{1, ..., L\}$ and proceeding analogously to Proposition 1 of [8]. This is possible for $N = 1$ e.g. if $n = 1$ and $p_1 = 1$. Note that $np_i \geq L$ is possible for diagonal matrices, which means that the sufficient condition for the existence of solutions for the inverse differential game may not be fulfilled.

We now introduce an approach to find a solution of the inverse differential game problem regardless of the presented properties.

E. Quadratic Programming Formulation for the Inverse Linear-Quadratic Differential Game

We define, analogously to the inverse optimal control literature (e.g. [26], [27]), the residual

$$\boldsymbol{r}_i = \boldsymbol{M}_i \boldsymbol{\theta}_i, \tag{20}$$

which represents the extent to which the necessary and sufficient conditions for Nash equilibria are violated. This may happen e.g. if $\boldsymbol{K}_i^*$ is not exactly known or if the parameters do not represent a Nash equilibrium for given $\boldsymbol{K}_i^*$. We aim to minimize this residual, leading to the optimization problem

$$\min_{\boldsymbol{\theta}_i} ||\boldsymbol{r}_i||_2^2 = \frac{1}{2}\boldsymbol{\theta}_i^\top \boldsymbol{H}_i \boldsymbol{\theta}_i,$$
$$\text{s.t.} \tag{21}$$
$$\boldsymbol{I}_L \boldsymbol{\theta}_i > \boldsymbol{0}$$
$$\boldsymbol{R}_{ii} > \boldsymbol{0}$$

where $\boldsymbol{H}_i = 2(\boldsymbol{M}_i^\top \boldsymbol{M}_i)$.

Proposition 1: A feasible solution, i.e. a local minimizer of the quadratic program (21) is guaranteed to exist. Furthermore, if $\boldsymbol{M}_i$ has full rank, i.e. $\text{rank}(\boldsymbol{M}_i) = \min(np_i, L)$, then (21) has a unique global solution.

Proof: It is clear that both the constraint set defined by $\boldsymbol{I}_L$ and $\boldsymbol{R}_{ii} > \boldsymbol{0}$ represent convex boundaries. The quadratic objective function is at least convex since $\boldsymbol{M}_i^\top \boldsymbol{M}_i \geq \boldsymbol{0}$ always holds. Hence, a solution of the quadratic program is guaranteed to exist. If $\boldsymbol{M}_i$ has full rank, then $\boldsymbol{H}_i > \boldsymbol{0}$ and the objective function is strictly convex, which means that a unique global minimizer of (21) exists. ∎

Remark 2: The constraint $\boldsymbol{I}_L \boldsymbol{\theta}_i > \boldsymbol{0}$ in (21) is introduced in order to avoid trivial solutions. Literature in inverse optimal control and inverse games often introduce the constraint $\boldsymbol{\theta}_{i,(j)} = 1$ for any $j \in \{1, ..., L\}$ (see [14] and references therein). The existence of the solution of (21) with this constraint under the same conditions as in Proposition 1 can easily be proved as well. Also note that, in case of diagonal cost function matrices, $\boldsymbol{I}_L \boldsymbol{\theta}_i > \boldsymbol{0}$ ensures $\boldsymbol{R}_{ii} > \boldsymbol{0}$.

Remark 3: Symmetry of the cost function matrices is considered by introducing corresponding parameter constraints in (21), which we omitted in favor of better readability. An alternative is to modify the matrix $\boldsymbol{M}_i$ according to a redefinition of the parameter vector of player i as $\boldsymbol{\theta}_i = \left[\text{vech}(\boldsymbol{Q}_i)^\top \; \text{vech}(\boldsymbol{R}_{i1})^\top \; \cdots \; \text{vech}(\boldsymbol{R}_{iN})^\top\right]^\top$, where $\text{vech}(\cdot)$ denotes the half-vectorization operator. This vectorizes only the lower triangular part of a matrix. A similar approach can be employed for diagonal matrices.

F. Identification of the Control Law

The optimization problem (21) always yields a solution which is associated with a given Nash equilibrium represented by $\boldsymbol{K}^*$. However, in practice it is often necessary to estimate the control laws out of observed Nash equilibrium control and state trajectories. Identification of $\boldsymbol{K}^*$ is performed e.g. in [10] and [12] for a one-player differential game. For the general inverse differential game, we apply a least-squares identification based on (2). For this, we introduce a finite sequence of sampling times $\mathcal{K}_i \triangleq \{t_k \in [0, T] : 1 \leq k \leq K_i \wedge 0 < t_1 < ... < t_{K_i} \leq T\}$ for each player $i \in \mathbb{P}$, where $[0, T]$ is the time interval of $\boldsymbol{x}(t)$ and $\boldsymbol{u}_i(t)$. Let the value of the state and control trajectories at t_k be denoted by $\boldsymbol{x}^{[k]}$ and $\boldsymbol{u}_i^{[k]}$, respectively. Then we estimate the feedback matrix by means of

$$\hat{\boldsymbol{K}}_i = \arg\min_{\boldsymbol{K}_i} \sum_{k=1}^{K_i} ||\boldsymbol{K}_i \boldsymbol{x}^{[k]} + \boldsymbol{u}_i^{[k]}||_2^2. \tag{22}$$

IV. Simulation Examples

In this section, we present two LQ differential game examples to illustrate the presented ideas and methods. We first provide a two-player game example to illustrate Theorem 2. Afterwards, we examine with a three-player game how the proposed method finds a valid solution of the problem based on noisy measurements of the Nash equilibrium trajectories.

A. 2-Player Scenario

We consider a scenario consisting of two players controlling a double-integrator system given by

$$\dot{\boldsymbol{x}}(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \boldsymbol{x}(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_1(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_2(t). \tag{23}$$

The ground truth cost functions of the two players are $\boldsymbol{Q}_1 = \text{diag}(1, 2)$ and $\boldsymbol{Q}_2 = \text{diag}(1, 0.7)$ as well as $\boldsymbol{R}_{11} = 1$, $\boldsymbol{R}_{12} = \boldsymbol{R}_{21} = 0$ and $\boldsymbol{R}_{22} = 1$. Therefore, the parameter vector is given by $\boldsymbol{\theta}_i = \begin{bmatrix} q_i^{11} & q_i^{22} & R_{ii} \end{bmatrix}$. We solved the LQ differential game by calculating the solution of the corresponding Riccati differential equations and extracting the converged

value of $\boldsymbol{P}_i$. We note that the resulting $\boldsymbol{K}^*$ represents a Nash equilibrium since the calculated $\boldsymbol{P}_i$ satisfies (7) for all players and the closed loop stability of (23) was confirmed (Theorem 1). In this case, the Nash equilibrium is given by $(\boldsymbol{K}_1^*, \boldsymbol{K}_2^*) = ([0.5773 \quad 1.2827], [0.5774 \quad 0.5882])$.

The kernels of the matrices $\hat{\boldsymbol{M}}_i \in \mathbb{R}^{2 \times 3}$ are given by

$$\boldsymbol{v}_1^{(1)} = [v_{1,(j)}^{(1)}]_{j=1,2,3} = [0.4083 \quad 0.8165 \quad 0.4083]^\top \quad (24)$$

$$\boldsymbol{v}_2^{(1)} = [v_{2,(j)}^{(1)}]_{j=1,2,3} = [0.6337 \quad 0.4437 \quad 0.6337]^\top \quad (25)$$

which result in the canonical parameter set

$$\boldsymbol{\Theta} = \{\nu_i \hat{\boldsymbol{Q}}_i, \tau_i \hat{R}_{ii}\}_{i=1,2}, \quad \nu_i \in \mathbb{R}, \tau_i \in \mathbb{R}^+, \quad (26)$$

where $\hat{\boldsymbol{Q}}_i = \mathrm{diag}(v_{i,1}^{(1)}, v_{i,2}^{(1)})$ and $\hat{R}_{ii} = v_{i,3}^{(1)}$. This means that the cost function parameters are unique up to a constant parameter. In particular, $\mu_1 = 2.4494$ and $\mu_2 = 1.5779$ lead to the defined ground truth parameters.

B. 3-Player Scenario

We now assume a more realistic scenario where the feedback matrices are not given and have to be estimated beforehand from observed data. Let us define 3 players which simultaneously control a dynamic system described by

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \sum_{i=1}^{3} \boldsymbol{B}_i \boldsymbol{u}_i(t) \quad (27)$$

with

$$\boldsymbol{A} = \begin{bmatrix} -8 & -6 & 1 & 0 \\ 1 & 0 & 2 & 1 \\ 0 & -2 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix}, \quad \boldsymbol{B}_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix},$$

$$\boldsymbol{B}_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{B}_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

The ground truth parameters of the three players were set to $\boldsymbol{Q}_1 = \mathrm{diag}(1, 0.4, 2, 1)$, $\boldsymbol{Q}_2 = \boldsymbol{Q}_3 = \mathrm{diag}(1, 1, 1, 1)$ and $\boldsymbol{R}_{11} = \mathrm{diag}(1, 1)$, $\boldsymbol{R}_{12} = \mathrm{diag}(0.5, 0.5)$, $\boldsymbol{R}_{13} = \mathrm{diag}(0.2, 0.2)$ and $\boldsymbol{R}_{21} = \boldsymbol{R}_{22} = \boldsymbol{R}_{23} = \mathrm{diag}(1, 1)$ and $\boldsymbol{R}_{31} = \mathrm{diag}(0.2, 0.5)$, $\boldsymbol{R}_{32} = \mathrm{diag}(1, 1)$, $\boldsymbol{R}_{33} = \mathrm{diag}(1, 2)$. The properties of this differential game lead to $\dim(\ker(\boldsymbol{M}_i)) = 2, i \in \{1, 2, 3\}$. We determined the Nash equilibrium in an analogous way as in the last example and calculated the trajectories $\boldsymbol{x}^*(t)$ and $\boldsymbol{u}_i^*(t)$. The Nash character of the equilibrium was also confirmed similarly. We added white Gaussian noise to all trajectories and obtained $\tilde{x}_l(t) = x_l^*(t) + \epsilon_l, \forall l \in \{1, ..., n\}$ and $\tilde{u}_{i,m}(t) = u_{i,m}^*(t) + \epsilon_{i,m}, \forall m \in \{1, ..., p\}, \forall i \in \mathbb{P}$. Gaussian noise was chosen in such a way that all signals have a particular signal-to-noise ratio (SNR). We used different SNR levels for trajectory generation.

In order to analyze the performance of the approach, we define the error measure for state trajectories

$$e^{\boldsymbol{x}} = \max \{e_1^{\boldsymbol{x}}, \dots, e_n^{\boldsymbol{x}}\}, \quad (28)$$

$$e_j^{\boldsymbol{x}} = \left\| \frac{\hat{x}_j(t)}{\|\tilde{x}_j(t)\|_{\max}} - \frac{\tilde{x}_j(t)}{\|\tilde{x}_j(t)\|_{\max}} \right\|_{\max}, \forall j \in \{1, \dots, n\} \quad (29)$$

TABLE I
RESULTS USING TRAJECTORIES WITH DIFFERENT SNR LEVELS

SNR in dB	e_x	e_u	$\|\boldsymbol{M}_1\hat{\boldsymbol{\theta}}_1\|$	$\|\boldsymbol{M}_2\hat{\boldsymbol{\theta}}_2\|$	$\|\boldsymbol{M}_3\hat{\boldsymbol{\theta}}_3\|$
5	0.259	0.535	0.190	0.382	0.280
10	0.152	0.424	0.176	0.291	0.229
15	0.102	0.261	0.152	0.238	0.160
20	0.052	0.204	0.102	0.177	0.178
25	0.024	0.088	0.049	0.079	0.060
30	0.014	0.028	0.018	0.034	0.041

and for the control trajectories

$$e^{\boldsymbol{u}} = \max \{e^{\boldsymbol{u}_1}, \dots, e^{\boldsymbol{u}_N}\} \quad (30)$$

$$e^{\boldsymbol{u}_i} = \max \{e_1^{\boldsymbol{u}_i}, \dots, e_{p_i}^{\boldsymbol{u}_i}\}, \quad \forall i \in \{1, \dots, N\}, \quad (31)$$

$$e_l^{\boldsymbol{u}_i} = \left\| \frac{\hat{u}_{i,(l)}(t)}{\|\tilde{u}_{i,(l)}(t)\|_{\max}} - \frac{\tilde{u}_{i,(l)}(t)}{\|\tilde{u}_{i,(l)}(t)\|_{\max}} \right\|_{\max}, \forall l \in \{1, \dots, p_i\}. \quad (32)$$

We further consider $\|\boldsymbol{M}_i\hat{\boldsymbol{\theta}}_i\|$ in order to get a measure of the Nash character of the parameters with respect to the ground truth feedback matrices.

We determined the cost function parameters by first estimating the feedback matrices using (22) and then solving (21). The results are given in Table I. The higher the SNR level, the better the parameters represent a Nash equilibrium and the better the model trajectories approximate the given state and control trajectories. However, even for a low SNR of 5 dB, the results are still reliable. An example of the trajectory approximation is given in Fig. 1 and Fig. 2 for an SNR of 5 dB. We note that when noisy measurements are considered, the estimated feedback matrices $\hat{\boldsymbol{K}}_i$ deviate slightly from the ground truth matrices $\boldsymbol{K}_i^*$. Therefore, the identified parameters lie within the span of $\ker(\hat{\boldsymbol{M}}_i)$, where $\hat{\boldsymbol{M}}_i$ is calculated by means of (14) with the estimated feedback matrices $\hat{\boldsymbol{K}}_i$. We found out that, for all scenarios with different SNR values, $\mathrm{span}(\boldsymbol{v}_i^{(1)}, \boldsymbol{v}_i^{(2)}) \neq \mathrm{span}(\hat{\boldsymbol{v}}_i^{(1)}, \hat{\boldsymbol{v}}_i^{(2)}), \forall i \in \mathbb{P}$. In spite of this fact, the approach is able to find parameters which describe the observed trajectories adequately.

V. CONCLUSION

We presented an approach and sufficient conditions for calculating solution sets of inverse infinite-horizon LQ games based on a reformulation of the coupled Riccati equations. The set includes all possible cost function matrices corresponding to a stabilizing Nash equilibrium. We also derived a quadratic program such that a solution which fulfills the coupled Riccati equations at least approximately can always be found. In this way, cost function parameters are found which best explain measured trajectories.

REFERENCES

[1] R. Isaacs, *Differential games: a mathematical theory with applications to warfare and pursuit, control and optimization*. Courier Corporation, 1999.
[2] M. Flad, L. Fröhlich, and S. Hohmann, "Cooperative Shared Control Driver Assistance Systems Based on Motion Primitives and Differential Games," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 5, pp. 711–722, Oct. 2017.
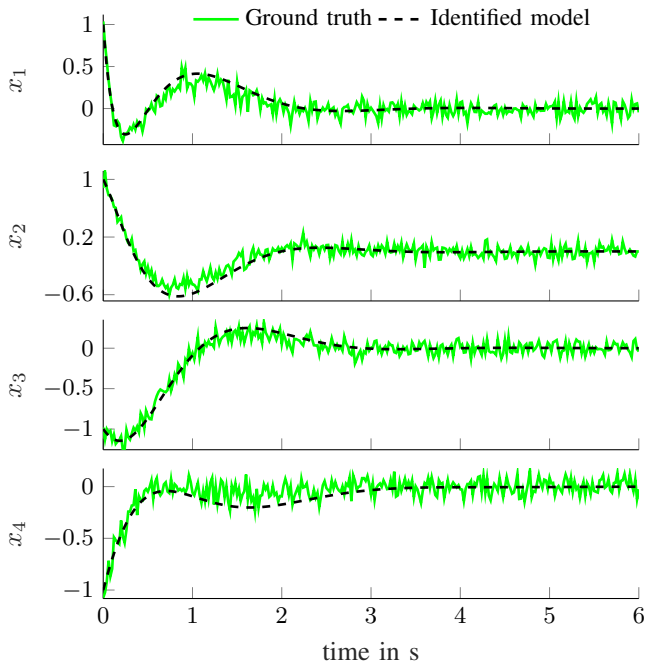
Fig. 1. State trajectories with SNR = 5dB used for identification and state trajectories generated by the identified model
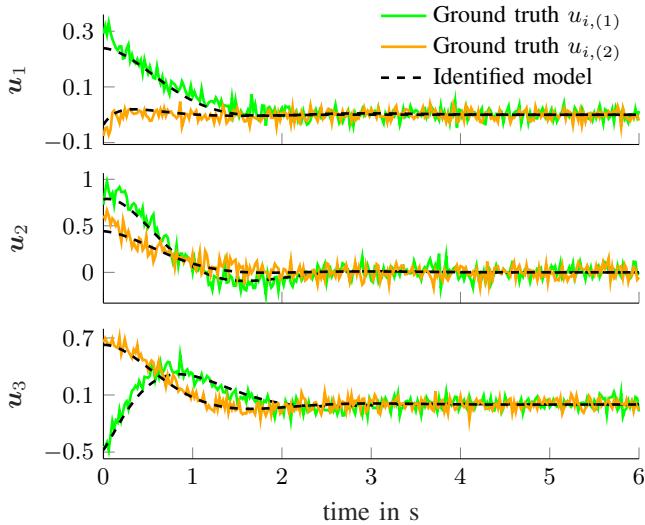

Fig. 2. Control trajectories with SNR = 5dB used for identification and control trajectories generated by the identified model

[3] T. Mylvaganam, M. Sassano, and A. Astolfi, "A Differential Game Approach to Multi-agent Collision Avoidance," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 4229–4235, Aug. 2017.

[4] D. Gu, "A Differential Game Approach to Formation Control," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 1, pp. 85–93, Jan. 2008.

[5] S. Zazo, S. V. Macua, M. Sánchez-Fernández, and J. Zazo, "Dynamic Potential Games With Constraints: Fundamentals and Applications in Communications," *IEEE Transactions on Signal Processing*, vol. 64, no. 14, pp. 3806–3821, Jul. 2016.

[6] S. Rothfuß, J. Inga, F. Köpf, M. Flad, and S. Hohmann, "Inverse Optimal Control for Identification in Non-Cooperative Differential Games," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 14 909–14 915, 2017.

[7] T. L. Molloy, G. S. Garden, T. Perez, I. Schiffner, D. Karmaker, and M. V. Srinivasan, "An Inverse Differential Game Approach to Modelling Bird Mid-Air Collision Avoidance Behaviours," *IFAC-PapersOnLine*, vol. 51, no. 15, Jan. 2018.

[8] M. Menner and M. N. Zeilinger, "Convex formulations and algebraic solutions for linear quadratic inverse optimal control problems," in *2018 European Control Conference (ECC)*, June 2018, pp. 2107–2112.

[9] F. Jean and S. Maslovskaya, "Inverse optimal control problem: the linear-quadratic case," in *2018 IEEE Conference on Decision and Control (CDC)*. FL, USA: IEEE, Dec. 2018, pp. 888–893.

[10] M. C. Priess, R. Conway, Jongeun Choi, J. M. Popovich, and C. Radcliffe, "Solutions to the Inverse LQR Problem With Application to Biological Systems Analysis," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 2, pp. 770–777, Mar. 2015.

[11] H. El-Hussieny, A. Abouelsoud, S. F. Assal, and S. M. Megahed, "Adaptive learning of human motor behaviors: An evolving inverse optimal control approach," *Engineering Applications of Artificial Intelligence*, vol. 50, pp. 115–124, Apr. 2016.

[12] I. A. Faruque, F. T. Muijres, K. M. Macfarlane, A. Kehlenbeck, and J. S. Humbert, "Identification of optimal feedback control rules from micro-quadrotor and insect flight trajectories," *Biological Cybernetics*, vol. 112, no. 3, pp. 165–179, Jun. 2018.

[13] 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, Apr. 2010.

[14] T. L. Molloy, J. Ford, and T. Perez, "Inverse Noncooperative Differential Games," *IEEE International Conference on Decision and Control*, pp. 5602–5608, 2017.

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

[16] 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 *IEEE Conference on Decision and Control (CDC)*. IEEE, Dec. 2016, pp. 6327–6334.

[17] D. Tsai, T. L. Molloy, and T. Perez, "Inverse two-player zero-sum dynamic games," in *2016 Australian Control Conference (AuCC)*, Nov. 2016, pp. 192–196.

[18] F. Köpf, J. Inga, S. Rothfuß, M. Flad, and S. Hohmann, "Inverse Reinforcement Learning for Identification in Linear-Quadratic Dynamic Games," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 14 902–14 908, 2017.

[19] C. Carraro, J. Flemming, and A. Giovannini, "The tastes of european central bankers," in *A European Central Bank?: Perspectives on Monetary Unification after Ten Years of the EMS*. Cambridge University Press, 1989, pp. 162–185.

[20] F. Nori and R. Frezza, "Linear optimal control problems and quadratic cost functions estimation," in *Proceedings of the Mediterranean Conference on Control and Automation*, vol. 4, 2004.

[21] J. C. Engwerda, W. A. v. d. Broek, and J. M. Schumacher, "Feedback Nash equilibria in uncertain infinite time horizon differential games," *Proceedings of the 14th International Symposium of Mathematical Theory of Networks and Systems, MTNS 2000*, pp. 1–6, 2000.

[22] J. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Transactions on Circuits and Systems*, vol. 25, no. 9, pp. 772–781, Sep. 1978.

[23] F. Zhang, *Matrix theory: basic results and techniques*, 2nd ed. New York: Springer, 2011.

[24] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory: Second Edition*. SIAM, 1999.

[25] A. J. T. M. Weeren, "A Geometric Approach to Infinite Horizon Linear Quadratic Differential Games," *IFAC Proceedings Volumes*, vol. 34, no. 20, pp. 35–40, Sep. 2001.

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

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