Preprints of the 23rd IFAC World Congress
Busan, Republic of Korea, August 23-28, 2026

On Data-based Nash Equilibria in LQ Nonzero-sum Differential Games ⋆

Victor G. Lopez, * **Matthias A. Müller** *

* *Leibniz University Hannover, Institute of Automatic Control, 30167 Hannover, Germany (e-mail: { lopez, mueller } @irt.uni-hannover.de).*

Abstract: This paper considers data-based solutions of linear-quadratic nonzero-sum differential games. Two cases are considered. First, the deterministic game is solved and Nash equilibrium strategies are obtained by using persistently excited data from the multiagent system. Then, a stochastic formulation of the game is considered, where each agent measures a different noisy output signal and state observers must be designed for each player. It is shown that the proposed data-based solutions of these games are equivalent to known model-based procedures. The resulting data-based solutions are validated in a numerical experiment.

Keywords: Differential or dynamic games, Data-driven control theory, Design methods for data-based control, State estimation, Linear systems.

1. INTRODUCTION

Nonzero-sum (NZS) differential games model the interaction of several systems or processes, regarded as players, that influence each other while trying to achieve their individual goals. The theory of differential games has been used to analyze optimal behavior in multiagent dynamical systems for applications in diverse areas of study including engineering, economics, sociology, among many others (Basar and Olsder, 1999). In a game among rational agents, the most important solution concept occurs when each player uses its best policy with respect to the best policy of the other players. These strategies are said to form a Nash equilibrium. Differential games have been studied for many decades (Starr and Ho, 1969; Basar and Olsder, 1999; Engwerda, 2005) and the conditions for the existence of Nash equilibria are well known. In the case of infinite-horizon linear-quadratic NZS differential games, the Nash policies are obtained by solving a set of coupled algebraic Riccati equations (AREs) (Lewis et al., 2012). Solving the set of coupled AREs is not trivial, and different algorithms to obtain such solutions have been proposed (Li and Gajic, 1995; Engwerda, 2007; Possieri and Sassano, 2015). These algorithms require full knowledge of the mathematical model of the system.

In recent years, there has been an increasing interest in control design using only measured data. Many of the recent results in data-based control are based on Willems' fundamental lemma (Willems et al., 2005), which states that every input-output trajectory of a dynamical linear system can be expressed in terms of a single persistently excited trajectory. In the current literature, there are now many different data-based control methods for discrete-time (see, e.g., Markovsky and Dörfler (2023) and the references therein) and continuous-time systems (Rapisarda et al., 2024; Eising and Cortés, 2025; Lopez and Müller, 2026). Some model-free solutions of nonzero-sum differential games have been obtained using dynamic programming methods (Vamvoudakis, 2015; Song et al., 2021; Xie et al., 2025). These methods often exploit the properties of the Kronecker product, resulting in inefficient computations (Lopez and Müller, 2023). Data-based solutions for discrete-time NZS games have also been obtained using ideas based on Willems' lemma in Nortmann et al. (2024), but those algorithms require collecting new data from the system at every iteration.

The literature discussed above concerns results for deterministic NZS games, where noise-free measurements of the full state vector are assumed to be available. In Rhodes and Luenberger (1969); Saksena and Cruz (1982), the stochastic nonzero-sum differential game is solved, where each player has access to a different noisy output signal. As it is discussed there, the truly optimal (Nash equilibrium) strategies require the use of infinite-dimensional observers. A practical solution is obtained by determining instead the optimal controllers based on an n-dimensional observer. The solution of this problem does not currently have a data-based counterpart. Note that in Zhang et al. (2024), a different stochastic differential game is solved.

The main contribution of this paper is leveraging the continuous-time Willems' lemma in (Lopez and Müller, 2022) to obtain model-free solutions of NZS games. We show how to collect persistently excited data only once from the multiagent system to design the data-based controllers. Both the deterministic and the stochastic NZS games are formulated and solved. For the stochastic case, as it is standard in the literature on data-based state estimation (Wolff et al., 2022; Turan and Ferrari-Trecate, 2022), we assume that the state of the system can be measured in an offline phase, whereas only output measurements are available during the online operation. This allows to use input-state-output persistently excited data to obtain the desired results.

⋆ This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 948679).

©2026 the authors. Accepted by IFAC for publication under a Creative Commons Licence CC-BY-NC-ND

The remainder of this paper is structured as follows. Section 2 presents preliminary results and definitions. The data-based solutions of NZS games are presented in Section 3, including the deterministic game (Section 3.2) and the stochastic one (Section 3.3). Section 4 presents a simulation example and Section 5 concludes the paper.

2. PRELIMINARIES

2.1 Deterministic nonzero-sum differential games

Consider a dynamical system of the form

$$\dot{x} = Ax + \sum_{i=1}^{N} B_i u_i \tag{1}$$

where each input $u_i \in \mathbb{R}^{m_i}$ is used to control the state $x \in \mathbb{R}^n$ while minimizing the cost function

$$J_i = \int_0^\infty \left(x^\top Q_i x + u_i^\top R_{ii} u_i + \sum_{j \neq i} u_j^\top R_{ij} u_j \right) dt, \tag{2}$$

with $Q_i \succeq 0$, $R_{ii} \succ 0$, and $R_{ij} \succeq 0$. This is the formulation of a nonzero-sum (NZS) game, where the input u_i corresponds to the control action of player i. Since the state x is affected by all the inputs u_i, each cost J_i depends on the actions of all players. A Nash equilibrium is defined as a set of control policies u_i^*, $i = 1, \ldots, N$, such that

$$J_i(u_i^*, u_{-i}^*) \leq J_i(u_i, u_{-i}^*)$$

for every input u_i and every player i. Here, u_{-i} denotes the set of inputs of all players except player i. In a NZS game there may be no Nash equilibrium solution, a unique solution, or multiple solutions (Engwerda, 2005).

When the model (1) is known, a Nash equilibrium can be obtained by policies of the form $u_i^* = -K_i^* x$, where

$$K_i^* = R_{ii}^{-1} B_i^\top P_i, \tag{3}$$

and the matrices $P_i \succ 0$ solve the coupled algebraic Riccati equations (ARE) (Lewis et al. (2012))

$$\begin{aligned} 0 &= P_i A + A^\top P_i + Q_i - P_i S_i P_i \\ &+ \sum_{j \neq i} \left(P_j B_j R_{jj}^{-1} R_{ij} R_{jj}^{-1} B_j^\top P_j - P_j S_j P_i - P_i S_j P_j \right), \end{aligned} \tag{4}$$

for $i = 1, \ldots, N$, where

$$S_i = B_i R_{ii}^{-1} B_i^\top. \tag{5}$$

Diverse methods to solve the coupled AREs (4) have been studied in the literature Li and Gajic (1995); Engwerda (2007); Possieri and Sassano (2015).

2.2 Stochastic nonzero-sum differential games

In Rhodes and Luenberger (1969), a stochastic zero-sum game was formulated and solved. This result was extended to nonzero-sum games in Saksena and Cruz (1982). In these articles and many others (Chong and Athans (1971); Li and Gajic (1995)), the game is formulated as having only two players for notation simplicity. The extension to the multi-player case can be straightforwardly obtained. Thus, consider now the dynamical system

$$\dot{x} = Ax + B_1 u_1 + B_2 u_2 + w \tag{6a}$$

$$y_i = C_i x + v_i, \quad i = 1, 2, \tag{6b}$$

where the pairs (A, C_1) and (A, C_2) are observable, $y_i \in \mathbb{R}^{p_i}$ is the output measurement of player i, and the noise vectors $w \in \mathbb{R}^n$, $v_1 \in \mathbb{R}^{p_1}$, and $v_2 \in \mathbb{R}^{p_2}$ are independent Gaussian disturbances with zero mean and covariances W, V_1, and V_2, respectively. The initial condition of the state is also a Gaussian random vector, independent of the noises w, v_1, and v_2, with mean $\bar{x}_0$ and covariance

$$E\{x(0)x(0)^\top\} = M_0.$$

The objective of each player is to minimize the expected cost $E\{J_i\}$ with J_i as in (2). Since the players do not have state measurements available, state observers are designed. It is known that, if no restriction is imposed over them, Nash equilibrium can only be achieved by using infinite-dimensional observers (Rhodes and Luenberger (1969)). To avoid this problem, the objective of this game is to determine the optimal controllers for all agents, when they are restricted to use n-dimensional observers.

The solution of this game is given by the dynamic controller (Saksena and Cruz (1982))

$$u_i = -R_{ii}^{-1} B_i^\top P_i \hat{x}^i, \tag{7}$$

$$\dot{\hat{x}}^i = F_i \hat{x}^i + B_i u_i + L_i \left(y_i - C_i \hat{x}^i \right), \tag{8}$$

where $\hat{x}^1(0) = \hat{x}^2(0) = \bar{x}_0$, P_i solves the AREs (4),

$$F_1 = A - S_2 P_2 \left(I + (M_{20} - M_{21})(M_{00} - M_{01})^{-1} \right), \tag{9}$$

$$F_2 = A - S_1 P_1 \left(I + (M_{10} - M_{12})(M_{00} - M_{02})^{-1} \right). \tag{10}$$

Moreover,

$$L_i = M_{ii} C_i^\top V_i^{-1}, \tag{11}$$

S_i is as in (5), and the matrix M, partitioned as

$$M = \begin{bmatrix} M_{00} & M_{01} & M_{02} \\ M_{10} & M_{11} & M_{12} \\ M_{20} & M_{21} & M_{22} \end{bmatrix} \tag{12}$$

with $M_{ij} \in \mathbb{R}^{n \times n}$, satisfies the differential equation

$$\dot{M} = \bar{A} M + M \bar{A}^\top + \bar{B} \bar{W} \bar{B}^\top \tag{13}$$

with initial conditions $M_{ij}(0) = M_0 + \bar{x}_0 \bar{x}_0^\top$ for $i = j = 0$, and $M_{ij}(0) = M_0$ otherwise. Here, $\bar{W} = \text{blockdiag}\{W, V_1, V_2\}$,

$$\bar{A} = \begin{bmatrix} A - S_1 P_1 - S_2 P_2 & S_1 P_1 & S_2 P_2 \\ A - F_1 - S_2 P_2 & F_1 - L_1 C_1 & S_2 P_2 \\ A - F_2 - S_1 P_1 & S_1 P_1 & F_2 - L_2 C_2 \end{bmatrix} \tag{14}$$

and

$$\bar{B} = \begin{bmatrix} -I & 0 & 0 \\ -I & L_1 & 0 \\ -I & 0 & L_2 \end{bmatrix}.$$

2.3 Persistence of excitation for continuous-time systems

Lopez and Müller (2022) showed a persistence of excitation condition for continuous-time systems. Here, these results are formulated for case of the multiagent system (1).

Consider an integer $\eta \in \mathbb{N}$ and a positive scalar $T \in \mathbb{R}_+$. Suppose that the input signals u_i, $i = 1, \ldots, N$, are applied to (1) and an input-state trajectory is collected in the time interval $[0, \eta T]$. Define the following matrices for $0 \leq t \leq T$

$$\mathcal{H}_T(x(t)) := [\, x(t) \;\; x(t+T) \;\; \cdots \;\; x(t+(\eta-1)T) \,], \tag{15}$$

$$\mathcal{H}_T(u_i(t)) := [\, u_i(t) \;\; u_i(t+T) \;\; \cdots \;\; u_i(t+(\eta-1)T) \,],$$

$i = 1, \ldots, N$. Notice that, from (1) and (15), we can write $\frac{d}{dt}\mathcal{H}_T(x(t)) = A\mathcal{H}_T(x(t)) + \sum_{i=1}^N B_i \mathcal{H}_T(u_i(t))$. Taking the integral on both sides of this expression, we get

$$\mathcal{H}_T(x(T)) - \mathcal{H}_T(x(0)) = A\int_0^T \mathcal{H}_T(x(\tau))d\tau + \sum_{i=1}^N B_i \int_0^T \mathcal{H}_T(u_i(\tau))d\tau. \quad (16)$$

To simplify the notation, define the data matrices

$$\tilde{\mathcal{H}}_S(x) := \mathcal{H}_T(x(T)) - \mathcal{H}_T(x(0)), \quad (17)$$

$$\mathcal{H}_S(x) := \int_0^T \mathcal{H}_T(x(\tau))d\tau, \quad \mathcal{H}_S(u_i) := \int_0^T \mathcal{H}_T(u_i(\tau))d\tau,$$

$$\mathcal{H}_S(u) := \left[\mathcal{H}_S(u_1)^\top \ \cdots \ \mathcal{H}_S(u_N)^\top\right]^\top.$$

We highlight again that, from (16), we can write

$$\tilde{\mathcal{H}}_S(x) = A\mathcal{H}_S(x) + \sum_{i=1}^N B_i \mathcal{H}_S(u_i). \quad (18)$$

The matrices $\mathcal{H}_S(x)$, $\mathcal{H}_S(u_i)$ satisfy an important rank property when the data used to compute them satisfy a persistence of excitation condition. For the following definition, consider the concatenation of control inputs $u := [u_1^\top \ \cdots \ u_N^\top]^\top \in \mathbb{R}^m$, with $m = \sum_{i=1}^N m_i$.

Definition 1. (Piecewise constant PE input). A piecewise constant persistently exciting (PCPE) input of order L is defined as $u(t + jT) = \mu_j$, $0 \le t < T$, $j = 0, \dots, \eta - 1$, where the sequence $\{\mu_j\}_{j=0}^{\eta-1}$, $\mu_j \in \mathbb{R}^m$, is such that

$$\operatorname{rank}\left(\begin{bmatrix} \mu_0 & \mu_1 & \cdots & \mu_{\eta-L} \\ \mu_1 & \mu_2 & \cdots & \mu_{\eta-L+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{L-1} & \mu_L & \cdots & \mu_{\eta-1} \end{bmatrix}\right) = mL,$$

and where the constant T does not correspond to a pathological sampling time (Lopez and Müller, 2022).

Lemma 2. (Lopez and Müller (2023), Lemma 4). Consider the system (1), let the pair $(A, [B_1 \ \cdots \ B_N])$ be controllable, and let $u = [u_1^\top \ \cdots \ u_N^\top]^\top$ be a PCPE input of order $n + 1$. Then,

$$\operatorname{rank}\left(\begin{bmatrix} \mathcal{H}_S(x) \\ \mathcal{H}_S(u) \end{bmatrix}\right) = n + m. \quad (19)$$

Note that collecting PE data from a multiagent system implies a degree of cooperation among the players, since the global input u must be persistently exciting.

3. DATA-BASED SOLUTION OF NZS GAMES

This section first shows useful data-based expressions for multiagent systems. Then, we present our data-based solutions to the deterministic NZS game described in Section 2.1, as well as the stochastic game in Section 2.2.

3.1 Useful data-based expressions

The rank condition (19) allows obtaining the data-based representation of closed loop systems (De Persis and Tesi, 2020). In particular, we show the following result.

Lemma 3. Consider the system (1) and let the pair $(A, [B_1 \ \cdots \ B_N])$ be controllable. Suppose that $u = [u_1^\top \cdots u_N^\top]^\top$ is a PCPE input of order $n + 1$ as in Definition 1. Then, for any set of matrices $K_i \in \mathbb{R}^{m_i \times n}$, $i = 1, \dots, N$, there exists a matrix Γ such that

$$\tilde{\mathcal{H}}_S(x)\Gamma = A - \sum_{i=1}^N B_i K_i. \quad (20)$$

Proof. Since the concatenation of inputs u is PCPE of order $n+1$, by Lemma 2 we have that there exist a matrix Γ such that

$$\begin{bmatrix} \mathcal{H}_S(x) \\ \mathcal{H}_S(u) \end{bmatrix} \Gamma = \begin{bmatrix} I & -K_1^\top & \cdots & -K_N^\top \end{bmatrix}^\top \quad (21)$$

for any matrices K_i. Moreover, from (18) we have

$$\tilde{\mathcal{H}}_S(x)\Gamma = [A \ B_1 \ \cdots \ B_N] \begin{bmatrix} \mathcal{H}_S(x) \\ \mathcal{H}_S(u) \end{bmatrix} \Gamma$$

which, together with (21), implies (20). □

The algorithm that is described in the next subsection to solve the coupled AREs (4) requires knowledge of initial stabilizing matrices K_i. Such stabilizing controllers can be obtained from data by exploiting Lemma 3. Here, we suggest to solve an auxiliary, global optimal control problem to determine such controllers, since they will be convenient later in this section. Notice that this auxiliary problem considers the global input $u := [u_1^\top \ \cdots \ u_N^\top]^\top$ and does not correspond to a game-theoretic problem.

Thus, define some auxiliary weighting matrices $\mathcal{Q} \in \mathbb{R}^{n \times n}$, $\mathcal{R} \in \mathbb{R}^{m \times m}$. The matrix $\mathcal{Q} \succeq 0$ can be chosen arbitrarily, but it will be useful to define $\mathcal{R} = \operatorname{blockdiag}_i\{R_{ii}\}$. The objective now is to determine the solution of the LQR problem defined by the matrices $(A, [B_1 \ \cdots \ B_N], \mathcal{Q}, \mathcal{R})$. In Lopez and Müller (2026), the LQR problem is solved for continuous-time systems using only measured data. Here, we rewrite this algorithm using our multiagent system formulation. First, determine the matrix $\mathcal{P} \in \mathbb{R}^{n \times n}$ that solves the optimization problem

$$\begin{aligned} \underset{P}{\text{maximize}} \quad & \operatorname{tr}(P) \\ \text{s.t.} \quad & P \succ 0, \quad \mathcal{L}(P) \succeq 0, \end{aligned} \quad (22)$$

where

$$\mathcal{L}(P) = \mathcal{H}_S(x)^\top \mathcal{Q} \mathcal{H}_S(x) + \mathcal{H}_S(u)^\top \mathcal{R} \mathcal{H}_S(u) + \mathcal{H}_S(x)^\top P \tilde{\mathcal{H}}_S(x) + \tilde{\mathcal{H}}_S(x)^\top P \mathcal{H}_S(x). \quad (23)$$

Then, the obtained matrix $\mathcal{P}$ is used to find Γ^0 such that

$$\begin{bmatrix} \mathcal{H}_S(x) \\ \mathcal{L}(\mathcal{P}) \end{bmatrix} \Gamma^0 = \begin{bmatrix} I \\ 0 \end{bmatrix}. \quad (24)$$

Finally, the matrices K_i^0 that solve this auxiliary LQR problem are given by

$$K_i^0 = -\mathcal{H}_S(u_i)\Gamma^0. \quad (25)$$

Lemma 4. Let the conditions in Lemma 2 hold and compute the matrix K_i^0 from the procedure described in (22)-(25). Then, K_i^0 is the optimal solution to the LQR problem defined by the matrices $(A, [B_1 \ \cdots \ B_N], \mathcal{Q}, \mathcal{R})$.

Proof. The proof follows as in (Lopez and Müller, 2026, Theorem 17), where it is shown that the global optimal controller has the form $K^0 = -\mathcal{H}_S(u)\Gamma^0$, which implies (25) for each player. □

Since the matrices K_i^0 are stabilizing, they will be used to initialize the algorithm in the next subsection. Besides, recall that these matrices satisfy $\left[K_1^{0\top} \ \cdots \ K_N^{0\top}\right]^\top = \mathcal{R}^{-1}[B_1 \ \cdots \ B_N]^\top \mathcal{P}$ and, therefore,

$$K_i^0 = R_{ii}^{-1} B_i^\top \mathcal{P}. \quad (26)$$

In the following, it will be useful to determine data-based expressions for the matrices S_i in (5) for all players. This can be achieved by exploiting (26) and Lemma 3.

First notice that, from (24) and (25), Γ^0 satisfies

$$\begin{bmatrix}\mathcal{H}_S(x)\\ \mathcal{H}_S(u)\end{bmatrix}\Gamma^0 = \begin{bmatrix}I & -K_1^{0\top} & \cdots & -K_N^{0\top}\end{bmatrix}^\top. \tag{27}$$

Moreover, let $\bar{\Gamma}_i$ satisfy (27) but replacing K_i with the zero matrix, where $\bar{\Gamma}_i$ is guaranteed to exist due to the rank condition (19). This means that, as in (20), $\tilde{\mathcal{H}}_S(x)\Gamma^0 = A - \sum_{i=1}^N B_i K_i^0$ and $\tilde{\mathcal{H}}_S(x)\bar{\Gamma}_i = A - \sum_{j\neq i} B_j K_j^0$. Therefore, we have

$$\tilde{\mathcal{H}}_S(x)(\bar{\Gamma}_i - \Gamma^0) = B_i K_i^0 = B_i R_{ii}^{-1} B_i^\top \mathcal{P}.$$

Now, we can compute S_i in (5) for all i using data as

$$S_i = \tilde{\mathcal{H}}_S(x)(\bar{\Gamma}_i - \Gamma^0)\mathcal{P}^{-1}. \tag{28}$$

3.2 Data-based solution of the deterministic NZS game

In Li and Gajic (1995), an iterative procedure to solve the coupled AREs (4) was proposed. Algorithm 1 below shows this method for the setting considered in this paper.

Algorithm 1: Model-based solution of coupled AREs (Li and Gajic (1995))

1: **procedure**
2: Let $k = 0$ and initialize the matrices K_i^0 such that $A - \sum_{i=1}^N B_i K_i^0$ is Hurwitz.
3: Solve the following N decoupled Lyapunov equations for $P_i^k \succ 0$

$$P_i^k\left(A - \sum_{i=1}^N B_i K_i^k\right) + \left(A - \sum_{i=1}^N B_i K_i^k\right)^\top P_i^k = - Q_i - (K_i^k)^\top R_{ii} K_i^k - \sum_{j\neq i}(K_j^k)^\top R_{ij} K_j^k \tag{29}$$

4: Using the solutions P_i^k, update the feedback matrices

$$K_i^{k+1} = R_{ii}^{-1} B_i^\top P_i^k, \quad i = 1,\dots,N. \tag{30}$$

5: Set $k = k+1$ and go to step 3 until convergence.
6: **end procedure**

Our data-based solution to this problem makes use of the following calculations. First, as in (27), given some matrices K_i^k, $i = 1,\dots,N$, compute Γ^k such that

$$\begin{bmatrix}\mathcal{H}_S(x)\\ \mathcal{H}_S(u)\end{bmatrix}\Gamma^k = \begin{bmatrix}I & -K_1^{k\top} & \cdots & -K_N^{k\top}\end{bmatrix}^\top. \tag{31}$$

Moreover, define $\Phi_i^k \in \mathbb{R}^{\eta\times\eta}$ such that

$$\begin{bmatrix}\mathcal{H}_S(x)\\ \mathcal{H}_S(u_1)\\ \vdots\\ \mathcal{H}_S(u_i)\\ \vdots\\ \mathcal{H}_S(u_N)\end{bmatrix}\Phi_i^k = \begin{bmatrix}\mathcal{H}_S(x)\\ -K_1^k\mathcal{H}_S(x)\\ \vdots\\ -K_i^k\mathcal{H}_S(x) + \mathcal{H}_S(u_i)\\ \vdots\\ -K_N^k\mathcal{H}_S(x)\end{bmatrix}, \tag{32}$$

where, on the right-hand side, only the block row corresponding to the ith input includes a term of the form $\mathcal{H}_S(u_i)$. Such matrices Φ_i^k always exist due to (19). From this definition and (18), it is easy to show that

$$\tilde{\mathcal{H}}_S(x)\Phi_i^k = \left(A - \sum_{i=1}^N B_i K_i^k\right)\mathcal{H}_S(x) + B_i\mathcal{H}_S(u_i). \tag{33}$$

Using Φ_i^k, define also the operator $\bar{\mathcal{L}}_i^k(P)$ as

$$\begin{aligned}\bar{\mathcal{L}}_i^k(P) := {}& \mathcal{H}_S(x)^\top(\bar{Q}_i + PS_iP)\mathcal{H}_S(x)\\ &+ \mathcal{H}_S(u_i)^\top R_{ii}\mathcal{H}_S(u_i) + \mathcal{H}_S(x)^\top P\tilde{\mathcal{H}}_S(x)\Phi_i^k\\ &+ (\Phi_i^k)^\top\tilde{\mathcal{H}}_S(x)^\top P\mathcal{H}_S(x),\end{aligned} \tag{34}$$

with $\bar{Q}_i = Q_i + (K_i^k)^\top R_{ii}K_i^k + \sum_{j\neq i}(K_j^k)^\top R_{ij}K_j^k$ and S_i is as in (28).

Our data-based solution to the deterministic NZS game is shown in Algorithm 2. In Theorem 5, we show the equivalence of the model-based Algorithm 1 and the data-based Algorithm 2.

Algorithm 2: Data-based solution of coupled AREs

1: **procedure**
2: Collect data from system (1) by applying inputs u_i such that $u = [u_1^\top \cdots u_N^\top]^\top$ is PCPE of order $n+1$.
3: Let $k = 0$, initialize the matrices K_i^0 by solving (22)-(25), and compute the matrices S_i as in (28).
4: Using K_i^k, compute Γ^k as in (31) and Φ_i^k as in (32).
5: Solve the following N decoupled Lyapunov equations for $P_i^k \succ 0$

$$\begin{aligned}P_i^k\tilde{\mathcal{H}}_S(x)\Gamma^k + (\Gamma^k)^\top\tilde{\mathcal{H}}_S(x)^\top P_i^k = {}& \\ - Q_i - (K_i^k)^\top R_{ii}K_i^k - \sum_{j\neq i}(K_j^k)^\top R_{ij}K_j^k.&\end{aligned} \tag{35}$$

6: Using the solutions P_i^k, compute $\bar{\Gamma}_i^k$, $i = 1,\dots,N$, from

$$\begin{bmatrix}\mathcal{H}_S(x)\\ \bar{\mathcal{L}}_i^k(P_i^k)\end{bmatrix}\bar{\Gamma}_i^k = \begin{bmatrix}I\\ 0\end{bmatrix}, \tag{36}$$

with $\bar{\mathcal{L}}_i^k(P_i^k)$ as in (34).
7: Update the feedback matrices

$$K_i^{k+1} = -\mathcal{H}_S(u_i)\bar{\Gamma}_i^k, \quad i = 1,\dots,N. \tag{37}$$

8: Set $k = k+1$ and go to step 4 until convergence.
9: **end procedure**

Theorem 5. Consider the system (1) and let the pair $(A, [B_1 \ \cdots \ B_N])$ be controllable. Then, Algorithm 2 is equivalent to Algorithm 1 in the sense that the same matrices P_i^k, K_i^{k+1}, $i = 1,\dots,N$, are obtained at every iteration.

Proof. The data-based version of (29) in Algorithm 1 is easy to obtain because, from Lemma 3, we have that $\tilde{\mathcal{H}}_S(x)\Gamma^k = A - \sum_{i=1}^N B_i K_i^k$. A direct substitution in (29) yields (35). Hence, it only remains to be shown that (37) and (30) are also equivalent.

The fact that P_i^k solves (29), implies also that it satisfies

$$P_i^k A_c^k + (A_c^k)^\top P_i^k + \tilde{Q}_i - P_i^k B_i R_{ii}^{-1} B_i^\top P_i^k = 0, \tag{38}$$

where $A_c^k := A - \sum_{i=1}^N B_i K_i^k$ and $\tilde{Q}_i := \bar{Q}_i + P_i^k B_i R_{ii}^{-1} B_i^\top P_i^k$. Thus, P_i^k can be interpreted as the solution of the ARE (38), corresponding to an LQR problem defined by the matrices $(A_c^k, B_i, \tilde{Q}_i, R_{ii})$. The solution of this auxiliary LQR problem is given by $K_i^{aux} = R_{ii}^{-1}B_i^\top P_i^k$ which, since P_i^k satisfies (29), corresponds precisely to (30) in Algoritm 1. Such an LQR problem could be solved in a data-based fashion by using a similar procedure to (22)-(25), but with $\mathcal{L}(P)$ replaced by $\bar{\mathcal{L}}_i^k(P)$ (see the analysis in Section 4.2 of Lopez and Müller (2026) for details about the data-based solution of the LQR problem). Since in our case the solution P_i^k of (38) was obtained from the Lyapunov equation (35), it only remains to compute K_i^{aux} similarly as in (24)-(25), from which the expressions in (36) and (37) are obtained. □

3.3 Data-based solution of the stochastic NZS game

Similarly as in the previous subsection, here we obtain a data-based formulation of the existing solution of the stochastic NZS game described in Section 2.2. Again, this is achieved by collecting persistently excited data from the system (6). However, in the following we assume that, during this (offline) data-collection phase, the players have access to noise-free measurements of the input-state-output of the system. Although this assumption is restrictive in general, it has been used in the data-based state estimation literature (Wolff et al., 2022; Turan and Ferrari-Trecate, 2022), and can be fulfilled if such offline measurements are obtained in a dedicated laboratory setting equipped with advanced sensors. Such sensors are not used during the online application of the controller/state estimator, when only noisy input-ouput information is available. Robustifying this method for the case of offline noisy data is a subject of future work.

Hence, recalling that this game is formulated for only two players for simplicity, assume that the pair $(A, [B_1 \quad B_2])$ is controllable. Then, after applying the inputs u_1, u_2 to the system (6) such that $u = [u_1^\top \quad u_2^\top]^\top$ is a PCPE input of order $n+1$, compute the data matrices $\tilde{\mathcal{H}}_S(x)$, $\mathcal{H}_S(x)$, and $\mathcal{H}_S(u_i)$, $i = 1, 2$, as in (17), as well as the matrices

$$\mathcal{H}_S(y_i) := \int_0^T \mathcal{H}_T(y_i(\tau))d\tau, \quad i = 1, 2,$$

with $\mathcal{H}_T(y_i(t))$ defined similarly as (15). Notice that, since noise-free data is assumed to be collected, it holds that

$$\mathcal{H}_S(y_i) = C_i \mathcal{H}_S(x). \tag{39}$$

Now, the data-based version of the solution of the stochastic NZS game described in Section 2.2 can be obtained as follows. First, notice that the feedback matrix in (7) for each player corresponds to the deterministic Nash equilibrium policy, which can be obtained from Algorithm 2. Moreover, substituting (7) in (8), the state observer can be written as

$$\dot{\hat{x}}^i = \bar{F}_i \hat{x}^i + L_i \left(y_i - C_i \hat{x}^i\right), \tag{40}$$

where, taking S_i as in (5),

$\bar{F}_1 = A - S_1P_1 - S_2P_2 \left(I + (M_{20} - M_{21})(M_{00} - M_{01})^{-1}\right)$,

$\bar{F}_2 = A - S_2P_2 - S_1P_1 \left(I + (M_{10} - M_{12})(M_{00} - M_{02})^{-1}\right)$.

Notice that we can also write these matrices as

$\bar{F}_i = A - B_1K_1^* - B_2K_2^* - S_jP_j(M_{j0} - M_{ji})(M_{00} - M_{0i})^{-1}$, where $i \neq j$, and with K_i^* as in (3). Hence, we obtain a fully data-based method by finding model-free expressions for the matrices $\bar{F}_i$, L_i and $\bar{A}$.

Obtaining L_i is straightforward by noticing that

$$\begin{aligned} C_i &= C_i \mathcal{H}_S(x) \mathcal{H}_S(x)^\top (\mathcal{H}_S(x) \mathcal{H}_S(x)^\top)^{-1} \\ &= \mathcal{H}_S(y_i) \mathcal{H}_S(x)^\top (\mathcal{H}_S(x) \mathcal{H}_S(x)^\top)^{-1}. \end{aligned}$$

Hence, from (11) we have

$$L_i = M_{ii} (\mathcal{H}_S(x) \mathcal{H}_S(x)^\top)^{-1} \mathcal{H}_S(x) \mathcal{H}_S(y_i)^\top W_i^{-1}.$$

Now, using the matrices K_i^* that result from Algorithm 2, and exploiting again the rank condition (19), compute Γ^* from the set of equations

$$\begin{bmatrix} \mathcal{H}_S(x) \\ \mathcal{H}_S(u_1) \\ \mathcal{H}_S(u_2) \end{bmatrix} \Gamma^* = \begin{bmatrix} I \\ -K_1^* \\ -K_2^* \end{bmatrix}. \tag{41}$$

As in Lemma (3), this matrix gives us a data-based representation of the closed loop system $\tilde{\mathcal{H}}_S(x)\Gamma^* = A - B_1K_1^* - B_2K_2^*$. Using this result, we can write

$$\begin{aligned} \bar{F}_1 &= \tilde{\mathcal{H}}_S(x)\Gamma^* - S_2P_2^*(M_{20} - M_{21})(M_{00} - M_{01})^{-1}, \\ \bar{F}_2 &= \tilde{\mathcal{H}}_S(x)\Gamma^* - S_1P_1^*(M_{10} - M_{12})(M_{00} - M_{02})^{-1}. \end{aligned}$$

Finally, the matrix $\bar{A}$ in (14) can be obtained as

$$\bar{A} = \begin{bmatrix} \tilde{\mathcal{H}}_S(x)\Gamma^* & S_1P_1 & S_2P_2 \\ S_2P_2\tilde{M}_1 & \bar{F}_1 + S_1P_1 - L_1C_1 & S_2P_2 \\ S_1P_1\tilde{M}_2 & S_1P_1 & \bar{F}_2 + S_2P_2 - L_2C_2 \end{bmatrix}, \tag{42}$$

where $\tilde{M}_1 = (M_{20} - M_{21})(M_{00} - M_{01})^{-1}$ and $\tilde{M}_2 = (M_{10} - M_{12})(M_{00} - M_{02})^{-1}$. This can be seen by substituting (9)-(10) in the first column of (14), and performing the resulting cancellations. Moreover, in the remaining columns of (14) we used the fact that $F_i = \bar{F}_i + S_iP_i$.

Remark 6. In this subsection, it was assumed that noise-free data could be collected from the system during the offline phase. Computing the matrices above using noisy data is akin to the presence of parametric uncertainties in the model-based case. See, e.g., (Lopez et al., 2023, Section IV) where convergence of an iterative algorithm with noisy data is analyzed. Notice that arbitrarily small magnitudes of the noise would imply that the obtained controllers are arbitrarily close to the optimal ones.

4. SIMULATION RESULTS

In this section we verify the applicability of the proposed methods with a numerical simulation. Since both the deterministic and the stochastic NZS games studied in this paper make use of Algorithm 2, here we focus only on simulating the stochastic case.

We consider a system in the form (6) where the matrices A, B_1 and B_2, as well as the cost function matrices Q_1, Q_2, R_{11}, R_{12}, R_{21}, and R_{22}, are the same used in Numerical Example 1 in (Li and Gajic, 1995). This is a 10-dimensional system, where each player uses a 2-dimensional input. For the output matrices we choose $C_1 = [I \quad 0]$ and $C_2 = [0 \quad I]$, such that each player measures different states. The noise signals are assumed to have covariances $W = 10^{-5}I$, and $V_1 = V_2 = 10^{-2}I$, where the identity matrices have the proper dimensions. The objective for each player is to stabilize the system states while minimizing its expected cost $E\{J_i\}$. Algorithm 2 is run for 15 iterations. It can be verified that the obtained matrices P_i^* satisfy the coupled AREs (4). Then, the data-based state observer designed in Section 3.3 is implemented in the control loop. The results are displayed in Figure 1, which shows that the 10 states are stabilized. Figure 2 shows the state estimation of the first state, for both players.

5. CONCLUSION

We proposed novel data-based solutions for two classes of nonzero-sum differential games. First, we solve the deterministic game where all players can measure the complete system state. Then, we solved the stochastic game in which the players have only access to noisy output measurements during the online application of the controller.

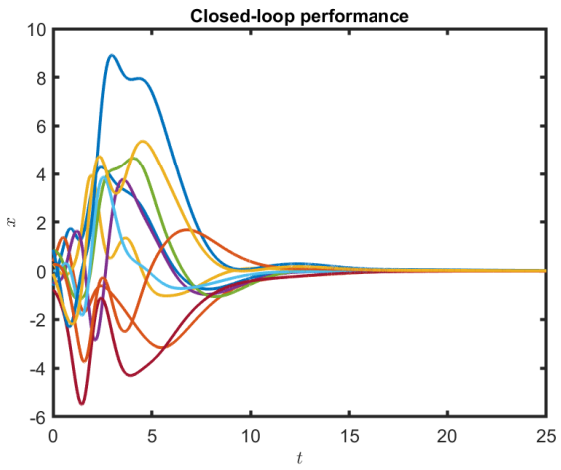


Fig. 1. Performance of the proposed data-based controller.

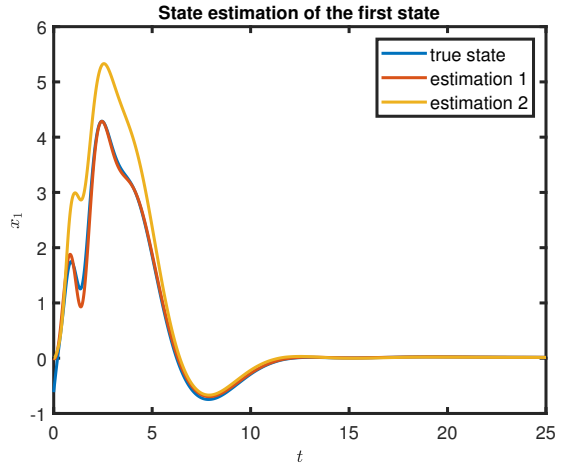


Fig. 2. State estimation of x_1 for both players

The obtained solutions are equivalent to their model-based counterparts and provide the desired outcomes under the same conditions. However, the proposed methods avoid the need for model knowledge by means of the collection of persistently excited data. As future work, we aim to relax the requirement in Section 3.3 of collecting noise-free data in the offline phase. Moreover, we will investigate the design of data-based solutions to differential graphical games, where distributed controllers must be obtained.

REFERENCES

Basar, T. and Olsder, G.J. (1999). *Dynamic Noncooperative Game Theory*. PA: SIAM, Philadelphia, USA, 2 edition.

Chong, C.Y. and Athans, M. (1971). On the stochastic control of linear systems with different information sets. *IEEE Transactions on Automatic Control*, 16(5), 423–430.

De Persis, C. and Tesi, P. (2020). Formulas for data-driven control: Stabilization, optimality, and robustness. *IEEE Transactions on Automatic Control*, 65(3), 909–924.

Eising, J. and Cortés, J. (2025). When sampling works in data-driven control: Informativity for stabilization in continuous time. *IEEE Transactions on Automatic Control*, 70(1), 565–572.

Engwerda, J. (2005). *LQ dynamic optimization and differential games*. John Wiley & Sons, West Sussex, England.

Engwerda, J. (2007). Algorithms for computing Nash equilibria in deterministic LQ games. *Computational Management Science*, 4(2), 113–140.

Lewis, F.L., Vrabie, D., and Syrmos, V.L. (2012). *Optimal Control*. John Wiley & Sons, New Jersey, USA, 3 edition.

Li, T.Y. and Gajic, Z. (1995). Lyapunov iterations for solving coupled algebraic Riccati equations of Nash differential games and algebraic Riccati equations of zero-sum games. In G.J. Olsder (ed.), *New Trends in Dynamic Games and Applications*, 333–351. Birkhäuser Boston, Boston, MA.

Lopez, V.G., Alsalti, M., and Müller, M.A. (2023). Efficient off-policy Q-learning for data-based discrete-time LQR problems. *IEEE Transactions on Automatic Control*, 68(5), 2922–2933.

Lopez, V.G. and Müller, M.A. (2022). On a continuous-time version of Willems' lemma. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, 2759–2764.

Lopez, V.G. and Müller, M.A. (2023). An efficient off-policy reinforcement learning algorithm for the continuous-time LQR problem. In *2023 62nd IEEE Conference on Decision and Control (CDC)*, 13–19.

Lopez, V.G. and Müller, M.A. (2026). Data-based control of continuous-time linear systems with performance specifications. *Automatica*, 189, 113002.

Markovsky, I. and Dörfler, F. (2023). Identifiability in the behavioral setting. *IEEE Transactions on Automatic Control*, 68(3), 1667–1677.

Nortmann, B., Monti, A., Sassano, M., and Mylvaganam, T. (2024). Nash equilibria for linear quadratic discrete-time dynamic games via iterative and data-driven algorithms. *IEEE Transactions on Automatic Control*, 69(10), 6561–6575.

Possieri, C. and Sassano, M. (2015). An algebraic geometry approach for the computation of all linear feedback Nash equilibria in LQ differential games. In *2015 54th IEEE Conference on Decision and Control (CDC)*, 5197–5202.

Rapisarda, P., van Waarde, H.J., and Çamlibel, M. (2024). Orthogonal polynomial bases for data-driven analysis and control of continuous-time systems. *IEEE Transactions on Automatic Control*, 69(7), 4307–4319.

Rhodes, I. and Luenberger, D. (1969). Stochastic differential games with constrained state estimators. *IEEE Transactions on Automatic Control*, 14(5), 476–481.

Saksena, V. and Cruz, J. (1982). A multimodel approach to stochastic Nash games. *Automatica*, 18(3), 295–305.

Song, R., Wei, Q., Zhang, H., and Lewis, F.L. (2021). Discrete-time non-zero-sum games with completely unknown dynamics. *IEEE Transactions on Cybernetics*, 51(6), 2929–2943.

Starr, A.W. and Ho, Y.C. (1969). Nonzero-sum differential games. *Journal of Optimization Theory and Applications*, 3, 184–206.

Turan, M.S. and Ferrari-Trecate, G. (2022). Data-driven unknown-input observers and state estimation. *IEEE Control Systems Letters*, 6, 1424–1429.

Vamvoudakis, K.G. (2015). Non-zero sum Nash Q-learning for unknown deterministic continuous-time linear systems. *Automatica*, 61, 274–281.

Willems, J.C., Rapisarda, P., Markovsky, I., and De Moor, B.L. (2005). A note on persistency of excitation. *Systems & Control Letters*, 54(4), 325–329.

Wolff, T.M., Lopez, V.G., and Müller, M.A. (2022). Data-based moving horizon estimation for linear discrete-time systems. In *2022 European Control Conference (ECC)*, 1778–1783.

Xie, K., Lu, M., Deng, F., Sun, J., and Chen, J. (2025). Data-driven dynamic output feedback Nash strategy for multi-player non-zero-sum games. *Journal of Systems Science and Complexity*, 38, 597–612.

Zhang, B.Q., Wang, B.C., and Cao, Y. (2024). An online Q-learning method for linear-quadratic nonzero-sum stochastic differential games with completely unknown dynamics. *Journal of Systems Science and Complexity*, 37, 1907–1922.