Mutually Quadratically Invariant Information Structures in Two-Team Stochastic Dynamic Games

Marcello Colombino[†], Roy S. Smith[†], and Tyler H. Summers[‡]

Abstract— **We formulate a two-team linear quadratic stochastic dynamic game featuring two opposing teams each with decentralized information structures. We introduce the concept of mutual quadratic invariance (MQI), which, analogously to quadratic invariance in (single team) decentralized control, defines a class of interacting information structures for the two teams under which optimal feedback control strategies are linear and easy to compute. We show that for zero-sum two-team dynamic games with MQI information structure, structured state feedback saddle-point equilibrium strategies can be computed from equivalent structured disturbance feedforward saddle point equilibrium strategies. We also show that there is a saddle-point equilibrium in linear strategies even when the teams are allowed to use nonlinear strategies. However, for nonzero-sum games we show via a counterexample that a similar equivalence fails to hold. The results are illustrated with a simple yet rich numerical example that illustrates the importance of the information structure for dynamic games.**

I. INTRODUCTION

Future cyber-physical systems (CPS) will feature cooperative networks of autonomous decision making agents equipped with embedded sensing, computation, communication, and actuation capabilities. These capabilities promise to significantly enhance performance, but also render the network vulnerable by increasing the number of access and influence points available to attackers. "Red team-blue team" scenarios, in which a defending team seeks to operate the network efficiently and securely while the attacking team seeks to disrupt network operation, have been used to qualitatively assess and improve security in military and intelligence organizations, but have not received formal mathematical analysis in a CPS context. Here we will study some fundamental properties in two-team stochastic dynamic games in cyber-physical networks, with a focus on interactions of the information structures of each team.

Dynamic game theory [1] offers a general framework for the study of optimal decision making in stochastic and non-cooperative environments. The theory can be viewed as a marriage of game theory [2], with a focus on interactions of multiple decision making agents, and optimal control theory [3], [4], with a focus on dynamics and feedback. The main elements of dynamic game theory are (1) a dynamical system along with a set of agents whose actions influence the state evolution of the system, (2) an objective function to be optimized associated with each agent, and (3) an information structure that specifies information sets for each agent, i.e., who knows what and when.

Several special classes of dynamic games have been extensively studied. In team decision theory all agents cooperate to optimize the same objective function. Static team theory traces back to the seminal work of Marschak and Radner [5]–[7]. Decentralized control theory has developed from team decision theory and control theory and introduces dynamics. The presence of dynamics makes available information depend on the actions of agents and significantly complicates the problem. Dynamic aspects were studied in important early work by Witsenhausen [8]–[11] and Ho [12], [13]. Witsenhausen's famous counterexample [8] vividly demonstrated the computational difficulties associated with team decision making in dynamic and stochastic environments. This still-unsolved counterexample described a simple team decision problem in which a nonlinear strategy strictly outperforms the optimal linear strategy and established deep connections between control, communication, and information theory. Research on decentralized and distributed control theory has continued, and there has been a recent resurgence of interest driven by the advent of large-scale cyber-physical networks. Recent work has elaborated on connections with communication and information theory [14], [15] and focused on computational and structural issues [16]–[21].

These important structural information aspects arising from cooperating agents are even more intricate in dynamic game settings [22] with non-cooperative and adversarial behavior, but have received much less attention in the literature. The most well studied case is the two-player problem, which features two opposing agents who have centralized information structures and has connections to robust control [23]. Our objective here is to study information structure aspects when there are *both* non-cooperative elements, as in general dynamic game theory, and cooperative elements, as in decentralized control theory. These aspects can be captured by a two-team stochastic dynamic game framework.

Two-team stochastic dynamic games feature two opposing teams with decentralized information structures for both the attacking and defending teams: each agent must act based on partial information measured or received locally in a way that coordinates its actions with team members and counters against the opposing team. This framework mathematically formalizes "red team-blue team" scenarios that qualitatively assess network security and resilience. In comparison to decentralized control theory, a team adversarial element is

[†]M. Colombino and R. Smith are with the Automatic Control Laboratory, ETH Zurich. [‡]T. Summers is with the Department of Mechanical Engineering, University of Texas at Dallas. E-mail addresses: {mcolombi, rsmith}@control.ee.ethz.ch, tyler.summers@utdallas.edu. This research is supported by the National Science Foundation under grant CNS-1566127 and partially supported by the Swiss National Science Foundation grant 2–773337–12.

added. In comparison to general dynamic game theory, a sharp contrast between cooperation with teammates and conflict against adversaries is preserved. Further, the stochastic element (modeled by a "chance" or "Nature" player in the game) allows the inclusion of random component failures and disturbance signals. There is currently a lack of deep theoretical and computational understanding in this class of dynamic games. A static version of the problem was studied in [24]. Many fundamental questions that been answered in static or single team decentralized control settings do not have counterparts in the two-team setting.

The main contributions of the present note are as follows. We formulate a two-team stochastic dynamic game problem and introduce a concept of mutual quadratic invariance (MQI), which defines a class of interacting information structures for the two teams under which optimal feedback control strategies are are linear and easy to compute. This is analogous to the concept of quadratic invariance in (single team) decentralized control [16]; however, our concept is distinct and is not equivalent to quadratic invariance for each team individually. We show that for zero-sum two-team dynamic games with MQI information structure, structured state feedback saddle point equilibrium strategies can be computed from equivalent structured disturbance feedforward saddle point equilibrium strategies. We also show that there is a saddle-point equilibrium in linear strategies even when the teams are allowed to use nonlinear strategies. However, for nonzero-sum games we show via a counterexample that a similar equivalence fails to hold for structured Nash equilibrium strategies. Finally, we present a numerical example, which illustrates the importance of the information structure on the value of the game.

The rest of the note is structured as follows. Section II provides preliminaries on static team games. Section III formulates a two-team stochastic dynamic game. Sections IV and V develop results on disturbance feedforward and state feedback strategies and introduce the concept of mutual quadratic invariance. Section VI presents illustrative numerical experiments, and Section VII gives some concluding remarks and future research directions.

II. Two-team stochastic static games

In this section we review basic results for a two-team stochastic static game. In this setting, two teams who both know the distribution parameters of a Gaussian random vector w need to decide strategies to compute vectors $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^q$ as a function of the realization $w \in \mathbb{R}^n$ in order to minimize the expectation of *different* quadratic forms in w, u, v. Each team is composed of multiple agents, each of which observes a different linear function of w and decides a portion of the vectors u or v.

More formally, given a vector $w \sim \mathcal{N}(m_w, \Sigma_w)$, where $\Sigma_w \succ 0$, consider the following game

$$T_1 : \begin{cases} \min\limits_{\kappa_i(\cdot)} \mathbb{E}_w \left(J_1(w,u,v)\right) \\ \text{s. t. } \; u_i = \kappa_i(C_i w) \\ \qquad \forall i \in \mathbb{Z}_{[1,N]} \end{cases}, \quad T_2 : \begin{cases} \min\limits_{\lambda_j(\cdot)} \mathbb{E}_w \left(J_2(w,u,v)\right) \\ \text{s. t. } \; v_j = \lambda_j(\Gamma_i w) \\ \qquad \forall j \in \mathbb{Z}_{[1,M]} \end{cases} \tag{1}$$

where for $i \in \{1,2\}$, $J_i(w,u,v) :=$

$$\begin{bmatrix} w \\ u \\ v \end{bmatrix}^\top \begin{bmatrix} \mathcal{H}_{i\,ww} & \mathcal{H}_{i\,wu} & \mathcal{H}_{i\,wv} \\ \mathcal{H}_{i\,wu}^\top & \mathcal{H}_{i\,uu} & \mathcal{H}_{i\,uv} \\ \mathcal{H}_{i\,wv}^\top & \mathcal{H}_{i\,uv}^\top & \mathcal{H}_{i\,vv} \end{bmatrix} \begin{bmatrix} w \\ u \\ v \end{bmatrix},$$

are the objective functions of each team, and $\kappa_i(\cdot), i \in \mathbb{Z}_{[1,N]}$ and $\lambda_j(\cdot), j \in \mathbb{Z}_{[1,M]}$ are Borel measurable functions corresponding to the decision strategies of agents on team 1 and 2, respectively.

Assumption 1: We assume

$$\begin{bmatrix} \mathcal{H}_{1\,uu} & \mathcal{H}_{1\,uv} \\ \mathcal{H}_{1\,uv}^\top & \mathcal{H}_{2\,vv} \end{bmatrix} \succ 0, \begin{bmatrix} \mathcal{H}_{1\,uu} & \mathcal{H}_{2\,uv} \\ \mathcal{H}_{2\,uv}^\top & \mathcal{H}_{2\,vv} \end{bmatrix} \succ 0.$$

Note that in the zero-sum case $(J_1 = -J_2)$, Assumption 1 is standard to guarantee the existence of a saddle point equilibrium to the game without decentralized information structure [25, condition 6.3.9]. If $J_1 = J_2$, Assumption 1 reduces to the standard positive definite assumption of team theory [6].

We now define the set of Nash optimal strategies for the game in (1).

Definition 1: A pair of strategies $(\kappa^\star(\cdot), \lambda^\star(\cdot))$ of the form $[\kappa_1^{\star\top}(C_1\,\cdot), \ldots, \kappa_N^{\star\top}(C_n\,\cdot)]^\top$ and $[\lambda_1^{\star\top}(\Gamma_1\,\cdot), \ldots, \lambda_M^{\star\top}(\Gamma_n\,\cdot)]^\top$ is Nash optimal for the game in (1) if

$$\begin{cases} \kappa^\star(\cdot) \in \arg\min\limits_{\kappa(\cdot)} \mathbb{E}_w J_1(w, \lambda^\star(w), \kappa(w)) \\ \lambda^\star(\cdot) \in \arg\min\limits_{\lambda(\cdot)} \mathbb{E}_w J_2(w, \lambda(w), \kappa^\star(w)). \end{cases} \tag{2}$$

Under Assumption 1, the game in (1) admits a unique set of linear Nash optimal strategies, which can be computed by solving a set of linear equations derived from stationarity conditions [24]. This turns out to be a special case of a general multi-player, multi-objective linear quadratic static game considered in [26].

III. Two-team stochastic dynamic games

Problem (1) is a static game. There is no concept of time and causality of the information pattern. In this section we formulate a dynamic game, where two teams can influence the state evolution of a dynamical system. The agents on each team decide a portion of an input signal based on different observations of the system state over time. Decision must be causal: each player is only allowed to use past or, at most, present information.

Our focus will be on the role of information structures for both teams in determining equilibrium strategies. Dynamic games offer a rich variety of information structures. Specific instances have been considered in the literature, with much work on various types of centralized structures [1] and some work on structures defined by spatiotemporal decentralization patterns. For example, a one-step-delay observation sharing pattern was shown in [26] to admit unique linear optimal strategies. There has been recent progress in (single team) decentralized control on information structure issues, including a characterization of information structures called quadratically invariant that yield convex control design problems [16]. Here we seek an analogous result in a two-team game setting.

Consider the system

$$x(t+1) = Ax(t) + B_1u(t) + B_2v(t) + w(t), \tag{3}$$

where $x(t) \in \mathbb{R}^n$ is the system state at time t with $x(0) \sim \mathcal{N}(0, \Sigma_0)$, $u(t) \in \mathbb{R}^{m_1}$ is the input for team 1 at time t, $v(t) \in \mathbb{R}^{m_2}$ is the input for team 2 at time t, and $w(t) \sim \mathcal{N}(0, \Sigma_t)$ is a random disturbance. The cost functions for each team are given by

$$J_i := \mathbb{E}\left(\sum_{t=0}^{N-1} x(t)^\top M_i(t)x(t) + u(t)^\top R_i(t)u(t) + v(t)^\top V_i(t)v(t)\right) + x(N)^\top M_i(N)x(N), \quad i \in \{1,2\}, \tag{4}$$

where $M_i(0) = \mathbf{0}_{n\times n}$, $M_i(t) = M_i(t)^\top \in \mathbb{R}^{n\times n}$, $R_i(t) = R_i(t)^\top \in \mathbb{R}^{m_1\times m_1}$ and $V_i(t) = V_i(t)^\top \in \mathbb{R}^{m_2\times m_2}$. By defining the matrices $\mathcal{A} = \text{blockdiag}(A,\ldots,A) \in \mathbb{R}^{n(N+1)\times n(N+1)}$,

$$\mathcal{B}_i = \begin{bmatrix} B_i & 0 & 0 \\ 0 & \ddots & 0 \\ \vdots & & B_i \\ 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{n(N+1)\times m_j N}, \quad i \in \{1,2\},$$

$\mathcal{M}_i = \text{blockdiag}(0, M_i(1),\ldots,M_i(N)) \in \mathbb{R}^{n(N+1)\times n(N+1)}$, $\mathcal{R}_i = \text{blockdiag}(R_i(0),\ldots,R_i(N-1)) \in \mathbb{R}^{m_1N\times m_1N}$ and $\mathcal{V}_i = \text{blockdiag}(V_i(0),\ldots,V_i(N-1)) \in \mathbb{R}^{m_2N\times M_2N}$ for $i \in \{1,2\}$, the vectors $\mathbf{x} = (x(0),...,x(N)) \in \mathbb{R}^{n(N+1)}$, $\mathbf{u} = (u(0),...,u(N-1)) \in \mathbb{R}^{m_1N}$, $\mathbf{v} = (v(0),...,v(N-1)) \in \mathbb{R}^{m_2N}$ and $\mathbf{w} = (x(0),w(0),...,w(N-1)) \in \mathbb{R}^{n(N+1)}$, and the shift matrix

$$\mathcal{Z} := \begin{bmatrix} 0 & & & \\ I & \ddots & & \\ & \ddots & \ddots & \\ & & I & 0 \end{bmatrix} \in \mathbb{R}^{n(N+1)\times n(N+1)},$$

we can write system (3) as $\mathbf{x} = \mathcal{Z}\mathcal{A}\mathbf{x} + \mathcal{Z}\mathcal{B}_1\mathbf{u} + \mathcal{Z}\mathcal{B}_2\mathbf{v} + \mathbf{w}$. The system can then be rewritten compactly as

$$\mathbf{x} = \begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} & \mathcal{P}_{13} \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{u} \\ \mathbf{v} \end{bmatrix}, \tag{5}$$

where $\mathcal{P}_{11} = (I - \mathcal{Z}\mathcal{A})^{-1}$, $\mathcal{P}_{12} = (I - \mathcal{Z}\mathcal{A})^{-1}\mathcal{Z}\mathcal{B}_1$ and $\mathcal{P}_{13} = (I - \mathcal{Z}\mathcal{A})^{-1}\mathcal{Z}\mathcal{B}_2$. The cost functions in (4) can be written in function of the vectorized inputs as

$$\mathbf{J}_i(\mathbf{u},\mathbf{v}) = \mathbb{E}_\mathbf{w}\left(\begin{bmatrix} \mathbf{w} \\ \mathbf{u} \\ \mathbf{v} \end{bmatrix}^\top \mathcal{H}_i \begin{bmatrix} \mathbf{w} \\ \mathbf{u} \\ \mathbf{v} \end{bmatrix}\right), \quad i \in \{1,2\},$$

where for $i = \{1,2\}$, $\mathcal{H}_i =$

$$\begin{bmatrix} \mathcal{P}_{11}^\top\mathcal{M}_i\mathcal{P}_{11} & \mathcal{P}_{11}^\top\mathcal{M}_i\mathcal{P}_{12} & \mathcal{P}_{11}^\top\mathcal{M}_i\mathcal{P}_{13} \\ \mathcal{P}_{12}^\top\mathcal{M}_i\mathcal{P}_{11} & \mathcal{P}_{12}^\top\mathcal{M}_i\mathcal{P}_{12} + \mathcal{R}_i & P_{12}^\top\mathcal{M}_i\mathcal{P}_{13} \\ \mathcal{P}_{13}^\top\mathcal{M}_i\mathcal{P}_{11} & \mathcal{P}_{13}^\top\mathcal{M}_i\mathcal{P}_{12} & \mathcal{P}_{13}^\top\mathcal{M}_i\mathcal{P}_{13} + \mathcal{V}_i \end{bmatrix}.$$

We are interested in the finite horizon, two-team stochastic dynamic game where Team 1 minimizes $\mathbf{J}_1(\mathbf{u},\mathbf{v})$, Team 2 minimizes $\mathbf{J}_2(\mathbf{u},\mathbf{v})$; and each team chooses a structured causal state feedback strategy of the form

$$\mathbf{u} = \mathcal{K}_1(\mathbf{x}), \quad \mathbf{v} = \mathcal{K}_2(\mathbf{x}). \quad \mathcal{K}_1 \in \mathcal{S}_1,\ \mathcal{K}_2 \in \mathcal{S}_2,$$

where $K_i : \mathbb{R}^{n(N+1)} \to \mathbb{R}^{m_iN}$, for $i \in \{1,2\}$ are measurable functions and $\mathcal{S}_1$ and $\mathcal{S}_2$ define an information structure for each team.

We define an information structure as a binary matrix $\mathcal{S}_i \in \{0,1\}^{n(N+1)\times m_iN}$, where $\mathcal{K}_i \in \mathcal{S}_i$ indicates that, if $[\mathcal{S}_i]_{jk} = 0$, then the j^{th} element of $\mathcal{K}_i$ is not a function of $\mathbf{x}_k$. By choosing the information structures one can enforce causality and a prescribed spatiotemporal structure on the strategies.

IV. Mutual Quadratic Invariance

In decentralized control with quadratically invariant information structures, the controller structure can be enforced on an affine parameter that defines the achievable set of closed-loop systems and recover a structured feedback controller. We now follow a similar approach in the two-team setting.

A. Disturbance feedforward strategies

By searching for measurable disturbance feedforward strategies of the type $\mathbf{u} = \mathcal{Q}_1(\mathcal{P}_{11}\mathbf{w})$ and $\mathbf{v} = \mathcal{Q}_2(\mathcal{P}_{11}\mathbf{w})$, where $\mathcal{Q}_1 \in \mathcal{S}_1$ and $\mathcal{Q}_2 \in \mathcal{S}_2$, we recover the formulation of (1). Provided that Assumption 1 is satisfied, there exists a unique Nash equilibrium of the form $\mathbf{u} = \bar{\mathcal{Q}}_1\mathcal{P}_{11}\mathbf{w},\ \mathbf{v} = \bar{\mathcal{Q}}_2\mathcal{P}_{11}\mathbf{w}$ in the space of linear strategies [24]. The matrices $\bar{\mathcal{Q}}_1, \bar{\mathcal{Q}}_2$ can be easily computed by solving a linear system of equations or a sequence of semidefinite programs [24], [26]. Assumption 1 for the dynamic game problem becomes

$$\begin{bmatrix} \mathcal{P}_{12}^\top\mathcal{M}_1\mathcal{P}_{12} + \mathcal{R}_1 & \mathcal{P}_{12}^\top\mathcal{M}_1\mathcal{P}_{13} \\ \mathcal{P}_{13}^\top\mathcal{M}_1\mathcal{P}_{12} & \mathcal{P}_{13}^\top\mathcal{M}_2\mathcal{P}_{13} + \mathcal{V}_2 \end{bmatrix} \succ 0,$$

$$\begin{bmatrix} \mathcal{P}_{12}^\top\mathcal{M}_2\mathcal{P}_{12} + \mathcal{R}_2 & \mathcal{P}_{12}^\top\mathcal{M}_2\mathcal{P}_{13} \\ \mathcal{P}_{13}^\top\mathcal{M}_2\mathcal{P}_{12} & \mathcal{P}_{13}^\top\mathcal{M}_1\mathcal{P}_{13} + \mathcal{V}_1 \end{bmatrix} \succ 0,$$

where We can define new cost functions that depend on the matrices $\mathcal{Q}_1$ and $\mathcal{Q}_2$ describing linear disturbance feedforward strategies as

$$\mathcal{J}_i\left(\begin{bmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \end{bmatrix}\right) = \mathbf{J}_i(\mathbf{u},\mathbf{v})\Big|_{\mathbf{u}=\mathcal{Q}_1\mathcal{P}_{11}\mathbf{w},\mathbf{v}=\mathcal{Q}_2\mathcal{P}_{11}\mathbf{w}}. \tag{6}$$

In particular,

$$\mathcal{J}_i\left(\begin{bmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \end{bmatrix}\right) = \|\mathcal{M}_i^{\frac{1}{2}}\left(I + \mathcal{P}_{12}\mathcal{Q}_1 + \mathcal{P}_{13}\mathcal{Q}_2\right)\mathcal{P}_{11}\Sigma_\mathbf{w}^{\frac{1}{2}}\|_F^2 + \|\mathcal{R}_i^{\frac{1}{2}}\mathcal{Q}_1\Sigma_\mathbf{w}^{\frac{1}{2}}\|_F^2 + \|\mathcal{V}_i^{\frac{1}{2}}\mathcal{Q}_2\Sigma_\mathbf{w}^{\frac{1}{2}}\|_F^2,$$

where $\Sigma_\mathbf{w}^{\frac{1}{2}}$ is the covariance of $\mathbf{w}$.

B. Equivalent state feedback strategies

It is easy to show that there exists a bijective relationship between a pair of linear disturbance feedforward strategies $(\mathcal{Q}_1, \mathcal{Q}_2)$ and an equivalent pair of linear state feedback strategies described by the matrices $(\mathcal{K}_1, \mathcal{K}_2)$. More precisely, using (5) we obtain

$$\begin{aligned} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} &= \begin{bmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \end{bmatrix} \mathcal{P}_{11}\mathbf{w} \\ &= \begin{bmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \end{bmatrix} \begin{bmatrix} \mathcal{P}_{12} & \mathcal{P}_{13} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}. \end{aligned} \tag{7}$$

We can define the function g such that $\begin{bmatrix} \mathcal{K}_1 \\ \mathcal{K}_2 \end{bmatrix} = g\left(\begin{bmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \end{bmatrix}\right)$, where

$$g\left(\begin{bmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \end{bmatrix}\right) = \left(I + \begin{bmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \end{bmatrix} \begin{bmatrix} \mathcal{P}_{12} & \mathcal{P}_{13} \end{bmatrix}\right)^{-1} \begin{bmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \end{bmatrix}.$$

Using a similar approach to (7), one can construct the inverse mapping that, given a pair of feedback strategies, recovers the equivalent feedforward strategies.

$$\begin{bmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \end{bmatrix} = g^{-1}\left(\begin{bmatrix} \mathcal{K}_1 \\ \mathcal{K}_2 \end{bmatrix}\right),$$

where the map g^{-1} takes the form

$$g^{-1}\left(\begin{bmatrix} \mathcal{K}_1 \\ \mathcal{K}_2 \end{bmatrix}\right) = \begin{bmatrix} \mathcal{K}_1 \\ \mathcal{K}_2 \end{bmatrix} \left(I - \begin{bmatrix} \mathcal{P}_{12} & \mathcal{P}_{13} \end{bmatrix} \begin{bmatrix} \mathcal{K}_1 \\ \mathcal{K}_2 \end{bmatrix}\right)^{-1}.$$

The inverses in the expressions exist due to the lower triangular structure of the matrices involved, so that computing them is a matter of simple forward substitution. Given a pair of linear feedback strategies $\mathcal{K}_1$ and $\mathcal{K}_2$, the cost for player i can be evaluated by considering the equivalent feedforward strategies as

$$\mathbf{J}_i(\mathbf{u}, \mathbf{v})\Big|_{\mathbf{u}=\mathcal{K}_1\mathbf{x}, \mathbf{v}=\mathcal{K}_2\mathbf{x}} = \mathcal{J}_i\left(g^{-1}\left(\begin{bmatrix} \mathcal{K}_1 \\ \mathcal{K}_2 \end{bmatrix}\right)\right),$$

where $\mathcal{J}_i$ is defined in (6).

Now that we have a way to construct feedback strategies which are equivalent to any set of linear feedforward strategies, we need to establish a condition that guarantees that such equivalent feedback strategies will preserve the desired structure.

C. Mutual Quadratic Invariance

We know form the quadratic invariance literature [16], [27] that $\mathcal{Q}_1 \in \mathcal{S}_1$ and $\mathcal{Q}_2 \in \mathcal{S}_2 \iff \mathcal{K}_1 \in \mathcal{S}_1$ and $\mathcal{K}_2 \in \mathcal{S}_2$ if and only if for all $(\mathcal{K}_1, \mathcal{K}_2) \in \mathcal{S}_1 \times \mathcal{S}_2$ it holds that

$$\begin{bmatrix} \mathcal{K}_1 \\ \mathcal{K}_2 \end{bmatrix} \begin{bmatrix} \mathcal{P}_{12} & \mathcal{P}_{13} \end{bmatrix} \begin{bmatrix} \mathcal{K}_1 \\ \mathcal{K}_2 \end{bmatrix} \in \mathcal{S}_1 \times \mathcal{S}_2, \tag{8}$$

in other words $\mathcal{S}_1 \times \mathcal{S}_2$ is quadratically invariant under $\begin{bmatrix} \mathcal{P}_{12} & \mathcal{P}_{13} \end{bmatrix}$. We define this property as **mutual quadratic invariance**. We can expand (8) as

$$\begin{bmatrix} \mathcal{K}_1 \\ \mathcal{K}_2 \end{bmatrix} \in \mathcal{S}_1 \times \mathcal{S}_2 \implies \begin{cases} \mathcal{K}_1\mathcal{P}_{12}\mathcal{K}_1, & \mathcal{K}_1\mathcal{P}_{13}\mathcal{K}_2 \in \mathcal{S}_1 \\ \mathcal{K}_2\mathcal{P}_{12}\mathcal{K}_1, & \mathcal{K}_2\mathcal{P}_{13}\mathcal{K}_2 \in \mathcal{S}_2. \end{cases} \tag{9}$$

By observing (8) we note that MQI is equivalent to QI for a control problem where both decisions $\mathbf{u}$ and $\mathbf{v}$ are taken by a single team. However, we emphasize that MQI is not equivalent to QI for each team individually; note that the cross term structural constraints $\mathcal{K}_1\mathcal{P}_{13}\mathcal{K}_2 \in \mathcal{S}_1$ and $\mathcal{K}_2\mathcal{P}_{12}\mathcal{K}_1 \in \mathcal{S}_2$ require availability of certain information that depends on the structure of how the team decisions interact through the system dynamics. Whereas QI precludes signaling incentives, in which decision making agents communicate to other team members through the system dynamics via their control actions, MQI also precludes obfuscation or deception incentives, in which decision making agents attempt to confound opposing team members through the system dynamics via their control actions. MQI information structures will allow us to compute equilibrium strategies in two-team games.

V. Computing Equilibrium Strategies

We have seen in Section IV that given structures $\mathcal{S}_1$ and $\mathcal{S}_2$ which are mutually quadratically invariant under $\mathcal{P}_{12}$ and $\mathcal{P}_{13}$, one can easily obtain a Nash equilibrium in the disturbance feedforward strategies. Furthermore once we recover the equivalent state feedback strategies the structure is preserved. There is still a nontrivial question that needs to be answered.

Problem 1: Given a Nash equilibrium in the feedforward strategies $(\bar{\mathcal{Q}}_1, \bar{\mathcal{Q}}_2)$, are the equivalent feedback strategies

$$\left(\begin{bmatrix} \bar{\mathcal{K}}_1 \\ \bar{\mathcal{K}}_2 \end{bmatrix}\right) = g\left(\begin{bmatrix} \bar{\mathcal{Q}}_1 \\ \bar{\mathcal{Q}}_2 \end{bmatrix}\right), \tag{10}$$

a Nash equilibrium in the feedback strategies?

In order to understand why the answer to Problem 1 is nontrivial, we first present a counterexample for a nonzero sum game.

A. Nonzero sum game

Consider the following problem instance with

$$\begin{aligned} &N = 2, \quad A = 2, \quad B_1 = 0.4, \quad B_2 = 0.1, \\ &\Sigma_0 = 1, \quad \Sigma_t = 1\ \forall t, \\ &M_1(t) = R_1(t) = V_1(t) = 1\ \forall t, \\ &M_2(t) = 70\ \forall t, R_2(t) = V_2(t) = 1\ \forall t. \end{aligned} \tag{11}$$

This is a single state, two-stage, two-player problem, and we will consider centralized, causal strategies, which are readily verified to be mutually quadratically invariant. Using disturbance feedforward strategies, the problem can be reduced to a static game whose unique Nash Equilibrium strategies are readily computed using methods from [24], [26]. This yields the Nash pair $(\bar{\mathcal{Q}}_1, \bar{\mathcal{Q}}_2)$, where

$$\begin{aligned} \bar{\mathcal{Q}}_1 &= \begin{bmatrix} -0.6795 & 0 & 0 \\ 0.6283 & -0.4301 & 0 \end{bmatrix}, \\ \bar{\mathcal{Q}}_2 &= \begin{bmatrix} -11.890 & 0 & 0 \\ 10.996 & -7.5269 & 0 \end{bmatrix}. \end{aligned} \tag{12}$$

The corresponding equilibrium value for player 1 is $\mathcal{J}_1^*\left(\begin{bmatrix}\bar{\mathcal{Q}}_1\\ \bar{\mathcal{Q}}_2\end{bmatrix}\right) = 220$. The corresponding state feedback strategies are

$$\begin{aligned}\bar{\mathcal{K}}_1 &= \begin{bmatrix} -0.6795 & 0 & 0 \\ 0 & -0.4301 & 0 \end{bmatrix},\\ \bar{\mathcal{K}}_2 &= \begin{bmatrix} -11.890 & 0 & 0 \\ 0 & -7.5269 & 0 \end{bmatrix},\end{aligned} \tag{13}$$

However, $(\bar{\mathcal{K}}_1, \bar{\mathcal{K}}_2)$ is not a Nash Equilibrium in state feedback strategies since it is readily verified that $\mathcal{J}_1\left(g^{-1}\left(\begin{bmatrix}\hat{\mathcal{K}}_1\\ \bar{\mathcal{K}}_2\end{bmatrix}\right)\right) = 206.1$, where

$$\hat{\mathcal{K}}_1 = \begin{bmatrix} -1.853 & 0 & 0 \\ 0 & -0.4301 & 0 \end{bmatrix}. \tag{14}$$

Although the sparsity structure is preserved between the disturbance feedforward and state feedback strategies since the information structure is mutually quadratically invariant, the Nash Equilibrium property is not preserved. Thus, for this particular non-zero sum dynamic game, the answer to the question posed in Problem 1 is negative. Note the function g in (10) couples $\bar{\mathcal{K}}_2, \bar{\mathcal{Q}}_1$ and $\bar{\mathcal{K}}_1, \bar{\mathcal{Q}}_2$. This means that the unilateral change in the feedback policy from $\bar{\mathcal{K}}_1$ to $\hat{\mathcal{K}}_1$ would require a bilateral change of both $\bar{\mathcal{Q}}_1$ and $\bar{\mathcal{Q}}_2$ to be implemented as a disturbance feedforward, thus losing all guarantees provided by the Nash equilibrium property in feedforward strategies. We show next that this situation does not occur in zero-sum games.

B. Zero sum game

Let us now consider the zero sum game case, where the objective function of one team is precisely the negative of that of the other team. Such problems can be re-written as a min-max problem of the form.

$$\min_{\mathbf{u}} \max_{\mathbf{v}} \quad \mathbf{J}(\mathbf{u}, \mathbf{v}) := \mathbb{E}_{\mathbf{w}}\left(\begin{bmatrix}\mathbf{w}\\ \mathbf{u}\\ \mathbf{v}\end{bmatrix}^\top \mathcal{H} \begin{bmatrix}\mathbf{w}\\ \mathbf{u}\\ \mathbf{v}\end{bmatrix}\right) \tag{15}$$

As before we are interested in strategies of the form

$$\mathbf{u} = \mathcal{K}_1(\mathbf{x}), \quad \mathbf{v} = \mathcal{K}_2(\mathbf{x}). \quad \mathcal{K}_1 \in \mathcal{S}_1,\ \mathcal{K}_2 \in \mathcal{S}_2,$$

where $\mathcal{S}_1$ and $\mathcal{S}_2$ are prescribed sets of structured causal controllers. As zero-sum games are a special case of nonzero sum games, provided Assumption 1 holds, we can find structured Nash equilibria (rather saddle point equilibrium in the zero sum context) if we consider linear strategies of the form $\mathbf{u} = \mathcal{Q}_1\mathcal{P}_{11}\mathbf{w}$ and $\mathbf{v} = \mathcal{Q}_2\mathcal{P}_{11}\mathbf{w}$. For the zero sum game of the form (15), Assumption 1 reads

$$\begin{bmatrix} \mathcal{P}_{12}^\top\mathcal{M}\mathcal{P}_{12} + \mathcal{R} & \mathcal{P}_{12}^\top\mathcal{M}\mathcal{P}_{13} \\ \mathcal{P}_{13}^\top\mathcal{M}\mathcal{P}_{12} & -\left(\mathcal{P}_{13}^\top\mathcal{M}\mathcal{P}_{13} + \mathcal{S}\right)\end{bmatrix} \succ 0.$$

As before, given the saddle point equilibrium in the feedforward strategies, when $\mathcal{S}_1$ and $\mathcal{S}_2$ are mutually quadratically invariant, we can compute equivalent linear feedback strategies that preserve the structure. In the zero-sum case, however, we can relate the equilibrium property of the feedforward strategies to that of the state feedback strategies.

We start with a result that shows that for a zero sum game, the maps g and g^{-1} preserve stationary points.

Lemma 1: Let $\mathcal{J}_1 = \alpha\mathcal{J}_2$, with $\alpha \in \mathbb{R}$ and

$$\begin{bmatrix}\bar{\mathcal{Q}}_1\\ \bar{\mathcal{Q}}_2\end{bmatrix} = g^{-1}\left(\begin{bmatrix}\bar{\mathcal{K}}_1\\ \bar{\mathcal{K}}_2\end{bmatrix}\right).$$

The following holds true

$$\begin{aligned}\frac{\partial \mathcal{J}_1\left(\begin{bmatrix}\mathcal{Q}_1\\ \mathcal{Q}_2\end{bmatrix}\right)}{\partial \mathcal{Q}_1}\Bigg|_{\mathcal{Q}_1=\bar{\mathcal{Q}}_1, \mathcal{Q}_2=\bar{\mathcal{Q}}_2} &\in \mathcal{S}_1^\perp,\\ \frac{\partial \mathcal{J}_2\left(\begin{bmatrix}\mathcal{Q}_1\\ \mathcal{Q}_2\end{bmatrix}\right)}{\partial \mathcal{Q}_2}\Bigg|_{\mathcal{Q}_1=\bar{\mathcal{Q}}_1, \mathcal{Q}_2=\bar{\mathcal{Q}}_2} &\in \mathcal{S}_2^\perp,\end{aligned} \tag{16}$$

if and only if

$$\begin{aligned}\frac{\partial \mathcal{J}_1\left(g^{-1}\left(\begin{bmatrix}\mathcal{K}_1\\ \mathcal{K}_2\end{bmatrix}\right)\right)}{\partial \mathcal{K}_1}\Bigg|_{\mathcal{K}_1=\bar{\mathcal{K}}_1, \mathcal{K}_2=\bar{\mathcal{K}}_2} &\in \mathcal{S}_1^\perp,\\ \frac{\partial \mathcal{J}_2\left(g^{-1}\left(\begin{bmatrix}\mathcal{K}_1\\ \mathcal{K}_2\end{bmatrix}\right)\right)}{\partial \mathcal{K}_2}\Bigg|_{\mathcal{K}_1=\bar{\mathcal{K}}_1, \mathcal{K}_2=\bar{\mathcal{K}}_2} &\in \mathcal{S}_2^\perp,\end{aligned} \tag{17}$$

Proof: Let us simplify the notation and define

$$\mathcal{Q} := \begin{bmatrix}\mathcal{Q}_1\\ \mathcal{Q}_2\end{bmatrix}, \quad \mathcal{K} := \begin{bmatrix}\mathcal{K}_1\\ \mathcal{K}_2\end{bmatrix}, \quad \mathcal{P} := \begin{bmatrix}\mathcal{P}_{12} & \mathcal{P}_{13}\end{bmatrix}.$$

We start by proving the *only if* part. Note that since $\mathcal{J}_1 = \alpha\mathcal{J}_2$,

$$\frac{\partial \mathcal{J}_2\left(\begin{bmatrix}\mathcal{Q}_1\\ \mathcal{Q}_2\end{bmatrix}\right)}{\partial \mathcal{Q}_2}\Bigg|_{\mathcal{Q}_1=\bar{\mathcal{Q}}_1, \mathcal{Q}_2=\bar{\mathcal{Q}}_2} \in \mathcal{S}_2^\perp$$

if and only if

$$\frac{\partial \mathcal{J}_1\left(\begin{bmatrix}\mathcal{Q}_1\\ \mathcal{Q}_2\end{bmatrix}\right)}{\partial \mathcal{Q}_2}\Bigg|_{\mathcal{Q}_1=\bar{\mathcal{Q}}_1, \mathcal{Q}_2=\bar{\mathcal{Q}}_2} \in \mathcal{S}_2^\perp.$$

Assume (17) holds, and suppose (16) does not hold. Then there exist $\tilde{\mathcal{Q}} \in \mathcal{S}_1 \times \mathcal{S}_2$, with $\tilde{\mathcal{Q}} \neq 0$ such that

$$\lim_{\varepsilon\to 0}\frac{\mathcal{J}_1(\bar{\mathcal{Q}} + \varepsilon\tilde{\mathcal{Q}}) - \mathcal{J}_1(\bar{\mathcal{Q}})}{\varepsilon} = \kappa \neq 0,$$

or equivalently

$$\lim_{\varepsilon\to 0}\frac{\mathcal{J}_1\left(g^{-1}\left(g\left(\bar{\mathcal{Q}} + \varepsilon\tilde{\mathcal{Q}}\right)\right)\right) - \mathcal{J}_1\left(\bar{\mathcal{Q}}\right)}{\varepsilon} = \kappa \neq 0. \tag{18}$$

We know that

$$\begin{aligned} g\left(\bar{\mathcal{Q}} + \varepsilon\tilde{\mathcal{Q}}\right) &= \left(I + \left(\bar{\mathcal{Q}} + \varepsilon\tilde{\mathcal{Q}}\right)\mathcal{P}\right)^{-1}\left(\bar{\mathcal{Q}} + \varepsilon\tilde{\mathcal{Q}}\right)\\ &= \left(I + \bar{\mathcal{Q}}\mathcal{P} + \varepsilon\tilde{\mathcal{Q}}\mathcal{P}\right)^{-1}\left(\bar{\mathcal{Q}} + \varepsilon\tilde{\mathcal{Q}}\right)\\ &= \Big[(I + \bar{\mathcal{Q}}\mathcal{P})^{-1} + \varepsilon(I + \bar{\mathcal{Q}}\mathcal{P})^{-1}\tilde{\mathcal{Q}}\mathcal{P}(I + \bar{\mathcal{Q}}\mathcal{P})^{-1}\\ &\quad + \mathcal{O}(\varepsilon^2)\Big]\left(\bar{\mathcal{Q}} + \varepsilon\tilde{\mathcal{Q}}\right) = \bar{\mathcal{K}} + \varepsilon\tilde{\mathcal{K}} + \mathcal{O}(\varepsilon^2),\end{aligned} \tag{19}$$

where

$$\tilde{\mathcal{K}} = (I + \bar{\mathcal{Q}}\mathcal{P})^{-1}\tilde{\mathcal{Q}} + (I + \bar{\mathcal{Q}}\mathcal{P})^{-1}\tilde{\mathcal{Q}}\mathcal{P}(I + \bar{\mathcal{Q}}\mathcal{P})^{-1}\bar{\mathcal{Q}}.$$

Using [16, Theorem 14 + Theorem 26], and mutual quadratic invariance it is easy to conclude that, $\tilde{\mathcal{K}} \in \mathcal{S}_1 \times \mathcal{S}_2$. Substituting (19) in (18) and using the fact that $\bar{\mathcal{Q}} = g\left(\bar{\mathcal{K}}\right)$ one obtains

$$\lim_{\varepsilon \to 0} \frac{\mathcal{J}_1\left(g^{-1}\left(\bar{\mathcal{K}} + \varepsilon\tilde{\mathcal{K}} + \mathcal{O}(\varepsilon^2)\right)\right) - \mathcal{J}_1\left(g^{-1}\left(\bar{\mathcal{K}}\right)\right)}{\varepsilon} =$$
$$\lim_{\varepsilon \to 0} \frac{\mathcal{J}_1\left(g^{-1}\left(\bar{\mathcal{K}} + \varepsilon\tilde{\mathcal{K}}\right)\right) - \mathcal{J}_1\left(g^{-1}\left(\bar{\mathcal{K}}\right)\right)}{\varepsilon} = \kappa \neq 0,$$

which, since $\tilde{\mathcal{K}} \in \mathcal{S}_1 \times \mathcal{S}_2$, is in contradiction with (17) and thus proves the claim.

The converse direction can be proven analogously. ∎

Remark 1: Note that Lemma 1 does not hold for general games as the proof relies to the fact that $\mathcal{J}_1 = \alpha\mathcal{J}_2$. In this work we are interested in the case in which $\mathcal{J}_1 = -\mathcal{J}_2$ and we have a zero-sum game. For the case $\mathcal{J}_1 = \mathcal{J}_2$ we recover well-known results from the literature on decentralized control with QI information structure.

We are now ready to state our main result, which allows us to construct a saddle point Equilibrium in the space of linear feedback strategies.

Theorem 2: Let $(\bar{\mathcal{Q}}_1, \bar{\mathcal{Q}}_2)$ be the unique saddle point equilibrium in the disturbance feedforward strategies. Then if there exists a saddle point equilibrium in state feedback strategies, it is unique and given by

$$\begin{bmatrix} \bar{\mathcal{K}}_1 \\ \bar{\mathcal{K}}_2 \end{bmatrix} = g\left(\begin{bmatrix} \bar{\mathcal{Q}}_1 \\ \bar{\mathcal{Q}}_2 \end{bmatrix}\right).$$

Proof: Let $(\hat{\mathcal{K}}_1, \hat{\mathcal{K}}_2)$ be any saddle point equilibrium in state-feedback linear strategies of a zero-sum mutually quadratic invariant game. Let $(\hat{\mathcal{Q}}_1, \hat{\mathcal{Q}}_2)$ be the corresponding disturbance feedforward policy. Since $(\hat{\mathcal{K}}_1, \hat{\mathcal{K}}_2)$ is stationary, so is $(\hat{\mathcal{Q}}_1, \hat{\mathcal{Q}}_2)$ by Lemma 1. Since $\mathbf{J}$ is convex quadratic in $\mathcal{Q}_1$ and concave quadratic in $\mathcal{Q}_2$, it follows from Assumption 1 that the stationary point is unique. Thus, $(\hat{\mathcal{Q}}_1, \hat{\mathcal{Q}}_2) = (\bar{\mathcal{Q}}_1, \bar{\mathcal{Q}}_2)$. Since g is bijective, $(\hat{\mathcal{K}}_1, \hat{\mathcal{K}}_2) = (\bar{\mathcal{K}}_1, \bar{\mathcal{K}}_2)$. Thus, $(\bar{\mathcal{K}}_1, \bar{\mathcal{K}}_2)$ is the unique saddle point equilibrium in state feedback linear strategies due to the uniqueness of $(\bar{\mathcal{Q}}_1, \bar{\mathcal{Q}}_2)$. ∎

This result allows computation of structured equilibrium feedback strategies in two-team stochastic dynamic games with mutually quadratically invariant information structures.

C. Generalization to arbitrary policies

So far we have been restricting the search to linear policies. In the following we show that a saddle point equilibrium in linear policies, under mutual quadratic invariance remains a saddle point equilibrium even if the players are allowed to use arbitrary nonlinear policies.

Lemma 3 ([28]): Suppose $A \in \mathbb{R}^{k\times k}$ is invertible, there exist $h_i \in \mathbb{R}$ with $i = 1, ..., k$ such that $A^{-1} = \sum_{i=1}^{k} h_i A^i$.

Lemma 4: Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be mutually quadratic invariant under $\mathcal{P}_{12}$ and $\mathcal{P}_{13}$. Then, for every $\mathcal{K}_2 \in \mathcal{S}_2$ such that $(I - \mathcal{P}_{13}\mathcal{K}_2)$ is invertible, $\mathcal{S}_1$ is quadratic invariant under $(I - \mathcal{P}_{13}\mathcal{K}_2)^{-1}\mathcal{P}_{12}$ and for every $\mathcal{K}_1 \in \mathcal{S}_1$ such that $(I - \mathcal{P}_{12}\mathcal{K}_1)$ is invertible, $\mathcal{S}_2$ is quadratic invariant under $(I - \mathcal{P}_{12}\mathcal{K}_1)^{-1}\mathcal{P}_{13}$.

Proof: Since the two statements are symmetric, we prove the first one. We need to show that for every $\mathcal{K}_2 \in \mathcal{S}_2$ such that $(I - \mathcal{P}_{13}\mathcal{K}_2)$ is invertible, we have

$$\mathcal{K}_1(I - \mathcal{P}_{13}\mathcal{K}_2)^{-1}\mathcal{P}_{12}\mathcal{K}_1 \in \mathcal{S}_1, \quad \forall \mathcal{K}_1 \in \mathcal{S}_1. \tag{20}$$

We begin by noting that $\mathcal{K}_1(\mathcal{P}_{13}\mathcal{K}_2)^k \in \mathcal{S}_1$ for all k. We can prove this by induction. We know that $\mathcal{K}_1 \in \mathcal{S}_1$, $\mathcal{K}_1\mathcal{P}_{13}\mathcal{K}_2 \in \mathcal{S}_1$ (from the MQI conditions). Then if $\mathcal{K}_1(\mathcal{P}_{13}\mathcal{K}_2)^{k-1} \in \mathcal{S}_1$ we have that $\mathcal{K}_1(\mathcal{P}_{13}K_2)^k = \mathcal{K}_1(\mathcal{P}_{13}\mathcal{K}_2)^{k-1}(\mathcal{P}_{13}\mathcal{K}_2)$, which is in $\mathcal{S}_1$ (again from the cross condition of MQI). Then, using Lemma 3 we can write

$$\begin{aligned} \mathcal{K}_1(I - \mathcal{P}_{13}\mathcal{K}_2)^{-1} &= \sum_{i=0}^{n(N+1)} h_i\mathcal{K}_1(I - \mathcal{P}_{13}\mathcal{K}_2)^i \\ &= \sum_{i=0}^{n(N+1)} p_i\mathcal{K}_1(\mathcal{P}_{13}\mathcal{K}_2)^i \in \mathcal{S}_1, \end{aligned}$$

where p_i are the coefficients that result after collecting each power. Since the structure $\mathcal{S}_1$ is a sparsity pattern, [16, Theorem 26] states that QI implies that $\mathcal{A}\mathcal{P}_{12}\mathcal{K}_1 \in \mathcal{S}_1$ for all $\mathcal{A}, \mathcal{K}_1 \in \mathcal{S}_1$. Therefore (20) follows and the proof is complete. ∎

Theorem 5: Let $(\bar{\mathcal{K}}_1, \bar{\mathcal{K}}_2)$ be a saddle point equilibrium in state-feedback linear strategies of a zero-sum mutually quadratic invariant game, then the set of linear strategies $(\bar{\mathcal{K}}_1, \bar{\mathcal{K}}_2)$ are a saddle point equilibrium even if the players are allowed to use state-feedback nonlinear strategies.

Proof: Let us assume that, *w.l.o.g.*, Player 2 plays the linear strategy $\mathbf{v} = \bar{\mathcal{K}}_2\mathbf{x}$. Closing the loop for Player 2, the dynamics given by (5) can now be written in the form $\mathbf{x} = \begin{bmatrix} \tilde{\mathcal{P}}_{11} & \tilde{\mathcal{P}}_{12} \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{u} \end{bmatrix}$, where

$$\tilde{\mathcal{P}}_{11} = (I - \mathcal{P}_{13}\bar{\mathcal{K}}_2)^{-1}\mathcal{P}_{11}, \quad \tilde{\mathcal{P}}_{12} = (I - \mathcal{P}_{13}\bar{\mathcal{K}}_2)^{-1}\mathcal{P}_{12}.$$

By Lemma 4, Player 1's best response $k_1^\star(\cdot)$ to the linear strategy of Player 2 is the solution of a finite horizon linear quadratic optimal control problem with QI information structure. The optimal strategy $k_1^\star(\cdot)$ is therefore be linear and unique [16], [29] under our assumptions and setting. Since the linear strategy $\bar{\mathcal{K}}_1$ is the optimal linear response to $\bar{\mathcal{K}}_2$ by hypothesis, we conclude that $k_1^\star(\cdot) = \bar{\mathcal{K}}_1\cdot$. A similar reasoning can be applied by fixing Player 1's strategy and evaluating the optimal response of Player 2. ∎

There may exist saddle point equilibria with nonlinear strategies, but any such equilibrium would result in the same value due to the ordered interchangeability of saddle point equilibria.

While Theorem 2 proves that the set of policies $(\bar{\mathcal{K}}_1, \bar{\mathcal{K}}_2)$ is a Nash equilibrium only if the players are allowed to use linear feedback policies, Theorem 5 guarantees that no player can unilaterally improve their cost by choosing a nonlinear policy.

VI. NUMERICAL EXAMPLE

We now present a very simple illustrative numerical example. The mutual quadratic invariance property allows us to compute equilibrium strategies and values under these information structures and thereby to understand the importance of different information structures in dynamic games. We consider a two player system depicted in Figure 1, which can be interpreted as a simple transportation network.

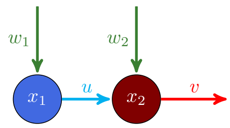

Fig. 1: Two player network

Each subsystem consists of a buffer with single integrator dynamics. System 1 stores x_1 and can control input u that transfers some of the good stored in its buffer to System 2. System 2 stores x_2 and can control input v to discard some of the good. Both systems are affected by random disturbances which are normally distributed with zero mean and unit variance. The dynamics of the system is

$$\begin{bmatrix} x_1^+ \\ x_2^+ \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix} u + \begin{bmatrix} 0 \\ -1 \end{bmatrix} v + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \qquad (21)$$

Given the dynamics in (21) with $w_1(t), w_2(t), x(0) \sim \mathcal{N}(0,1)$ we consider the following zero sum dynamic game

$$\begin{aligned} \min_{\mathbf{u}} \max_{\mathbf{v}} \quad & \sum_{t=1}^{10} 2\,\mathbb{E}x_1^2(t) + \mathbb{E}u^2(t-1) + \\ & - \mathbb{E}x_2^2(t) - 2\,\mathbb{E}v^2(t-1) \qquad (22) \\ \text{s. t.} \quad & \mathbf{u} = \mathcal{K}_1\mathbf{x}, \quad \mathcal{K}_1 \in \mathcal{S}_1 \\ & \mathbf{v} = \mathcal{K}_2\mathbf{x}, \quad \mathcal{K}_2 \in \mathcal{S}_2, \end{aligned}$$

where $\mathbf{x} = (x(0), ..., x(10))$, $\mathbf{u} = (u(0), ..., u(9))$, $\mathbf{v} = (v(0), ..., v(9))$. Both players benefit from keeping the variance of their state and input low and from increasing the variance of the opponent's state and input. We will compare the results for different information structures $\mathcal{S}_1$ and $\mathcal{S}_2$, all of which are mutually quadratically invariant. In particular, we consider:

- **Causal controllers with full information (FI)**. That is both players have access to all past and present information:

$$\mathcal{S}_1 = \mathcal{S}_2 = \begin{bmatrix} \star & \star & & & & & & & & \\ \star & \star & \star & \star & & & & & & \\ \star & \star & \star & \star & \star & \star & & & & \\ \star & \star & \star & \star & \star & \star & \star & \star & & \\ \star & \star & \star & \star & \star & \star & \star & \star & \star & \star \\ & & & & & \vdots & & & & \end{bmatrix}.$$

- **One step delay information sharing (1SDIS)**. At time t both players do not know the opponent's current state but have full information up to time $t-1$:

$$\mathcal{S}_1 = \begin{bmatrix} \star & 0 & & & & & & & & \\ \star & \star & \star & 0 & & & & & & \\ \star & \star & \star & \star & \star & 0 & & & & \\ \star & \star & \star & \star & \star & \star & \star & 0 & & \\ \star & \star & \star & \star & \star & \star & \star & \star & \star & 0 \\ & & & & & & \vdots & & & \end{bmatrix},$$

$$\mathcal{S}_2 = \begin{bmatrix} 0 & \star & & & & & & & & \\ \star & \star & 0 & \star & & & & & & \\ \star & \star & \star & \star & 0 & \star & & & & \\ \star & \star & \star & \star & \star & \star & 0 & \star & & \\ \star & \star & \star & \star & \star & \star & \star & \star & 0 & \star \\ & & & & & & \vdots & & & \end{bmatrix}.$$

- **Decentralized control for Player 1 (DP1)**. Player 1 only has access to present and past information about its own state, and Player 2 has full information:

$$\mathcal{S}_1 = \begin{bmatrix} \star & 0 & & & & & & & & \\ \star & 0 & \star & 0 & & & & & & \\ \star & 0 & \star & 0 & \star & 0 & & & & \\ \star & 0 & \star & 0 & \star & 0 & \star & 0 & & \\ \star & 0 & \star & 0 & \star & 0 & \star & 0 & \star & 0 \\ & & & & & & \vdots & & & \end{bmatrix},$$

and $\mathcal{S}_2$ is full information as above in case FI.

It is easily verified that all of these structures respect the mutual quadratic invariance condition for the given system dynamics. We computed the saddle point equilibrium feedforward strategies $(\bar{\mathcal{Q}}_1, \bar{\mathcal{Q}}_2)$ for the three different information structures using the method proposed in [24]. We applied Theorem 2 to compute the linear saddle point equilibrium in the state feedback strategies as

$$\begin{bmatrix} \bar{\mathcal{K}}_1 \\ \bar{\mathcal{K}}_2 \end{bmatrix} = g\left(\begin{bmatrix} \bar{\mathcal{Q}}_1 \\ \bar{\mathcal{Q}}_2 \end{bmatrix}\right).$$

In Table I we observe the different costs at equilibrium. Note the large difference in cost function that is achieved for different information structures. Under the full information structure, the equilibrium value is -1.58. Using this structure as a baseline, Player 1 (the minimizing player) is penalized by using decentralized information (DP1), as expected, where the cost becomes 0.00. On the other hand, Player 1 obtains a great advantage with the one step delay information sharing (1SDIS) structure, where the cost becomes -10.02, as the reduction in information significantly limits the relative performance of Player 2.

TABLE I: The cost at the saddle point equilibrium for the three different information structures. A smaller cost is indicates an advantage for Player 1 who is minimizing in (22), while a larger cost is an advantage for Player 2.

Equilibrium cost for the different information structures	
Structure	Equilibrium Cost (22)
FI	-1.58
1SDIS	-10.02
DP1	0.00

Figure 2 shows the breakdown of the cost function of (22) for the different information structures and allows us to understand why certain information structures are more beneficial for different players. For example, if we consider 1SDIS we notice that Player 1 can exploit the fact that its opponent has no information on Player 1's current state and input and it uses this to dramatically increase the variance of x_2 (as indicated by the dark red bars third from the left in each group). To do so Player 1 needs to 'spend' some variance in u (as indicated by the light blue bars second from the left in each group), which is also increased. The net gain, however, is clearly in favor of Player 1. On the other hand, for DP1, Player 2 can exploit the lack of information of Player 1 to significantly reduce its input and state variance. Such calculations not only illustrate the importance of information structures in dynamic team games, but also may be used as a basis for the *design* of information structures that improve team performance and resilience to antagonistic teams, as considered in [30].

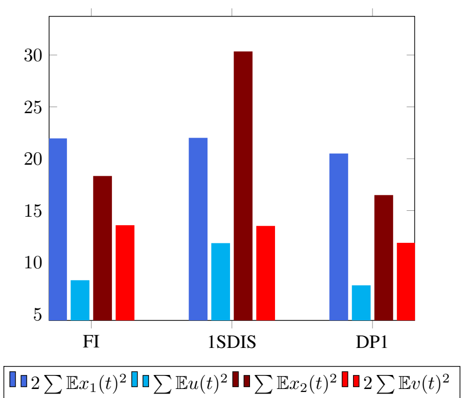

Fig. 2: Breakdown of the cost function of (22) for the three information structures

VII. Conclusion

We have considered a two-team linear quadratic stochastic dynamic game with decentralized information structures for both teams. We introduced the concept of Mutual Quadratic Invariance (MQI), which defines a class of interacting team information structures for which equilibrium strategies are linear and can be easily computed in the zero-sum setting. In particular, we demonstrated an equivalence of disturbance feedforward and state feedback saddle point equilibrium strategies that facilitates this computation in zero-sum games, and we showed such an equivalence fails to hold for Nash equilibrium strategies in nonzero-sum games. A numerical example showed how mutually quadratically invariant information structures can be evaluated and how different structures can lead to significantly different equilibrium values.

Many fundamental questions remain open in two-team stochastic dynamic games. For example, issues involving infinite horizon and boundedness of the equilibrium value (stability), separation and certainty equivalence, games with incomplete model information, and design of information structures can be considered. Some of these results may take inspiration from recent progress on information structure issues in decentralized control.

References

[1] T. Başar and G. Olsder, *Dynamic noncooperative game theory*, vol. 200. SIAM, 1995.

[2] J. Von Neumann and O. Morgenstern, *Theory of games and economic behavior*. Princeton University Press, 1944.

[3] R. Bellman, "On the theory of dynamic programming," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 38, no. 8, p. 716, 1952.

[4] L. Pontryagin, "Optimal control processes," *Usp. Mat. Nauk*, vol. 14, no. 3, 1959.

[5] J. Marschak, "Elements for a theory of teams," *Management Science*, vol. 1, no. 2, pp. 127–137, 1955.

[6] R. Radner, "Team decision problems," *The Annals of Mathematical Statistics*, pp. 857–881, 1962.

[7] J. Marshak and R. Radner, *Economic theory of teams*. Yale University Press, 1972.

[8] H. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.

[9] H. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proceedings of the IEEE*, vol. 59, no. 11, pp. 1557–1566, 1971.

[10] H. Witsenhausen, "On information structures, feedback and causality," *SIAM Journal on Control*, vol. 9, no. 2, pp. 149–160, 1971.

[11] H. S. Witsenhausen, "Equivalent stochastic control problems," *Mathematics of Control, Signals, and Systems (MCSS)*, vol. 1, no. 1, pp. 3–11, 1988.

[12] Y. C. Ho and K.-C. Chu, "Team decision theory and information structures in optimal control problems–part i," *IEEE Transactions on Automatic Control*, vol. 17, no. 1, pp. 15–22, 1972.

[13] Y.-C. Ho, "Team decision theory and information structures," *Proceedings of the IEEE*, vol. 68, no. 6, pp. 644–654, 1980.

[14] P. Grover, S. Y. Park, and A. Sahai, "Approximately optimal solutions to the finite-dimensional Witsenhausen counterexample," *IEEE Trans. on Automatic Control*, vol. 58, no. 9, pp. 2189–2204, 2013.

[15] P. Grover, A. B. Wagner, and A. Sahai, "Information embedding and the triple role of control," *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1539–1549, 2015.

[16] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *Automatic Control, IEEE Transactions on*, vol. 51, no. 2, pp. 274–286, 2006.

[17] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1644–1658, 2013.

[18] S. Yüksel and T. Başar, "Stochastic networked control systems," *AMC*, vol. 10, p. 12, 2013.

[19] J. Swigart and S. Lall, "Optimal controller synthesis for decentralized systems over graphs via spectral factorization," *IEEE Transactions on Automatic Control*, vol. 59, no. 9, pp. 2311–2323, 2014.

[20] L. Lessard and S. Lall, "Optimal control of two-player systems with output feedback. to appear," *IEEE Transactions on Automatic Control*, vol. 60, no. 8, pp. 2129 – 2144, 2015.

[21] A. Gattami, B. M. Bernhardsson, and A. Rantzer, "Robust team decision theory," *Automatic Control, IEEE Transactions on*, vol. 57, no. 3, pp. 794–798, 2012.

[22] T. Başar, "Stochastic differential games and intricacy of information structures," in *Dynamic Games in Economics*, pp. 23–49, Springer, 2014.

[23] T. Başar and P. Bernhard, *H-infinity optimal control and related minimax design problems: a dynamic game approach*. Springer Science & Business Media, 1990.

[24] M. Colombino, T. Summers, and R. Smith, "Quadratic two-team games," in *IEEE Conference on Decision and Control, Osaka, Japan*, pp. 3784–3789, 2015.

[25] B. Hassibi, A. H. Sayed, and T. Kailath, "Indefinite-quadratic estimation and control," *Studies in Applied & Numerical Math.*, 1999.

[26] T. Basar, "Decentralized multicriteria optimization of linear stochastic systems," *Automatic Control, IEEE Transactions on*, vol. 23, no. 2, pp. 233–243, 1978.

[27] J. Swigart and S. Lall, "An explicit state-space solution for a decentralized two-player optimal linear-quadratic regulator," in *Proceedings of the 2010 American Control Conference*, pp. 6385–6390, IEEE, 2010.

[28] L. Lessard and S. Lall, "An algebraic framework for quadratic invariance," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, pp. 2698–2703, IEEE, 2010.

[29] M. Rotkowitz, "On information structures, convexity, and linear optimality," in *IEEE Conf. on Decision and Control*, pp. 1642–47, 2008.
[30] T. Summers, C. Li, and M. Kamgarpour, "Information structure design in team decision problems," *IFAC World Congress*, 2017.