Decentralized Control Problems with Substitutable Actions

Seyed Mohammad Asghari and Ashutosh Nayyar

Abstract— **We consider a decentralized system with multiple controllers and define substitutability of one controller by another in open-loop strategies. We explore the implications of this property on the optimization of closed-loop strategies. In particular, we focus on the decentralized LQG problem with substitutable actions. Even though the problem we formulate does not belong to the known classes of "simpler" decentralized problems such as partially nested or quadratically invariant problems, our results show that, under the substitutability assumption, linear strategies are optimal and we provide a complete state space characterization of optimal strategies. We also identify a family of information structures that all give the same optimal cost as the centralized information structure under the substitutability assumption. Our results suggest that open-loop substitutability can work as a counterpart of the information structure requirements that enable simplification of decentralized control problems.**

I. INTRODUCTION

The difficulty of finding optimal strategies in decentralized control problems has been well-established in the literature [1], [2], [3]. In general, the optimization of strategies can be a non-convex problem over infinite dimensional spaces [4]. Even the celebrated linear quadratic Gaussian (LQG) model of centralized control presents difficulties in the decentralized setting [1], [2], [5]. There has been significant interest in identifying classes of decentralized control problems that are more tractable. Information structures of decentralized control problems, which describe what information is available to which controller, have been closely associated with their tractability. Problems with partially nested [6] or stochastically nested information structures [7] and problems that satisfy quadratic invariance [8] or funnel causality [9] properties have been identified as "simpler" than the general decentralized control problems.

In this paper, instead of starting from the information structure of the problem, we first look at open-loop strategies under which controllers take actions without any observations. Stated another way, we start with a trivially simple information structure: no controller knows anything (except, of course, the model of the system and the cost objective).

We define a property of open-loop decentralized control, namely the substitutability of one controller by another, and explore its implications on optimization of closed-loop strategies (under which controllers take actions as functions of their observations). In particular, we focus on the decentralized LQG problem with substitutable actions. Even

though the problem we formulate does not belong to one of the simpler classes mentioned earlier (partially nested, quadratic invariant etc.), (i) our results show that linear strategies are optimal; (ii) we provide a complete state space characterization of optimal strategies; (iii) we also identify a family of information structures that all achieve the same cost as the centralized information structure. Our results suggest that open-loop substitutability can work as a counterpart of the information structure requirements that enable simplification of decentralized control problems.

Our work shares conceptual similarities with the work on internal quadratic variance [10], [11] which identified problems that are not quadratically invariant but can still be reduced to (infinite dimensional) convex programs. In contrast to this work, we explicitly identify optimal control strategies.

A. Notation

Uppercase letters denote random variables/vectors and their corresponding realizations are represented by lowercase letters. Uppercase letters are also used to denote matrices. $\mathrm{E}[\cdot]$ denotes the expectation of a random variable. When random variable X is normally distributed with mean μ and variance Σ, it is shown as $X \sim \mathcal{N}(\mu, \Sigma)$.

For a sequence of column vectors $X, Y, Z, ...$, the notation $vec(X, Y, Z, ...)$ denotes vector $[X^\mathsf{T}, Y^\mathsf{T}, Z^\mathsf{T}, ...]^\mathsf{T}$. Furthermore, the vector $vec(X_1, X_2, ..., X_t)$ is denoted by $X_{1:t}$. The transpose and Moore-Penrose pseudo-inverse of matrix A are denoted by A^T and $A^\dagger$, respectively. The identity matrix and zero vector are denoted by I and 0 respectively and their dimensions are inferred from the context.

II. SUBSTITUTABLE ACTIONS

We consider a stochastic system with n controllers. The dynamics of the system are given as:

$$X_{t+1} = f(X_t, U_t^1, \ldots, U_t^n, W_t), \quad t = 1, \ldots, T-1. \quad (1)$$

where X_t is the state of the system at time t, U_t^i is the action of controller i at time t and W_t is a random noise variable. The state takes value in the set $\mathcal{X}$, the control action of the ith controller takes value in the set $\mathcal{U}^i$ and the noise W_t takes value in the set $\mathcal{W}$. We use U_t to denote the vector $vec(U_t^1, \ldots, U_t^n)$.

The system operates in discrete time for a horizon T. At time step t, the system incurs a cost given as a function of the state and control actions: $c(X_t, U_t^1, \ldots, U_t^n)$. The control objective is to minimize the expected value of the total cost accumulated over the T time steps: $\mathrm{E}[\sum_{t=1}^{T} c(X_t, U_t^1, \ldots, U_t^n)]$.

S. M. Asghari and A. Nayyar are with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089 USA (e-mail: asgharip@usc.edu; ashutosn@usc.edu).

We say that controller i is using open-loop control strategy if its control actions are a function only of time and not of any information (observations) obtained from the system. Otherwise, we say that controller i is using a closed-loop control strategy. Based on the dynamics and the cost function, we can define a notion of open-loop substitutability among controllers.

Definition 1: We say that controller 1 can substitute for controller 2 in open-loop control if for every $u^1 \in \mathcal{U}^1, u^2 \in \mathcal{U}^2$, there exists a control action $v^1 \in \mathcal{U}^1$ for controller 1 such that for all $x \in \mathcal{X}, w \in \mathcal{W}, u^i \in \mathcal{U}^i, i = 3, \ldots, n,$

$$f(x, u^1, u^2, u^3, \ldots, u^n, w) = f(x, v^1, 0, u^3, \ldots, u^n, w), \tag{2a}$$

$$\text{and } c(x, u^1, u^2, u^3, \ldots, u^n) = c(x, v^1, 0, u^3, \ldots, u^n). \tag{2b}$$

v^1 is a function only of u^1 and u^2, that is, $v^1 = l^{1,2}(u^1, u^2)$, for some function $l^{1,2}$. We will call $l^{1,2}$ the *substitution function* for controller 1 and 2.

A similar notion of open-loop substitutability can be defined for any pair of controllers in the system.

If we are considering only open-loop control strategies for all controllers and if controller 1 can substitute for controller 2 in open-loop control (as per Definition 1), then there is no loss of optimality in fixing all actions of controller 2 to 0. This is the intuitive meaning of substitutability — since controller 1 can substitute for controller 2, controller 2 does not need to do anything.

The open-loop substitutability property has some closed-loop implications. Let us denote by I_t^i the collection of all observations and past control actions that are available to controller i at time t. Under a closed-loop strategy, controller i selects its control action as a function of I_t^i, that is, $U_t^i = g_t^i(I_t^i)$. The collection of functions $g^i := \{g_1^i, g_2^i, \ldots, g_T^i\}$ is referred to as controller i's (closed-loop) strategy.

Lemma 1: Suppose that controller 1 can substitute for controller 2 in open-loop control and that for all time instants t, $I_t^1 \supseteq I_t^2$, then for any strategies $h^1, h^2, \ldots, h^n$ of n controllers, there exists strategies g^1, g^2 for controllers 1 and 2, with $g_t^2(I_t^2) = 0$ for all t, such that $g^1, g^2, h^3, \ldots, h^n$ achieve the same cost as $h^1, h^2, \ldots, h^n$.

Proof: Consider any arbitrary strategies $h^1, h^2, \ldots, h^n$ for the n controllers. Define new strategy for controller 1 as follows:

$$g_t^1(I_t^1) = l^{1,2}(h_t^1(I_t^1), h_t^2(I_t^2)), \tag{3}$$

where $l^{1,2}$ is the substitution function from Definition 1. Firstly, note that (3) is a valid strategy for controller 1 because $I_t^2 \subseteq I_t^1$. If this was not the case, the right hand side of (3) would be using information that controller 1 may not have.

The result of the lemma then follows from the observation that the pair $(U_t^1 = g_t^1(I_t^1), U_t^2 = 0)$ will always have the same effects on dynamics and cost as $(U_t^1 = h_t^1(I_t^1), U_t^2 = h_t^2(I_t^2))$ because of the substitutability conditions. ■

The condition $I_t^1 \supseteq I_t^2$ is necessary for Lemma 1 to hold. It is easy to construct examples where I_t^1 does not include I_t^2

and the second controller cannot be restricted to the "always zero" strategy without losing optimality.

The statement of Lemma 1 can be intuitively interpreted as follows: The open-loop substitutability of controller 2 by controller 1 *and* the fact that controller 1 is better informed make controller 2 essentially redundant for the purpose of cost optimization.

Lemma 1 suggests that open-loop substitutability, combined with the information structure of the problem, can have implications about the closed-loop problem. In the rest of the paper, we consider a LQG control problem with multiple controllers and obtain results much sharper than Lemma 1 for such problems.

III. LQG PROBLEM WITH STATE FEEDBACK

A. System Model

We consider a stochastic system with n controllers where
1) The state dynamics are given as

$$X_{t+1} = AX_t + BU_t + W_t, \quad t = 1, \ldots, T - 1, \tag{4}$$

where $X_t, W_t \in \mathbb{R}^{d_x}$ and $U_t \in \mathbb{R}^{d_u}$.
2) The cost at time t is given as

$$c(X_t, U_t) = (MX_t + NU_t)^{\mathsf{T}}(MX_t + NU_t). \tag{5}$$

3) The initial state X_1 and the noise variables $W_t, t = 1, \ldots, T - 1$ are independent and have Gaussian distributions.

We will make the following assumption about the system.

Assumption 1: Each controller can substitute for any other controller in open loop control. In other words, for every vector $u = \text{vec}(u^1, u^2, \ldots, u^n)$, there exist control actions $v^i = l^i(u)$ for controller i, $i = 1, \ldots, n$, such that

$$Bu = B \begin{bmatrix} 0 \\ \vdots \\ v^i \\ \vdots \\ 0 \end{bmatrix} \quad \text{and} \quad Nu = N \begin{bmatrix} 0 \\ \vdots \\ v^i \\ \vdots \\ 0 \end{bmatrix}. \tag{6}$$

Lemma 2: We can write the B and N matrices in terms of their blocks as

$$B = \begin{bmatrix} B^1 & \cdots & B^n \end{bmatrix},$$

$$N = \begin{bmatrix} N^1 & \cdots & N^n \end{bmatrix}.$$

If Assumption 1 is true, then $v^i = \Lambda^i u$ [1] satisfies (6) for $i = 1, \ldots, n$ where

$$\Lambda^i = \begin{bmatrix} B^i \\ N^i \end{bmatrix}^{\dagger} \begin{bmatrix} B \\ N \end{bmatrix}. \tag{7}$$

---

[1] In the previous version of this work [12], we stated that if Assumption 1 is true, then $v^i = (B^i)^{\dagger}Bu = (N^i)^{\dagger}Nu$ and $(B^i)^{\dagger}B = (N^i)^{\dagger}N$. This statement is correct only if B^i and N^i are invertible so that $(B^i)^{\dagger} = (B^i)^{-1}$ and $(N^i)^{\dagger} = (N^i)^{-1}$. However, in general, the statement is incorrect. The correct expression for v^i is $v^i = \Lambda^i u$ where Λ^i is as given by (7).

Proof: The system of equations of (6) is equivalent to a matrix equation of the form

$$Pv^i = b, \text{ where } P = \begin{bmatrix} B^i \\ N^i \end{bmatrix}, b = \begin{bmatrix} B \\ N \end{bmatrix} u. \quad (8)$$

The general solution of (8) is $v^i = P^\dagger b + (I - P^\dagger P)y$ for arbitrary y [13]. By setting $y = 0$, we have $v^i = P^\dagger b$. ∎

An example of a system satisfying Assumption 1 is a two-controller LQG problem where the dynamics and the cost are functions only of the sum of the control actions, that is, $(u_t^1 + u_t^2)$. This happens if $B^1 = B^2$ and $N^1 = N^2$. In this case, using $v_t^1 = v_t^2 = u_t^1 + u_t^2$ satisfies (6), which means that controller 1 can substitute for controller 2 and vice versa.

B. Information Structure

We assume that the state vector X_t consists of n sub-vectors, that is $X_t = vec(X_t^1, X_t^2, \ldots, X_t^n)$. X_t^i can be interpreted as the state of the ith sub-system. We assume a local state feedback with perfect recall information structure, that is, the information available to controller i at time t is:

$$I_t^i = \{X_{1:t}^i, U_{1:t-1}^i\}, \qquad i = 1, 2, \ldots, n. \quad (9)$$

Controller i chooses action U_t^i as a function of the information available to it. Specifically, for $i = 1, \ldots, n$,

$$U_t^i = g_t^i(I_t^i), \qquad t = 1, \ldots, T. \quad (10)$$

The collection $g^i = (g_1^i, \ldots, g_T^i)$ is called the control strategy of controller i. The performance of the control strategies g^1, $g^2, \ldots, g^n$ is measured by the expected cost

$$\mathcal{J}(g^1, \ldots, g^n)$$
$$= E^{g^1, \ldots, g^n} \left[\sum_{t=1}^T (MX_t + NU_t)^\mathsf{T}(MX_t + NU_t) \right] \quad (11)$$

where the expectation is with respect to the joint probability distribution on $(X_{1:T}, U_{1:T})$ induced by the choice of $g^1, \ldots, g^n$.

C. Optimal Strategies

The optimization problem is defined as follows.

Problem 1: For the model described in section III-A and III-B, find control strategies $g^1, \ldots, g^n$ for the n controllers that minimize the expected cost given by (11).

Remark 1: Since we have not imposed any constraints on the matrices A and B in system dynamics, Problem 1 may not have partially nested information structure. Thus, we cannot guarantee, at this point, the optimality of linear control strategies for this problem.

In addition to the decentralized information structure described above, we will also consider the centralized information structure where all controllers have access to the entire state and action history.

Problem 2: For the model described in section III-A, assume that the information available to each controller is,

$$\tilde{I}_t = \{X_{1:t}, U_{1:t-1}\}. \quad (12)$$

Controller i chooses its action according to strategy $h^i = (h_1^i, \ldots, h_T^i)$,

$$U_t^i = h_t^i(\tilde{I}_t).$$

The objective is to select controller strategies that minimize

$$\mathcal{J}(h^1, \ldots, h^n)$$
$$= E^{h^1, \ldots, h^n} \left[\sum_{t=1}^T (MX_t + NU_t)^\mathsf{T}(MX_t + NU_t) \right] \quad (13)$$

where the expectation is with respect to the joint probability distribution on $(X_{1:T}, U_{1:T})$ induced by the choice of $h^1, \ldots, h^n$.

D. Main results

In this section, we will show that we can construct optimal strategies in Problem 1 from the optimal control strategies of the centralized problem (Problem 2). We start with the following observations.

Lemma 3: 1) The optimal cost in Problem 2 (with centralized information structure) is a lower bound on the optimal cost in Problem 1 (with decentralized information structure).

2) The optimal strategies in Problem 2 are linear functions of the state, that is, the control vector $U_t = vec(U_t^1, \ldots, U_t^n)$ is given as

$$U_t = K_t X_t = [K_t^1 \ldots K_t^n] X_t$$
$$= K_t^1 X_t^1 + \ldots + K_t^n X_t^n. \quad (14)$$

K_t is the centralized gain matrix and K_t^i is its ith block.

We can now state our main result for the state feedback case.

Theorem 1: The optimal control strategies in Problem 1 are given as

$$U_t^i = l_t^i(K_t^i X_t^i) = \Lambda^i K_t^i X_t^i, \qquad i = 1, 2, \ldots, n, \quad (15)$$

where Λ^i is given by (7) and K_t^i is the ith block of the centralized gain matrix in (14). Moreover, the optimal strategies in Problem 1 achieve the same cost as the optimal strategies in Problem 2.

Proof Outline: Firstly, observe that the strategies given by (15) are valid control strategies under the information structure of Problem 1. The optimal control vector U_t under the centralized strategy is a superposition of terms of the form $K_t^i X_t^i$. Note that the term $K_t^i X_t^i$ consists of n sub-vectors (one corresponding to each controller's action).

$$K_t^i X_t^i = \begin{bmatrix} K_t^{i1} \\ K_t^{i2} \\ \vdots \\ K_t^{in} \end{bmatrix} X_t^i.$$

Such a control vector cannot be implemented in the decentralized information structure since it requires each controller to have access to X_t^i. We now exploit the open loop substitutability of the problem to state that the vector $vec(0, \ldots, l_t^i(K_t^i X_t^i), \ldots, 0)$ will have the same effect as

$K_t^i X_t^i$. This allows us to construct a decentralized strategy with the same performance as the centralized one. We provide a detailed proof for the more general case of the output feedback problem in the next section.

We can also derive the following corollary of Theorem 1.

Corollary 1: For the model described in section III-A, consider any information structure under which the information of controller i at time t, $\hat{I}_t^i$, satisfies

$$\{X_t^i\} \subseteq \hat{I}_t^i \subseteq \{X_{1:t}, U_{1:t-1}\},$$

for all $i = 1, \ldots, n$ and $t = 1, \ldots, T$. Then, the optimal strategies in this information structure are the same as in Theorem 1.

Corollary 1 identifies memoryless local state feedback as the minimal information structure that achieves the optimal centralized cost. In other words, it describes the minimal communication and memory requirements for controllers to achieve the optimal centralized cost.

IV. LQG PROBLEM WITH OUTPUT FEEDBACK

A. System Model

We consider the system model described in section III-A and assume that each controller makes a noisy observation of the system state given as

$$Y_t^i = C^i X_t + V_t^i, \qquad i = 1, \ldots, n. \tag{16}$$

Combining (16) for all controllers gives:

$$Y_t = \begin{bmatrix} C^1 \\ C^2 \\ \vdots \\ C^n \end{bmatrix} X_t + V_t, \tag{17}$$

where Y_t denotes $vec(Y_t^1, Y_t^2, \ldots, Y_t^n)$ and V_t denotes $vec(V_t^1, V_t^2, \ldots, V_t^n)$. The initial state X_1 and the noise variables $W_t, t = 1, \ldots, T-1$, and $V_t, t = 1, \ldots, T-1$, are mutually independent and jointly Gaussian with the following probability distributions:

$$X_1 \sim \mathcal{N}(0, \Sigma_x), \qquad W_t \sim \mathcal{N}(0, \Sigma_w), \qquad V_t \sim \mathcal{N}(0, \Sigma_v).$$

The information available to the ith controller at time t is:

$$I_t^i = \{Y_{1:t}^i, U_{1:t-1}^i\} \qquad i = 1, \ldots, n. \tag{18}$$

Each controller i, chooses its action U_t^i according to $U_t^i = g_t^i(I_t^i)$ and the performance of the control strategies of all controllers, $(g^1, \ldots, g^n)$, is measured by (11).

The optimization problem is defined as follows.

Problem 3: For the model described above, find control strategies $g^1, \ldots, g^n$ for the n controllers that minimize the expected cost given by (11).

In addition to the decentralized information structure described above, we will also consider the centralized information structure and the corresponding strategy optimization.

Problem 4: For the model described above, assume that the information available to each controller is

$$\tilde{I}_t = \{Y_{1:t}, U_{1:t-1}\}. \tag{19}$$

Controller i chooses its action according to strategy $U_t^i = h_t^i(\tilde{I}_t)$. The objective is to select control strategies that minimize (13).

The following lemma follows directly from the problem descriptions above and well-known results for the centralized LQG problem with output feedback [14].

Lemma 4: 1) The optimal cost in Problem 4 (with centralized information structure) is a lower bound on the optimal cost in Problem 3 (with decentralized information structure).

2) The optimal strategies in Problem 4 have the form of $U_t = K_t Z_t$ where $Z_t = \mathrm{E}(X_t | \tilde{I}_t)$. Z_t evolves according to the following equations:

$$Z_1 = L_1 Y_1$$
$$Z_{t+1} = (I - L_{t+1}C)(AZ_t + BU_t) + L_{t+1}Y_{t+1}. \tag{20}$$

We define $\Sigma_t = \mathrm{E}[(X_t - Z_t)(X_t - Z_t)^\mathsf{T} | \tilde{I}_t]$ which satisfies the following update equations:

$$\Sigma_1 = (I - L_1 C)\Sigma_x$$
$$\Sigma_{t+1} = (I - L_{t+1}C)(A\Sigma_t A^\mathsf{T} + \Sigma_w). \tag{21}$$

The matrices $L_1, \ldots, L_T$ in (20) and (21) satisfy the forward recursion:

$$L_1 = \Sigma_1 C^\mathsf{T}[C\Sigma_1 C^\mathsf{T} + \Sigma_v]^{-1}$$
$$L_{t+1} = (A\Sigma_t A^\mathsf{T} + \Sigma_w)C^\mathsf{T}[C(A\Sigma_t A^\mathsf{T} + \Sigma_w)C^\mathsf{T} + \Sigma_v]^{-1}. \tag{22}$$

B. Main results

In this section, we show that it is possible to construct optimal strategies in Problem 3 from the optimal control strategy of Problem 4.

Theorem 2: Consider Problems 3 and 4, and consider the optimal strategy, $U_t = K_t Z_t$, of Problem 4, where K_t and Z_t are as defined in Lemma 4. We write L_{t+1} of Lemma 4 as $L_{t+1} = \begin{bmatrix} L_{t+1}^1 & L_{t+1}^2 & \ldots & L_{t+1}^n \end{bmatrix}$. The optimal control strategies of Problem 3 can be written as

$$U_t^i = \Lambda^i K_t S_t^i \tag{23}$$

where Λ^i is given by (7) and S_t^i satisfies the following update equations:

$$S_1^i = L_1^i Y_1^i$$
$$S_{t+1}^i = (I - L_{t+1}C)(AS_t^i + B^i U_t^i) + L_{t+1}^i Y_{t+1}^i. \tag{24}$$

Moreover, the optimal strategies in Problem 3 achieve the same cost as the optimal strategies in Problem 4.

Observe that the strategies given by (23) and (24) are valid control strategies under the information structure of Problem 3 because they depend only on $Y_{1:t}^i, U_{1:t-1}^i$ which are included in I_t^i. The states S_t^i defined in (24) are related to the centralized estimate Z_t by the following result.

Lemma 5: The centralized state estimate Z_t and the states S_t^i defined in (24) satisfy the following equation:

$$Z_t = \sum_{i=1}^{n} S_t^i. \tag{25}$$

Proof: We prove the result by induction. For $t = 1$, from (20), we have $Z_1 = L_1 Y_1$ and according to (24),

$$\sum_{i=1}^{n} S_1^i = L_1^1 Y_1^1 + L_1^2 Y_1^2 + \ldots + L_1^n Y_1^n = L_1 Y_1. \tag{26}$$

Now assume that $Z_t = \sum_{i=1}^{n} S_t^i$. We need to show that $Z_{t+1} = \sum_{i=1}^{n} S_{t+1}^i$. From (20), it follows that

$$Z_{t+1} = (I - L_{t+1}C)(AZ_t + BU_t) + L_{t+1}Y_{t+1}. \tag{27}$$

From (24), we have

$$\sum_{i=1}^{n} S_{t+1}^i = \sum_{i=1}^{n} [(I - L_{t+1}C)(AS_t^i + B^i U_t^i) + L_{t+1}^i Y_{t+1}^i]$$
$$= (I - L_{t+1}C)(A\sum_{i=1}^{n} S_t^i + \sum_{i=1}^{n} B^i U_t^i) + \sum_{i=1}^{n} L_{t+1}^i Y_{t+1}^i]$$
$$= (I - L_{t+1}C)(AZ_t + BU_t) + L_{t+1}Y_{t+1}. \tag{28}$$

Therefore, $Z_{t+1} = \sum_{i=1}^{n} S_{t+1}^i$. ∎

Remark 2: For the case of state feedback, it can be easily shown that $S_t^i = vec(0, \ldots, X_t^i, \ldots, 0)$.

The following result is an immediate consequence of Theorem 2.

Corollary 2: For the model described in section IV-A, consider any information structure under which the information of controller i at time t, $\hat{I}_t^i$, satisfies

$$\{Y_{1:t}^i, U_{1:t-1}^i\} \subseteq \hat{I}_t^i \subseteq \{Y_{1:t}, U_{1:t-1}\},$$

for all $i = 1, \ldots, n$ and $t = 1, \ldots, T$. Then, the optimal strategies in this information structure are the same as in Theorem 2.

C. Proof of Theorem 2

For notational conveniences, we will describe the proof for $n = 2$. If $U_t = K_t Z_t$ is the optimal control strategy of Problem 4, then from Lemma 5, we have:

$$U_t = K_t Z_t = K_t(S_t^1 + S_t^2) \tag{29}$$

We claim that the decentralized control strategies defined in Theorem 2, that is

$$U_t = \begin{bmatrix} U_t^1 \\ U_t^2 \end{bmatrix} = \begin{bmatrix} \Lambda^1 K_t S_t^1 \\ \Lambda^2 K_t S_t^2 \end{bmatrix}, \tag{30}$$

yield the same expected cost as the optimal centralized control strategies $U_t = K_t Z_t$.

To establish the above claim, we define cost-to-go functions under the optimal centralized strategy and the strategies defined in Theorem 2. These functions, denoted by $\mathbf{V}_r(z)$

and $\hat{\mathbf{V}}_r(z, s^1, s^2)$ for $r = T, T-1, \ldots, 1$, are defined as follows:

$$\mathbf{V}_r(z) =$$
$$\mathrm{E}\Big[\sum_{t=r}^{T}(MX_t + NU_t)^\intercal(MX_t + NU_t)|Z_r = z, U_r = K_r z\Big], \tag{31}$$

where U_t is given by (29) for all t, and

$$\hat{\mathbf{V}}_r(z, s^1, s^2) =$$
$$\mathrm{E}\Big[\sum_{t=r}^{T}(MX_t + NU_t)^\intercal(MX_t + NU_t)|Z_r = z, S_r^1 = s^1,$$
$$S_r^2 = s^2, U_r^1 = \Lambda^1 K_r s^1, U_r^2 = \Lambda^2 K_r s^2\Big], \tag{32}$$

where U_t is given by (30) for all t. The function $\hat{\mathbf{V}}_r(z, s^1, s^2)$ in (32) is defined only for $z = s^1 + s^2$; $\hat{\mathbf{V}}_r(z, s^1, s^2)$ is undefined for $z \neq s^1 + s^2$.

We will show that for $r = 1, \ldots, T$, $\mathbf{V}_r(z) = \hat{\mathbf{V}}_r(z, s^1, s^2)$ $\forall z, s^1, s^2 \in \mathbb{R}^{d_x}$ such that $z = s^1 + s^2$. We follow a backward induction argument. For $r = T$, we have,

$$\mathbf{V}_T(z) = \mathrm{E}[(MX_T + NU_T)^\intercal(MX_T + NU_T)|Z_T = z, U_T = K_T z] \tag{33}$$
$$\hat{\mathbf{V}}_T(z, s^1, s^2) =$$
$$\mathrm{E}[(MX_T + NU_T)^\intercal(MX_T + NU_T)|Z_T = z, S_T^1 = s^1,$$
$$S_T^2 = s^2, U_T^1 = \Lambda^1 K_T s^1, U_T^2 = \Lambda^2 K_T s^2]. \tag{34}$$

Since the only difference between (33) and (34) is with respect to their different control strategies, it suffices to show that the term NU_T is the same under these two control strategies.

Under control action $u_T = K_T z$, we have $Nu_T = NK_T z$. Under control actions $u_T^1 = \Lambda^1 K_T s^1, u_T^2 = \Lambda^2 K_T s^2$, we have

$$Nu_T = \begin{bmatrix} N^1 & N^2 \end{bmatrix} u_T = N^1 u_T^1 + N^2 u_T^2$$
$$= N^1 \Lambda^1 K_T s^1 + N^2 \Lambda^2 K_T s^2 \tag{35}$$

From the substitutability assumption (Assumption 1) and Lemma 2, for any vector u, $Nu = N^i l^i(u) = N^i \Lambda^i u$. Therefore,

$$N^1 \Lambda^1 K_T s^1 = NK_T s^1,$$
$$N^2 \Lambda^2 K_T s^2 = NK_T s^2. \tag{36}$$

(35) can now be written as,

$$N^1 \Lambda^1 K_T s^1 + N^2 \Lambda^2 K_T s^2 = N(K_T s^1 + K_T s^2) = NK_T z \tag{37}$$

where the last equality is true because $z = s^1 + s^2$. Therefore, $\mathbf{V}_T(z) = \hat{\mathbf{V}}_T(z, s^1, s^2)$ $\forall z, s^1, s^2 \in \mathbb{R}^{d_x}$ such that $z = s^1 + s^2$.

Now, assume that $\mathbf{V}_{k+1}(z) = \hat{\mathbf{V}}_{k+1}(z, s^1, s^2)$ $\forall z, s^1, s^2 \in \mathbb{R}^{d_x}$ such that $z = s^1 + s^2$. We need to show that $\mathbf{V}_k(z) = \hat{\mathbf{V}}_k(z, s^1, s^2)$ $\forall z, s^1, s^2 \in \mathbb{R}^{d_x}$ with $z = s^1 + s^2$. For this, note that one can use dynamic

programming arguments to write the cost-to-go functions $\mathbf{V}_k$ and $\hat{\mathbf{V}}_k$ in terms of instantaneous cost and the next stage cost-to-go functions:

$$\mathbf{V}_k(z) = \mathrm{E}[(MX_t + NU_t)^\intercal(MX_t + NU_t)|Z_k = z, U_k = K_k z] + \mathrm{E}[\mathbf{V}_{k+1}(Z_{k+1})|Z_k = z, U_k = K_k z], \quad (38)$$

and

$$\hat{\mathbf{V}}_k(z, s^1, s^2) = \mathrm{E}[(MX_t + NU_t)^\intercal(MX_t + NU_t)|Z_k = z, S_k^1 = s^1, S_k^2 = s^2, U_k^1 = \Lambda^1 K_k s^1, U_k^2 = \Lambda^2 K_k s^2] + \mathrm{E}[\hat{\mathbf{V}}_{k+1}(Z_{k+1}, S_{k+1}^1, S_{k+1}^2)|Z_k = z, S_k^1 = s^1, S_k^2 = s^2, U_k^1 = \Lambda^1 K_k s^1, U_k^2 = \Lambda^2 K_k s^2]. \quad (39)$$

The first expectation on the right hand side of (38) can be shown to be equal to the first expectation on the right hand side of (39) by repeating the arguments used at time T. Using Lemma 4, the second expectation on the right hand side of (38) can be written as,

$$\mathrm{E}[\mathbf{V}_{k+1}(Z_{k+1})|Z_k = z, U_k = K_k z]$$
$$= \mathrm{E}[\mathbf{V}_{k+1}\big((I - L_{k+1}C)(AZ_k + BU_k) + L_{k+1}(CX_{k+1} + V_{k+1})\big)|Z_k = z, U_k = K_k z]$$
$$= \mathrm{E}[\mathbf{V}_{k+1}\big((I - L_{k+1}C)(A + BK_k)z + L_{k+1}(CAX_k + CBK_k z + CW_k + V_{k+1})\big)|Z_k = z, U_k = K_k z]$$
$$= \mathrm{E}[\mathbf{V}_{k+1}\big((A + BK_k)z + L_{k+1}CA(X_k - z) + L_{k+1}(CW_k + V_{k+1})\big)|Z_k = z]. \quad (40)$$

Furthermore, because of the induction hypothesis, the second expectation on the right hand side of (39) can be written as,

$$\mathrm{E}[\hat{\mathbf{V}}_{k+1}(Z_{k+1}, S_{k+1}^1, S_{k+1}^2)|Z_k = z, S_k^1 = s^1, S_k^2 = s^2, U_k^1 = \Lambda^1 K_k s^1, U_k^2 = \Lambda^2 K_k s^2]$$
$$= \mathrm{E}[\mathbf{V}_{k+1}(Z_{k+1})|Z_k = z, S_k^1 = s^1, S_k^2 = s^2, U_k^1 = \Lambda^1 K_k s^1, U_k^2 = \Lambda^2 K_k s^2]. \quad (41)$$

(41) can be further written as

$$\mathrm{E}[\mathbf{V}_{k+1}\big((I - L_{k+1}C)(Az + B^1 U_k^1 + B^2 U_k^2) + L_{k+1}(CAX_k + CB^1 U_k^1 + CB^2 U_k^2 + CW_k + V_{k+1})\big)|Z_k = z, S_k^1 = s^1, S_k^2 = s^2, U_k^1 = \Lambda^1 K_k s^1, U_k^2 = \Lambda^2 K_k s^2]. \quad (42)$$

From the substitutability assumption (Assumption 1) and Lemma 2, for any vector u, $Bu = B^i l^i(u) = B^i \Lambda^i u$. Therefore,

$$B^1 \Lambda^1 K_k s^1 = BK_k s^1,$$
$$B^2 \Lambda^2 K_k s^2 = BK_k s^2. \quad (43)$$

(42) can now be written as

$$\mathrm{E}[\mathbf{V}_{k+1}\big((I - L_{k+1}C)(Az + BK_k s^1 + BK_k s^2) + L_{k+1}(CAX_k + CBK_k s^1 + CBK_k s^2 + CW_k + V_{k+1})\big)|Z_k = z, S_k^1 = s^1, S_k^2 = s^2, U_k^1 = \Lambda^1 K_k s^1, U_k^2 = \Lambda^2 K_k s^2] = \mathrm{E}[\mathbf{V}_{k+1}\big((Az + BK_k z) + L_{k+1}CA(X_k - z) + L_{k+1}(CW_k + V_{k+1})\big)|Z_k = z]. \quad (44)$$

(44) is the same as (40). Therefore, $\mathbf{V}_k(z) = \hat{\mathbf{V}}_k(z, s^1, s^2) \ \forall z, s^1, s^2 \in \mathbb{R}^{d_x}$ such that $z = s^1 + s^2$.

Now, the expected cost under the centralized control strategy, $U_t = K_t Z_t$, can be written as,

$$\mathrm{E}\left[\sum_{t=1}^{T} c(X_t, K_t Z_t)\right]$$
$$= \mathrm{E}\left[\mathrm{E}\left[\sum_{t=1}^{T} c(X_t, K_t Z_t)\Big|Z_1, U_1 = K_1 Z_1\right]\right]$$
$$= \mathrm{E}[\mathbf{V}_1(Z_1)], \quad (45)$$

while the expected cost under the decentralized strategies of Theorem 2 can be written as

$$\mathrm{E}\left[\sum_{t=1}^{T} c\left(X_t, \begin{bmatrix} \Lambda^1 K_t S_t^1 \\ \Lambda^2 K_t S_t^2 \end{bmatrix}\right)\right]$$
$$= \mathrm{E}\Big[\mathrm{E}\big[\sum_{t=1}^{T} c\left(X_t, \begin{bmatrix} \Lambda^1 K_t S_t^1 \\ \Lambda^2 K_t S_t^2 \end{bmatrix}\right)\Big|Z_1, S_1^1, S_1^2, U_1^1 = \Lambda^1 K_1 S_1^1, U_1^2 = \Lambda^2 K_1 S_1^2\big]\Big]$$
$$= \mathrm{E}[\hat{\mathbf{V}}_1(Z_1, S_1^1, S_1^2)] \quad (46)$$

Because $\mathbf{V}_1(z) = \hat{\mathbf{V}}_1(z, s^1, s^2) \ \forall z, s^1, s^2$ such that $z = s^1 + s^2$, (45) and (46) are equal. Thus, the decentralized control strategies of Theorem 2 achieve the same expected cost as the optimal centralized strategies.

V. CONCLUDING REMARKS

We considered a decentralized system with multiple controllers and defined a property called substitutability of one controller by another in open-loop strategies. For the LQG problem, our results show that, under the substitutability assumption, linear strategies are optimal and we provide a complete state space characterization of optimal strategies. Our results suggest that open-loop substitutability can work as a counterpart of the information structure requirements that enable simplification of decentralized control problems.

REFERENCES

[1] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.
[2] G. M. Lipsa and N. C. Martins, "Optimal memoryless control in Gaussian noise: A simple counterexample," *Automatica*, vol. 47, no. 3, pp. 552–558, 2011.
[3] V. D. Blondel and J. N. Tsitsiklis, "A survey of computational complexity results in systems and control," *Automatica*, vol. 36, no. 9, pp. 1249–1274, 2000.

[4] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Boston, MA: Birkhäuser, 2013.

[5] A. Mahajan, N. Martins, M. Rotkowitz, and S. Yuksel, "Information structures in optimal decentralized control," in *IEEE Conference on Decision and Control*, 2012, pp. 1291–1306.

[6] Y.-C. Ho and K.-C. Chu, "Team decision theory and information structures in optimal control problems–Part I," vol. 17, no. 1, pp. 15–22, 1972.

[7] S. Yüksel, "Stochastic nestedness and the belief sharing information pattern," pp. 2773–2786, Dec. 2009.

[8] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," vol. 51, no. 2, pp. 274–286, 2006.

[9] B. Bamieh and P. Voulgaris, "A convex characterization of distributed control problems in spatially invariant systems with communication constraints," *Systems and Control Letters*, vol. 54, no. 6, pp. 575–583, 2005.

[10] L. Lessard and S. Lall, "Internal quadratic invariance and decentralized control," in *American Control Conference (ACC), 2010*, June 2010, pp. 5596–5601.

[11] L. Lessard, "Tractability of complex control systems," Ph.D. dissertation, Stanford University, 2011.

[12] S. M. Asghari and A. Nayyar, "Decentralized control problems with substitutable actions," in *IEEE 54th Annual Conference on Decision and Control (CDC), 2015*, Dec 2015.

[13] A. Ben-Israel and T. Greville, *Generalized Inverses: Theory and Applications*, ser. CMS Books in Mathematics. Springer New York, 2006. [Online]. Available: https://books.google.com/books?id=abEPBwAAQBAJ

[14] P. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control*. Prentice-Hall, 1986.