Control Communication Complexity of Distributed Actions

Wing Shing Wong and John Baillieul

Abstract

Recent papers have treated *control communication complexity* in the context of information-based, multiple agent control systems including nonlinear systems of the type that have been studied in connection with quantum information processing. The present paper continues this line of investigation into a class of two-agent distributed control systems in which the agents cooperate in order to realize common goals that are determined via independent actions undertaken individually by the agents. A basic assumption is that the actions taken are unknown in advance to the other agent. These goals can be conveniently summarized in the form of a *target matrix*, whose entries are computed by the control system responding to the choices of inputs made by the two agents. We show how to realize such target matrices for a broad class of systems that possess an input-output mapping that is bilinear. One can classify control-communication strategies, known as *control protocols*, according to the amount of information sharing occurring between the two agents. Protocols that assume no information sharing on the inputs that each agent selects and protocols that allow sufficient information sharing for identifying the common goals are the two extreme cases. Control protocols will also be evaluated and compared in terms of cost functionals given by integrated quadratic functions of the control inputs. The minimal control cost of the two classes of control protocols are analyzed and compared. The difference in the control costs between the two classes reflects an inherent trade-off between communication complexity and control cost.

Index Terms

Information-based control system, Control communication complexity, Brockett-Heisenberg system

W. S. Wong is with the Chinese University of Hong Kong, email: wswong@ie.cuhk.edu.hk. John Baillieul is with the Boston University, email: johnb@bu.edu. Research work is partially supported by the Research Grants Council of the Hong Kong Special Administrative Region under project 417207 and 417105. Also, JB acknowledges support from ODDR&E MURI07 Program Grant Number FA9550-07-1-0528, and the National Science Foundation ITR Program Grant Number DMI-0330171.

1. INTRODUCTON

In [34] and [35] the authors proposed the concept of *control communication complexity* as a formal approach for studying a group of distributed agents exercising independent actions to achieve common goals. For distributed cooperative systems, it is natural to expect that communication can help improve system performance, such as reducing the control cost. In this paper, we demonstrate how the concept of control communication complexity can lead to an inherent estimate of the value of communication bits in reducing the control cost.

Information-based control theory aims to deal with systems in which the interplay between control and communication are closely intertwined. Over the past decade a substantial volume of literature has appeared on this topic, focusing mostly on issues that arise due to the locality separation between a dynamic system and a single decision-maker, (for some early work see for example [12], [36], [37], [14], [15], [22], [28], and [32].) In this paper, we investigate information-based systems controlled by two distributed agents. In particular, we focus on cooperative control, the goal of which is for the agents, Alice and Bob, to induce a system output that depends jointly on the controls they independently select from their respective finite sets of control inputs.

The concept of multiple distributed selection of control actions from specified set of possible choices has not received much attention in the control literature until recent work by the authors. The significance of this new perspective can be illustrated by a generalization of the rendezvous problem in operations research. (See for example [1] and also [18] for an early team-theoretic perspective.) The problem seeks to find touring strategies for two agents to meet as quickly as possible, as exemplified by the scenario of a mother and her child separated in a crowded park. For that particular problem, a practical solution, not necessarily optimal, is to ask the child to stay in one place and for the mother to conduct a complete tour of the park. For a more complex situation, consider Alice and Bob who jog in the same park at the same time every day. By tacit understanding, they wish their paths to cross or not to cross according to their moods (the choices) as prescribed in the following table

	Alice in good mood	Alice in bad mood
Bob in good mood	Paths cross	Paths do not cross
Bob in bad mood	Paths do not cross	Paths do not cross

Table 1

If Alice and Bob can call each other to communicate their choice of inputs, the problem is trivial to solve. However, if direct communication is not available or allowed, a basic question is whether it is possible for Alice and Bob to follow different tour paths based on their moods to accomplish the stated objective. Moreover, if multiple feasible solutions exist, we are interested in identifying those that are optimal with regard to some performance measures.

Although Table 1 bears resemblance to the payoff functions in classical game theory, one cannot over-emphasize the point that there is no optimization of the table values in our problem formulation.

A second example that is more closely aligned to the model studied here is the mobile sensor network positioning problem introduced in [34]. Consider a mobile sensor network, such as a network of remote sensing satellites. The sensor network serves two agents, Alice and Bob; each of them wants to monitor a geographic region of interest selected from a pre-defined list. The agents do not communicate to each other directly; in fact, they may not even know the identity of the other party. If for any given ordered pair of choices there is an optimal configuration for the sensors, it is natural to investigate whether one can design communication and control strategies, known as *control protocols*, for the agents to jointly maneuver the mobile sensors to the optimal configuration based on their own individual selected choice. Moreover, if such feasible control protocols do exist, a second objective is to find the optimal protocols under appropriately defined performance measures.

To fix ideas for subsequent discussions, throughout the paper we index the finite collection of control inputs that Alice can send to the system by a set of labels $\mathcal{A} \equiv \{1, \ldots, m\}$. SImilarly, a finite set $\mathcal{B} \equiv \{1, \ldots, n\}$ is used to label the controls available to Bob. It is assumed that over many enactments of the protocol the inputs used by Alice and Bob will appear to be randomly chosen samples from uniformly distributed random variables with sample spaces cA and cB. If one represents the target output when Alice chooses $\alpha = i$ and Bob chooses $\beta = j$ by $\mathbf{H}_{ij}$, then the m-by-n matrix

$$\mathbf{H} = [\mathbf{H}_{ij}] \tag{1.1}$$

provides a compact representation of the set of target outputs for all possible choices of inputs and will be referred to as the *target matrix*.

While the structure of control protocols will be explained in the following section, a basic observation is that the input labels of the agents, $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$, are key inputs to these protocols. Once α and β are specified, following the basic premise of [19], the control protocol is assumed to run to completion. For a control protocol, $\mathcal{P}$, let $\mathbf{x}(\mathcal{P}(i,j),t)$ represent the state at time t when $\alpha = i$ and $\beta = j$. If the system output mapping is represented by F, the *feasibility problem* of *protocol-realizing control* is to determine whether it is possible to design a control protocol, $\mathcal{P}$, so that at the termination time, T, the following condition is satisfied for all $i \in \{1,\ldots,m\}$ and $j \in \{1,\ldots,n\}$:

$$F(\mathbf{x}(\mathcal{P}(i,j),T)) = \mathbf{H}_{ij}. \tag{1.2}$$

Solutions to the feasibility problem may involve control protocols that require observations of the system state; even though there is no direct communication links between the agents, it is possible for them to signal to one another via the dynamic system. An example of such a control protocol was discussed in [35]. The number of bits exchanged during the execution of a control protocol is a useful indicator of its complexity. On the other hand, control protocols can also be evaluated by means of the control cost incurred, for example as measured by the control energy required. Although communication complexity and control cost seem unrelated, we will demonstrate that there is a close relation between the two.

To make our results concrete, we will focus on a class of systems whose input-output mappings are bilinear. A prototypical example of such a system is the *Brockett-Heisenberg system*, which we hereafter refer to as the B-H system. The Heisenberg group, to which this system is related, has appeared in literature of one-dimensional problems in quantum mechanics, [33]. Brockett's interest, on the other hand, arose from his fundamental work on sub-Riemannian geometry. (See [4] and [27] for details.) For B-H systems one can characterize exactly when protocol-realizing control problems have feasible solutions in the absence of any communication between the agents. Moreover, when a problem is feasible it is possible to determine the minimal control cost needed to achieve it. On the other hand, in the case that the agents have partial information about each others choice of inputs, it will sometimes be possible to decompose the control communication problem into simpler parts on each of which a smaller control cost can be calculated. The concept of ϵ-signaling will be introduced, by means of which the agents may share partial information with each other. Using this idea, it will be possible to estimate the minimal control costs for two extreme scenarios, one without side communication or partial prior knowledge and the other with enough communication to allow the players to precisely compute intermediate results regarding the target matrix. The two extreme scenarios suggest a natural framework for appraising the inherent value of a communication bit.

The organization of the paper is as follows. In section 2, we provide a description of the basic model as well as the definition of key concepts. In particular, we introduce the idea of multi-round protocols. In section 3, background results on a bilinear input-output system are presented. In section 4 we describe how to transform the optimization of a single round protocol into a matrix optimization problem, the solution to which is presented in section 5. Implications of the result for understanding the trade-off issue between communication complexity and control cost is explained in section 6. Multi-round protocols are discussed in section 7. In particular, a special class of multi-round protocols known as *two-phase protocols* is studied in some detail. Section 8 provides a brief conclusion of the paper. Two short appendices are also included.

2. THE BASIC MODEL

The dynamical systems considered here are inherently continuous time systems of the form:

$$\begin{cases} \frac{d}{dt}\mathbf{x}(t) &= \mathbf{d}(\mathbf{x}(t), u(t), v(t)),\ \mathbf{x}(0) = \mathbf{x}_0 \in \mathbb{R}^N, \\ \mathbf{z}(t) &= \mathbf{c}(\mathbf{x}(t)) \in \mathbb{R}, \end{cases} \tag{2.1}$$

where $\mathbf{d}$ is an arbitrary smooth vector field and u and v are scalar control functions that, once chosen, are applied over a time interval of standard length T. The output $z(\cdot)$ is sampled at discrete time instants $t_0 < t_1 < \ldots$, where $t_0 = 0$ and $t_{k+1} = t_k + T$. Information about the state is made available to the two agents through encoded observations $\mathbf{b}_A(\mathbf{x}(t))$ and $\mathbf{b}_B(\mathbf{x}(t))$ which are made at the same time instants. The fact there are standard intervals of time over which observations and selected control inputs are applied allows the analysis to make contact with prior work on information-based control of discrete-time systems. Under our assumptions, we thus consider the following simplified version of the model introduced in [34]. For $k = 1, 2, \ldots$, let t_k represent the fixed time when observations are taken. Define

$$\mathbf{x}_k = \mathbf{x}(t_k). \tag{2.2}$$

Observations of the state, $\mathbf{b}_A(\mathbf{x}_k)$ and $\mathbf{b}_B(\mathbf{x}_k)$, are made available to each agent as encoded messages, $Q_k^{(A)}$ and $Q_k^{(B)}$, consisting of finite bit-length codewords; in other words, the ranges of these quantization functions are finite sets. Computation and communication delays in reporting observations to the agents are assumed to be negligible. The agents select control actions—respectively $u_k = P_k^A(Q_k^{(A)}(\mathbf{b}_A(\mathbf{x}_k)), \alpha)$ and $v_k = P_k^B(Q_k^{(B)}(\mathbf{b}_B(\mathbf{x}_k)), \beta)$—that may be viewed as random variables that are assumed to be measurable with respect to the input selections $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$. For reasons of simplicity, we assume that the controls u and v are scalar functions. The codewords identifying the selected controls are transmitted to the dynamic system. Computation and communication delays associated with this step are also assumed to be negligible. Thus at time t_k the dynamic system can determine the control selected by Alice and Bob.

The state transition (2.1) between times t_k and t_{k+1} is thus described by the following discrete time control model where the controls are square integrable scalar functions:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{a}(\mathbf{x}_k, u_k, v_k), \quad \mathbf{x}(0) = \mathbf{x}_0 \in \mathbb{R}^N, \\ \mathbf{y}_k^{(A)} = \mathbf{b}_A(\mathbf{x}_k) \in \mathbb{R}^{\ell_A}, \ \mathbf{y}_k^{(B)} = \mathbf{b}_B(\mathbf{x}_k) \in \mathbb{R}^{\ell_B}, \\ u_k = P_k^A(Q_k^{(A)}(\mathbf{b}_A(\mathbf{x}_k)), \alpha), \ v_k = P_k^B(Q_k^{(B)}(\mathbf{b}_B(\mathbf{x}_k)), \beta), \\ \mathbf{z}_k = \mathbf{c}(\mathbf{x}_k) \in \mathbb{R}. \end{cases} \tag{2.3}$$

The quantity $\mathbf{z}_k = \mathbf{c}(\mathbf{x}_k)$ is a global system output that is observable to Alice, Bob, and possibly to exogenous observers as well. The protocol paramters, α and β, are specified at time t_0 and remain unchanged for all the time. The case where these parameters are allowed to change over time is an interesting extension which is not considered here.

Definition 2.1. *A control protocol, $\mathcal{P}$, consists of the functions:*
$\{Q_k^{(A)}\}_{k=0}^{\infty}, \{Q_k^{(B)}\}_{k=0}^{\infty}, \{P_k^{(A)}\}_{k=0}^{\infty}$, *and* $\{P_k^{(B)}\}_{k=0}^{\infty}$.

We define the epoch between time t_k and t_{k+1} as round $k+1$. In round $k+1$, an agent first observes and then selects a control to be applied to the system. The selected controls $u_k = P_k^A(Q_k^{(A)}(\mathbf{b}_A(\mathbf{x}_k)), \alpha)$ and $v_k = P_k^B(Q_k^{(B)}(\mathbf{b}_B(\mathbf{x}_k)), \beta)$ will typically be time-varying in the epoch between time t_k and t_{k+1}, but they are otherwise independent of the state. Hence, the agents use essentially open-loop controls during the round, but the selection of the control at the beginning of the round can depend on partial state information.

Definition 2.2. *Consider a dynamic system, $\mathbf{\Sigma}$, with parameters defined by $(\mathbf{a}, \mathbf{b}_A, \mathbf{b}_B, \mathbf{c}, \mathbf{x}_0)$. A target matrix $\mathbf{H}$ is said to be realizable at termination time T_f if there exists a k-round protocol,* *such that*

$$t_k = T_f,$$

and for any choice of indices $i \in \mathcal{A}$ and $j \in \mathcal{B}$, the k-step evolution prescribed by the protocol of (2.3) defined with $\alpha = i$ and $\beta = j$ terminates with

$$\mathbf{c}(\mathbf{x}_k) = \mathbf{H}_{ij}. \tag{2.4}$$

In this paper, the initial state is assumed to be fixed and known to the agents. Hence, for the first round, there is no need for any communication to the agents. That is, we assume the quantization functions $Q_0^{(A)}$ and $Q_0^{(B)}$ always take the same value for all control choices and need not be transmitted.

Because a target matrix may be realized by different protocols, there is an interest in identifying those that are optimal with respect to some performance measures. One such measure classifies a control protocol in terms of the number of rounds it takes to realize all the possible outputs listed in a target matrix. This can be regarded as a *time complexity*. A somewhat related measure counts the number of communication bits exchanged during the protocol execution. To make this precise, [34] introduced the concept of *control communication complexity* by extending the concept of *communication complexity* that was introduced in computer science by Yao [38]. Briefly speaking, given a dynamic system and a target matrix, one defines the protocol complexity of a feasible protocol to be the maximum number of bits exchanged by the agents in running the protocol to completion. The control communication complexity is then the minimum protocol complexity over the set of all feasible protocols. A caveat: unlike classical communication complexity, control communication complexity is defined with regard to a fixed dynamical system. A fundamental conceptual difference is that while communication complexity theory is aimed at analyzing rather than executing algorithms, the focus of control communication complexity as discussed in this paper is the optimal cost of running specific distributed computation algorithms.

In the models considered in the present paper, control inputs are square integrable functions, and this suggests measuring the complexity of a protocol in terms of the integral of a quadratic function of the control. This makes contact with performance measures commonly used in centralized control. Indeed, control communication complexity provides a rich new class of optimal control problems. While the relation between communication complexity and integral control cost may not be transparent, the analysis of the B-H system in [35] demonstrates that there is an intriguing relation between them. Intuition suggests that protocols in which there is limited communication between the agents *except for their common observations of the system dynamics* may require a larger integral control cost than those that employ a large number of communication bits in addition to the system observations. If the agents have partial information about each others choices of inputs, control laws can be more precisely tailored. A contribution of this paper is to take the first step towards analyzing this trade-off by comparing two limiting types of control protocols, namely, single round protocols which entail no communication bits, and protocols in which agents share partial information about which elements $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$ are governing the execution of the protocol. We provide a fairly complete treatment of single round protocols, but multi-round protocols are less well understood at this point. As a lower bound on the control cost, we assume that it is possible for the agents to communicate their choices of inputs to each other without incurring any control cost or state change, thereby decomposing the original problem into a sequence of standard optimal control problems. The difference in the minimal control costs between this type of protocol and the single round protocols having no communication provides an upper bound on the saving afforded by the bits of information that are communicated between the two agents as described in [34].

In the current model the only unknown parameter to Alice is Bob's choice of control input and vice versa the only unknown parameter to Bob is Alice's choice of input. The single choice problem in which each agent has only one choice of control input is equivalent to the scenario where the agents have complete information on the system. Since the initial state is known and there is no state noise, optimal distributed control can be devised without resorting to any state observations. In other words, for optimizing single choice problems, we only need to consider single round protocols.

For the rest of the paper we focus on scalar output functions to simplify the analysis. Vector-valued outputs, as necessitated by the sensor network example, present a significantly larger technical challenge, and will not be discussed here.

To analyze the cost of letting our control system evolve under different input curves in single round protocols, we lift the evolution dynamics (2.3) back up to the continuous domain as described by equation (2.1) but with $\mathbf{c}$ replaced by a scalar function.

Unless stated otherwise, the controls u and v exercised by the agents are assumed to lie in a closed subspace $\mathcal{L} \subset L^2[0, T]$. Let $\mathcal{L} \otimes \mathcal{L}$ represent the tensor product Hilbert space with inner product defined by

$$< u_1 \otimes v_1, u_2 \otimes v_2 >=< u_1, u_2 >< v_1, v_2 > . \tag{2.5}$$

At time T, the output of system (2.1) can be regarded as a functional from $\mathcal{L} \otimes \mathcal{L}$ to $\mathbb{R}$, denoted by F. F of course depends on the initial state $\mathbf{x}_0$ but as the state is assumed to be fixed and since for the time being we consider only single round protocols, this dependency can be hidden for simplification. Without loss of generality we assume that $t_0 = 0$ and $t_1 = T = 1$. (Related problems in which termination time cannot necessarily be normalized are treated in [31].)

F is a bounded functional if there exists a finite $\|F\|$ so that for all $(u, v) \in \mathcal{L} \times \mathcal{L}$,

$$|F(u, v)| \leq \|F\| \|(u \otimes v)\|_{\mathcal{L} \otimes \mathcal{L}} = \|F\| \|u\|_{\mathcal{L}} \|v\|_{\mathcal{L}}. \tag{2.6}$$

To realize a given target matrix $\mathbf{H}$, the optimal controls in general depend jointly on the parameters, α and β. However, if the agents make their choices independently and there is no communication between them, u can only depend on Alice's parameterα and v on Bob's parameter β..

A target matrix $\mathbf{H}$ is realized by a single round protocol $\mathcal{P}$ if there exist sets of controls, $\mathcal{U} = \{u_1, \ldots, u_m\}$ and $\mathcal{V} = \{v_1, \ldots, v_n\}$, so that

$$F(u_i, v_j) = H_{ij}. \tag{2.7}$$

We emphasize that F represents the system output at time T. Such a single round protocol solution may not always exist as we will see in subsequent sections. The problem may become feasible if Alice and Bob can exchange information about to their choices to each other.

Given the system (2.1) and the parameter sets $\mathcal{A}$ and $\mathcal{B}$, the cost of a single round control protocol $\mathcal{P}$ can be defined as an average of the required control energy, given explicitly by the formula,

$$I(\mathcal{U}, \mathcal{V}) = \frac{1}{m} \sum_{i=1}^{m} \int_0^1 u_i^2(t) dt + \frac{1}{n} \sum_{j=1}^{n} \int_0^1 v_j^2(t) dt. \tag{2.8}$$

One can also write equation (2.8) in the form,

$$I(\mathcal{U}, \mathcal{V}) = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \left(\int_0^1 u_i^2(t) dt + \int_0^1 v_j^2(t) dt \right), \tag{2.9}$$

which highlights the fact that the control cost is averaged over all possible event outcomes based on the control actions that are chosen by the agents.

In subsequent sections, we compute the minimum averaged control energy for an arbitrary target function $\mathbf{H}$. That is, our aim is to compute

$$\hat{C}_F(\mathbf{H}) \equiv \min_{\mathcal{U},\mathcal{V}\subset\mathcal{L}} I(\mathcal{U},\mathcal{V}) \tag{2.10}$$

subject to the constraints that for $i = 1, \dots, m$ and $j = 1, \dots, n$,

$$F(u_i, v_j) = H_{i,j}. \tag{2.11}$$

Before concluding this section we note that there are some similarities between the cooperative control communication protocols studied in this paper and more classical dynamic game strategies as studied in, say, [3]. Yet there are fundamental differences, for example, optimization of payoff function is not the focus of our investigation. Using the rendezvous problem as an example, once the moods of the the agents are fixed, the outcome to be achieved is automatically defined by the target matrix. The investigation focus is on how to ensure that the target objective—paths crossing or not crossing—can be guaranteed. The work here also has some connection with the many papers in the literature dealing with distributed control of mobile agents, multi-agent consensus problems, and classical team decision theory. See, for instance, [9],[10],[13],[16],[18], and [29]. In these papers the dynamics of the subsystems controlled by the agents are usually not tightly coupled and the control cost is not explicitly calculated, unlike the models we are considering here. Moreover, allowing agents to select controls from sets of standard inputs is also a fundamental point of departure.

3. BILINEAR INPUT-OUTPUT MAPPINGS: BACKGROUND INFORMATION

The simplest input-output mapping to investigate is probably the affine mapping related to the following system:

$$\begin{cases} \dfrac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + u(t)\mathbf{b}_1 + v(t)\mathbf{b}_2,\ \mathbf{x}_0 \in \mathbb{R}^N, \\ y_A(t) = y_B(t) = z(t) = \mathbf{c}^T\mathbf{x}(t) \in \mathbb{R}. \end{cases} \tag{3.1}$$

Since the control effects of the agents enter into the input-output mapping linearly and independent of each other, it is easy to see that if $F(u_i, v_k) = H_{i,k}$, $F(u_i, v_l) = H_{i,l}$, $F(u_j, v_k) = H_{j,k}$, and $F(u_j, v_l) = H_{j,l}$ for some u_i, u_j, v_k, and v_l, then

$$H_{i,k} - H_{j,k} = H_{i,l} - H_{j,l}. \tag{3.2}$$

Thus, there are severe restrictions on the structure of realizable target functions for such an input-output mapping. For further details on the affine case one can refer to [17].

A richer class of input-output mappings are the bilinear mappings:

Definition 3.1. *The system defined by equation (2.1) is a* bilinear input-output system *if for any time $t \geq t_0$, the output at t, regarded as a mapping $(u(\cdot), v(\cdot)) \mapsto z(\cdot)$ from $\mathcal{L} \otimes \mathcal{L}$ to $\mathbb{R}$ is bilinear in (u, v).*

An important example of such a bilinear input-output mapping is the Brocket-Heisenberg (B-H) system, [4],[5]. The system, denoted by $\mathbf{\Sigma}_B$, can be described as follows.

$$\begin{cases} \dfrac{d}{dt}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} u \\ v \\ vx - uy \end{pmatrix}, \begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \in \mathbb{R}^3, \\ y_A(t) = y_B(t) = z(t) \equiv c((x(t), y(t), z(t)). \end{cases} \tag{3.3}$$

The fact that the input-output mapping of this system is bilinear in controls u and v can be easily verified. There are other systems that define bilinear input-output mappings, for example:

$$\begin{cases} \dfrac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + u(t)v(t)\mathbf{b},\ \mathbf{x}(0) = \mathbf{0} \in \mathbb{R}^N, \\ y_A(t) = y_B(t) = z(t) = \mathbf{c}^T\mathbf{x}(t) \in \mathbb{R}. \end{cases} \tag{3.4}$$

While the input-output mapping is bilinear, (3.4) can be thought of as a linear system in which the input is the product of the control chosen by the two agents.

Both (3.3) and (3.4) are representative of a general class of bilinear input-output systems of the form depicted in Figure 1. As detailed in [8], the basic structure here plays a role in understanding the ways in which cyclic processes govern the dynamics of both engineered and natural systems. In what follows we shall develop a theory of distributed control in which pairs of agents select periodic inputs to steer bilinear systems in order to achieve common objectives. In studying such systems, we are interested in the dependence of the output z on inputs that can be described by trigonometric Fourier series of the form

$$\sum_{k=1}^{\infty}(\alpha_k \cos(2\pi kt) + \beta_k \sin(2\pi kt)). \tag{3.5}$$

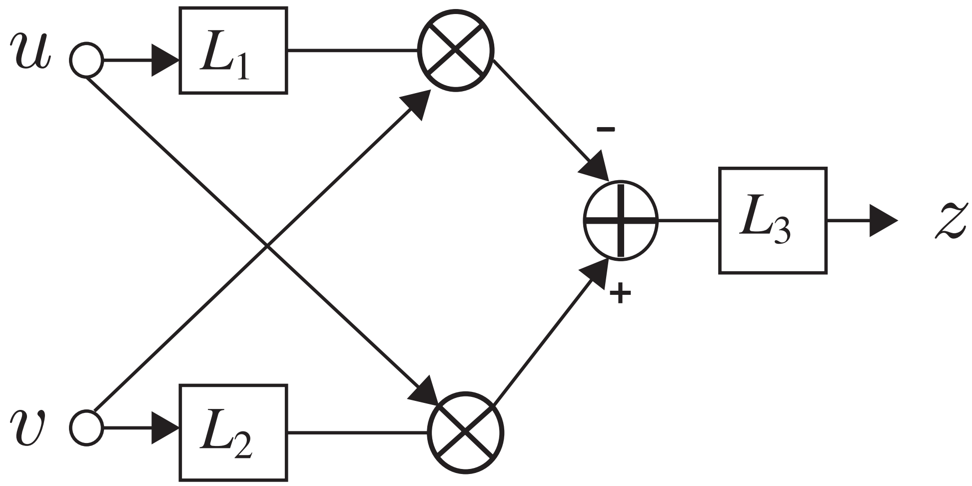


Fig. 1. Block diagram of a general bilinear input-output (i.o.) system. L_1, L_2, L_3 are linear operators, and in the case of the bilinear i.o. system (3.3), $L_i = \int$ for $i = 1, 2, 3$.

The members of the Fourier basis $\{\cos(2\pi kt), \sin(2\pi kt) \ : \ k = 1, 2, \dots\}$ appear as matrix coefficients in the canonical (real) irreducible representations of $SO(2)$:

$$\left\{ \begin{pmatrix} \cos(2\pi nt) & -\sin(2\pi nt) \\ \sin(2\pi nt) & \cos(2\pi nt) \end{pmatrix} \ : \ n = 1, 2, \dots \right\},$$

and by the Peter-Weyl theorem for compact groups we know that this Fourier basis is complete in the space of periodic (period-one) functions in $L^2[0, 1]$ (which can be thought of as $L^2[SO(2)]$). With respect to this basis, the Hilbert space $L^2[SO(2)]$ decomposes into the direct sum of two dimensional subspaces V_k spanned by $\{\cos(2\pi kt), \sin(2\pi kt)\}$.

For systems of the form depicted in Figure 1 with control inputs having the form (3.5) we make the following assumptions:

1) For each k, the operators L_1 and L_2 leave the subspaces V_k invariant, and
2) When restricted to V_k, the operators L_1 and L_2 are invertible and normal.

These assumptions are easy to check in the examples we consider, and they will allow us to conclude that a certain matrix representation of (3.3) is *strongly regular* in the sense that we define below.

In the case of (3.4), we take $L_1 = 0$, $L_2 : L^2[SO(2)] \to L^2[SO(2)]$ is the identity mapping, and L_3 is the bounded operator given by

$$w \mapsto \int_0^t \mathbf{c}^T e^{(t-s)} w(s)\, ds.$$

For (3.3), we take $L_1 = L_2 = L : L^2[SO(2)] \to L^2[SO(2)]$ to be bounded operators given by

$$w \mapsto \int_0^t w(s)\, ds.$$

$L_3 : L^2[SO(2)] \to \mathbb{R}$ is given by

$$w \mapsto \int_0^1 w(t)\, dt.$$

While our main results make use of the general properties of systems of the form shown in Figure 1 and to some extent of properties that are particular to (3.3), it is important to note that the system (3.3) has features in common with a much larger class of two-input "drift-free" control systems whose output mappings are not necessarily bilinear functionals of the inputs. These systems arise in applications including the kinematic control of nonholonomic wheeled vehicles (See, e.g. [21].) and the control of ensembles of spin systems arising in coherent spectroscopy and quantum information processing. (See [24].) We refer the reader to [35] for other examples of such two-input systems and to [11], [23], [24], and [25] for details regarding the control of Bloch equations and other applications to quantum mechanics.

To illustrate how important features of a broad class of two-input systems are held in common with (3.3), we consider the Bloch equations in a spatially fixed frame in which there is an rf-inhomogeneity but no Larmor dispersion. The equations for this system are:

$$\dot{X} = [u(t)\Omega_y + v(t)\Omega_x]X, \quad X(0) = I, \tag{3.6}$$

where

$$\Omega_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \Omega_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \text{and} \quad \Omega_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

are the infinitessimal generators of rotations about the x-,y-, and z-axes in $\mathbb{R}^3$. This system does not have a bilinear form as depicted in Figure 1, but input pairs defined for (3.3) can be used to steer (3.6) in useful ways. To briefly explore the connection between (3.3) and (3.6), we recall the *area rule*.

Lemma 3.1. (Brockett, [6].) *Let $u, v \in L^2[SO(2)]$ be as above, and let x, y be defined by (3.3). Then $x, y \in L^2[SO(2)]$, and $z(\cdot)$ as defined by (3.3) satisfies*

$$z(1) = \int_0^1 v(t)x(t) - u(t)y(t)\,dt = \oint x\,dy - y\,dx = 2A,$$

where A is the signed area enclosed by $(x(t), y(t))$.

The connection between (3.3) and (3.6) is made by the following.

Example 3.1. Let the inputs

$$u(t) = \epsilon\cos(2\pi t), \quad v(t) = \epsilon\sin(2\pi t)$$

be applied to both (3.3) and (3.6). For (3.3), corresponding to these inputs,

$$x(t) = \frac{\epsilon}{2\pi}\sin(2\pi t) \quad \text{and} \quad y(t) = -\frac{\epsilon}{2\pi}\cos(2\pi t).$$

By Lemma 3.1 (or a direct calculation), the output of (3.3) is $z(1) = \epsilon^2/(2\pi)$. For this choice of input, (3.6) may be solved explicitly yielding

$$X(t) = \exp\left[\begin{pmatrix} 0 & -2\pi & 0 \\ 2\pi & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} t\right] \exp\left[\begin{pmatrix} 0 & 2\pi & \epsilon \\ -2\pi & 0 & 0 \\ -\epsilon & 0 & 0 \end{pmatrix} t\right].$$

Note that the first factor in this expression has period 1, while the second factor is periodic with period $T = 1/\sqrt{1 + (\epsilon/(2\pi))^2}$. At time T,

$$X(T) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where $\theta = -(1/2)(\epsilon^2/(2\pi)) + 0(\epsilon^2)$. Hence, the output of (3.3) provides a good approximation of the z-axis rotation $\theta = -z(1)/2$ resulting from application the inputs u, v to (3.6) over the time interval $[0, 1/\sqrt{1 + (\epsilon/(2\pi))^2}]$. We refer to [35] for further information regarding area rules and approximate motions of (3.6) on $SO(3)$ and to [21] and [7] for information on the general application of area rules in two-input control systems. As detailed in [25], systems of the form (3.6) can be used to guide the design of rf pulses for quantum control experiments. While (3.6) does not have the same natural bilinear input-output structure as (3.3) the approximate relationship between the two system responses to closed curve inputs augurs well for potential applications of the computability results to be established in the remainder of the paper.

With such broader applications in mind, we confine our attention in the remainder of the paper to studying the response of (3.3) to those closed input curves $(u(t), v(t))$ that give rise to closed curves $(x(t), y(t))$. Let $\mathcal{L}$ be the closed subspace of $L^2[0,1]$ consisting of functions that can be represented by the following type of Fourier series with square summable coefficients:

$$\begin{aligned} u_i(t) &= \sqrt{2}\sum_{k=1}^{\infty}\left[a_{i,2k-1}\sin(2\pi kt) + a_{i,2k}\cos(2\pi kt)\right], \\ v_j(t) &= \sqrt{2}\sum_{k=1}^{\infty}\left[-b_{j,2k-1}\cos(2\pi kt) + b_{j,2k}\sin(2\pi kt)\right]. \end{aligned} \tag{3.7}$$

This restriction is not essential to the investigation reported in this paper, but it allows connection to the results presented in [35] and [2]. Note that $\mathcal{L}$ contains continuous periodic functions with zero mean and period one.

From basic orthogonality properties of the sine and cosine functions, one can show:

Lemma 3.2. *If the controls u_i and v_j are used, then at time $t = 1$,*

$$z(1) = \sum_{k=1}^{\infty}\frac{a_{i,2k-1}b_{j,2k-1}}{\pi k} + \sum_{k=1}^{\infty}\frac{a_{i,2k}b_{j,2k}}{\pi k}. \tag{3.8}$$

Moreover,

$$\int_0^1 u_i^2(t)dt = \sum_{k=1}^{\infty}(a_{i,2k-1}^2 + a_{i,2k}^2), \quad \int_0^1 v_j^2(t)dt = \sum_{k=1}^{\infty}(b_{j,2k-1}^2 + b_{j,2k}^2). \tag{3.9}$$

The proof is a straightforward calculation and is omitted. Define matrices $\mathbf{U_B}$, $\mathbf{V_B}$, and $\mathbf{F_B}$ as follows:

$$\mathbf{U_B} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots \\ \vdots & \vdots & \vdots & \ddots \\ a_{m,1} & a_{m,2} & a_{m,3} & \dots \end{bmatrix}, \mathbf{V_B} = \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} & \dots \\ b_{2,1} & b_{2,2} & b_{2,3} & \dots \\ \vdots & \vdots & \vdots & \ddots \\ b_{n,1} & b_{n,2} & b_{n,3} & \dots \end{bmatrix}, \tag{3.10}$$

$$\mathbf{F_B} \equiv \frac{1}{\pi}\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 2^{-1} & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 2^{-1} & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 3^{-1} & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 3^{-1} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{3.11}$$

Lemma 3.3. *The output of the B-H system is given by* $\mathbf{U_B F_B V_B}^T$. *Moreover,*

$$\sum_{i=1}^{m} \int_0^1 u_i^2(t)dt = tr\mathbf{U_B U_B}^T, \quad \sum_{i=1}^{n} \int_0^1 v_i^2(t)dt = tr\mathbf{V_B V_B}^T. \tag{3.12}$$

The proof is straightforward and is omitted. It should be noted that the weight assignment and the indexing in (3.7) are chosen to allow the matrices $\mathbf{U_B}, \mathbf{V_B}$, and $\mathbf{F_B}$ to assume these simple representations. The control cost defined in (2.8) can be rewritten as

$$I(\mathcal{U}, \mathcal{V}) = \left(\frac{1}{m} tr\mathbf{U_B U_B}^T + \frac{1}{n} tr\mathbf{V_B V_B}^T \right). \tag{3.13}$$

To extend these results to a general, bounded functional, F, on $\mathcal{L} \times \mathcal{L}$, let $\{e_1, e_2, \ldots\}$ and $\{f_1, f_2, \ldots\}$ be orthonormal bases (possibly the same) for $\mathcal{L}$ with respect to the standard inner product on $L^2[0,1]$. If $u = \sum_{i=1}^{\infty} r_i e_i$ and $v = \sum_{j=1}^{\infty} s_j f_j$, by the bounded bilinear

$$F(u, v) = F(\sum_{i=1}^{\infty} r_i e_i, \sum_{j=1}^{\infty} s_j f_j) = \sum_{i,j=1}^{\infty} r_i s_j F(e_i, f_j). \tag{3.14}$$

We can represent u and v as infinite dimensional row vectors in terms of their coefficients with respect to these orthonormal bases and F as an infinite dimensional matrix:

$$\mathbf{F} \equiv \begin{bmatrix} F(e_1, f_1) & F(e_1, f_2) & F(e_1, f_3) & \ldots \\ F(e_2, f_1) & F(e_2, f_2) & F(e_2, f_3) & \ldots \\ F(e_3, f_1) & F(e_3, f_2) & F(e_3, f_3) & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{3.15}$$

Then, the condition that the targets are achievable can be expressed by the equation:

$$\mathbf{H} = \mathbf{UFV}^T, \tag{3.16}$$

where $\mathbf{U}$ is the matrix whose i row is the vector representation of u_i and similarly, $\mathbf{V}$ is the matrix whose j row is the vector representation of v_j.

As usual, define the rank of $\mathbf{F}$ to be the number of its independent columns (or rows). The rank of $\mathbf{F}$ is said to be *infinite* if there is no finite subset of columns (rows) in terms of which all columns (rows) can be expressed as linear combinations. One can show via arguments presented in Appendix A that the rank is independent of the bases chosen. Hence, we can speak of the rank of a bilinear map without ambiguity.

Proposition 3.1. *There exists a single round protocol that realizes an m-by-n target matrix,* $\mathbf{H}$*, of rank k, if and only if the rank of the bilinear input-output mapping F is at least k.*

Proof: Let $\mathbf{H}_k$ be a full rank k-by-k submatrix of $\mathbf{H}$. If there is a single round protocol to realize $\mathbf{H}$, we can select controls from the solution protocol to obtain a realization for $\mathbf{H}_k$. Conversely, if a single round protocol exists for $\mathbf{H}_k$, we can extend it to obtain a protocol for realizing $\mathbf{H}$ by adding control functions to the control sets according to the additional choices allowed; these added controls can be constructed as linear combinations of the original controls used in realizing $\mathbf{H}_k$. Hence, we can assume without loss of generality that $m = n = k$ and $\mathbf{H}$ is a k-by-k full rank matrix.

Suppose a single round protocol exists for such a target matrix. The solutions consist of k controls for Alice and k for Bob. We can construct an orthonormal basis for $\mathcal{L}$ so that the controls used by Alice are spanned by the first k basis elements. Construct similarly a basis for Bob. Represent F in terms of these bases and let $\hat{\mathbf{F}}$ denote the restriction of F to the first k basis elements. Then, $\hat{\mathbf{F}}$ is a k-by-k matrix and the equation

$$\mathbf{H} = \mathbf{U}\hat{\mathbf{F}}\mathbf{V}^T \tag{3.17}$$

has a solution. It follows that the rank of $\hat{\mathbf{F}}$, and hence, F is at least k.

Conversely, if there exists a k-by-k matrix $\hat{\mathbf{F}}$ of full rank representing the restriction of F to some control subspaces, equation (3.17) always has a solution for any k-by-k matrix, $\mathbf{H}$. ■

The previous proposition implies that the existence of a target realizing single round protocol depends on the rank of the bilinear input-output mapping and the rank of the target matrix. For a bilinear input-output mapping of finite rank, there exists target matrices that cannot be realized without communication between the agents. For a bilinear input-output mapping of infinite order, any finite dimensional target matrix can be realized by a single round protocol.

Given a matrix representation, $\mathbf{F}$, for $l = 1, 2, \ldots$, define its *l-dimensional leading principal minors* by:

$$\mathbf{F}_l = \begin{bmatrix} F_{1,1} & \ldots & F_{1,l} \\ \vdots & \ddots & \vdots \\ F_{l,1} & \ldots & F_{l,l} \end{bmatrix}. \tag{3.18}$$

Definition 3.2. *A matrix representation of a bilinear map is regular if it is of infinite rank and for all positive integer l, its l-dimensional leading principal minors are nonsingular.*

For subsequent discussions we adopt the following notation. For an l-by-l matrix, $\mathbf{M}$, let $\sigma_i(\mathbf{M})$ represent its i-th singular value under the ordering

$$\sigma_1(\mathbf{M}) \geq \sigma_2(\mathbf{M}) \ldots \geq \sigma_l(\mathbf{M}). \tag{3.19}$$

If $\mathbf{M}$ is symmetric, represent by $\lambda_i(\mathbf{M})$ the i-th eigenvalue under the ordering

$$\lambda_1(\mathbf{M}) \geq \lambda_2(\mathbf{M}) \ldots \geq \lambda_l(\mathbf{M}). \tag{3.20}$$

Note that for a rectangular $\mathbf{M}$, say of dimensions p-by-q, there are p singular values for $\mathbf{M}$, while $\mathbf{M}^T$ has q singular values. It is well-known that the nonzero singular values of the two matrices are identical (see [26] for details).

Lemma 3.4. *Consider a k-by-k matrix, $\mathbf{M}$, with a matrix decomposition so that*

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{1,1} & \mathbf{M}_{1,2} \\ \mathbf{M}_{2,1} & \mathbf{M}_{2,2} \end{bmatrix}, \tag{3.21}$$

where $\mathbf{M}_{1,1}$ is l-by-l, $\mathbf{M}_{1,2}$ is l-by-$(k-l)$, $\mathbf{M}_{2,1}$ is $(k-l)$-by-l, and $\mathbf{M}_{2,2}$ is $(k-l)$-by-$(k-l)$. Then, for $i = 1, \ldots, l$

$$\sigma_i(\mathbf{M}) \geq \sigma_i(\mathbf{M}_{1,1}). \tag{3.22}$$

Proof: Consider the following positive semi-definite matrix:

$$\mathbf{M}\mathbf{M}^T = \begin{bmatrix} \mathbf{M}_{1,1}\mathbf{M}_{1,1}^T + \mathbf{M}_{1,2}\mathbf{M}_{1,2}^T & \mathbf{M}_{1,1}\mathbf{M}_{2,1}^T + \mathbf{M}_{1,2}\mathbf{M}_{2,2}^T \\ \mathbf{M}_{2,1}\mathbf{M}_{1,1}^T + \mathbf{M}_{2,2}\mathbf{M}_{1,2}^T & \mathbf{M}_{2,1}\mathbf{M}_{2,1}^T + \mathbf{M}_{2,2}\mathbf{M}_{2,2}^T \end{bmatrix}. \tag{3.23}$$

It follows from Fischer's Minimax Theorem (page 510, A.1.c [26]) that for $i = 1, \ldots, l$

$$\lambda_i(\mathbf{M}\mathbf{M}^T) \geq \lambda_i(\mathbf{M}_{1,1}\mathbf{M}_{1,1}^T + \mathbf{M}_{1,2}\mathbf{M}_{1,2}^T). \tag{3.24}$$

By means of a result of Loewner (page 510, A.1.b [26]):

$$\lambda_i(\mathbf{M}_{1,1}\mathbf{M}_{1,1}^T + \mathbf{M}_{1,2}\mathbf{M}_{1,2}^T) \geq \lambda_i(\mathbf{M}_{1,1}\mathbf{M}_{1,1}^T). \tag{3.25}$$

The result then follows by combining these two inequalities. ■

It follows that for all i,

$$\sigma_i(\mathbf{F}_l) \leq \sigma_i(\mathbf{F}_{l+1}). \tag{3.26}$$

On the other hand, it is shown in Appendix B that

$$\sigma_i(\mathbf{F}_l) \leq \|F\| \tag{3.27}$$

for all i. Hence the limit $\lim_{l\to\infty}\sigma_i(\mathbf{F}_l)$ exists and is finite. Denote this limit by $\sigma_i(\mathbf{F})$.

Definition 3.3. *A matrix representation of a bilinear map is strongly regular if it is regular and if for all i, there exists an index l_i such that*

$$\sigma_i(\mathbf{F}_{l_i}) = \sigma_i(\mathbf{F}). \tag{3.28}$$

For illustration, for a bilinear map F with a diagonal matrix representation, the square of the diagonal entries must be in non-increasing order in order to satisfy the above condition.

For the B-H system, the input-output mapping that takes (u, v) to $z(1)$ by means of (3.3) has a matrix representation $\mathbf{F}_B$ described by (3.11) with respect to the orthonormal basis,

$$\mathcal{B}_0 = \{\sqrt{2}\sin(2\pi t), \sqrt{2}\cos(2\pi t), \sqrt{2}\sin(4\pi t), \sqrt{2}\cos(4\pi t), \ldots\}. \tag{3.29}$$

The matrix is diagonal, of infinite rank, and provides a strongly regular representation of F.

4. MATRIX REPRESENTATION OF THE OPTIMIZATION PROBLEM

The control cost of $\mathcal{U}$ and $\mathcal{V}$ defined in (2.8) can be rewritten in matrix form as

$$\frac{1}{m} tr\mathbf{U}\mathbf{U}^T + \frac{1}{n} tr\mathbf{V}\mathbf{V}^T. \tag{4.1}$$

Our goal of finding the minimum cost controls to realize a given target matrix can now be represented as an infinite dimensional matrix optimization problem. One solution approach is to consider finite dimensional control subspace approximations. In particular, for any positive integer l, satisfying $l \geq \max(m, n)$, consider the following optimization problem involving the l-by-l leading principal minor of $\mathbf{F}$, $\mathbf{F}_l$:

Optimization Problem $(\mathbf{H}, \mathbf{F}_l)$: Let $\mathbf{U}$ and $\mathbf{V}$ be m-by-l and n-by-l matrices respectively. The optimization problem is defined by:

$$\min_{\mathbf{U},\mathbf{V}} \left(\frac{1}{m} tr\mathbf{U}\mathbf{U}^T + \frac{1}{n} tr\mathbf{V}\mathbf{V}^T \right) \tag{4.2}$$

subject to the constraint:

$$\mathbf{H} = \mathbf{U}\mathbf{F}_l\mathbf{V}^T. \tag{4.3}$$

While there exists bilinear $\mathbf{F}$ for which the minimum control cost does not exist, the infimum, denoted by $\hat{C}_F(\mathbf{H})$, is always well-defined if the target function is realizable.

One can formulate a slightly more general version of this optimization problem by allowing the weights in the cost function to be arbitrary positive integers:

Generalized Optimization Problem $(\mathbf{H}, \mathbf{F}_l; p, q)$: Let $\mathbf{H}$ be an m-by-n target matrix, $\mathbf{F}_l$ be an l-by-l leading principal minor of $\mathbf{F}$, with $l \geq max(m, n)$, p and q be arbitrary positive integers. The optimization problem is defined by:

$$\min_{\mathbf{U},\mathbf{V}} \left(\frac{1}{p} tr \mathbf{U}\mathbf{U}^T + \frac{1}{q} tr \mathbf{V}\mathbf{V}^T \right) \tag{4.4}$$

subject to the condition

$$\mathbf{H} = \mathbf{U}\mathbf{F}_l\mathbf{V}^T. \tag{4.5}$$

There is an important connection between these two classes of problems. Given an m-by-n target matrix, $\mathbf{H}$, one can obtain another target matrix by appending rows of zeros and columns of zeros to obtain an l-by-l matrix, with $l \geq n, l \geq m$, so that

$$\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{H} & \mathbf{0}_{m,l-n} \\ \mathbf{0}_{l-m,n} & \mathbf{0}_{l-m,l-n} \end{bmatrix}, \tag{4.6}$$

where $\mathbf{0}_{i,j}$ is an i-by-j matrix with all zeros.

Now consider the Generalized Optimization Problem $(\tilde{\mathbf{H}}, \mathbf{F}_l; m, n)$. Any optimal solution to the Generalized Optimization Problem must satisfy the property:

$$u_i = 0, \ v_j = 0 \tag{4.7}$$

for $i > m$ and $j > n$. Otherwise, a lower cost solution can be obtained by substituting with these zero controls. From this, one can conclude that the solution is also optimal for the lower dimensional Optimization Problem, $(\mathbf{H}, \mathbf{F}_l)$. Conversely, an optimal solution to the latter problem can be extended to an optimal solution to the Generalized Optimization Problem $(\tilde{\mathbf{H}}, \mathbf{F}_l; m, n)$. Hence, by using the Generalized Optimization Problem formulation, we can assume without loss of generality that the target matrices are square matrices with the same dimensions as $\mathbf{F}_l$.

5. SINGLE ROUND PROTOCOLS: MINIMUM ENERGY CONTROL

One of the key results in this paper is summarized by the following theorem. Although this result is based on regularity or strong regularity, for any m-by-n target matrix, $\mathbf{H}$, the theorem also holds for any finite bilinear input-output mapping with a matrix representation, $\mathbf{F}$, that is nonsingular and l-by-l in dimensional with $l \geq max(m, n)$. It is easy to construct examples satisfyig this condition from equation (3.4) if the control functions are restricted to l-dimensional subspaces of $\mathcal{L}$.

Theorem 5.1. *Consider a bounded bilinear input-output mapping, F, with a regular matrix representation $\mathbf{F}$. Let $\mathbf{H}$ be an m-by-n target matrix. The infimum control cost of any single round protocol that realizes $\mathbf{H}$ is given by:*

$$\hat{C}_F(\mathbf{H}) = \frac{2}{\sqrt{mn}} \sum_{k=1}^{\min(m,n)} \sigma_k(\mathbf{H})/\sigma_k(\mathbf{F}). \tag{5.1}$$

If the matrix representation is strongly regular, there exists a single round protocol that achieves this infimum control cost.

Before proving Theorem 5.1, we present a corollary that localizes the result to the B-H system (3.3). First note that the diagonal entries of $\mathbf{F_B}$ in (3.11) can be represented as $(\lceil k/2 \rceil \pi)^{-1}$, and this leads to the following result.

Corollary 5.1. *Consider the input-output system (3.3) and an m-by-n target matrix $\mathbf{H}$. For single round protocols that realize $\mathbf{H}$ the infimum control cost is given by*

$$\hat{C}_{F_B}(\mathbf{H}) = \frac{2\pi}{\sqrt{mn}} \sum_{k=1}^{\min(m,n)} \lceil k/2 \rceil \sigma_k(\mathbf{H}), \tag{5.2}$$

and this is achieved by Alice choosing m controls and Bob choosing n controls from the space spanned by the basis $\mathcal{B}_0$ defined in (3.29).

To prove the main theorem, we first establish a proposition in which we prove the result for a square target matrix $\mathbf{H}$ and for controls restricted to finite dimensional subspaces.

Proposition 5.1. *Consider an l-by-l target matrix $\mathbf{H}$ with rank r, and an invertible matrix $\mathbf{F}_l$. The minimum control cost for the Generalized Optimization Problem $(\mathbf{H}, \mathbf{F}_l; p, q)$ is achievable and is given by the formula,*

$$\hat{C}_{\mathbf{F}_l}(\mathbf{H}) = \frac{2}{\sqrt{pq}} \sum_{k=1}^{r} \sigma_k(\mathbf{H})/\sigma_k(\mathbf{F}_l). \tag{5.3}$$

Proof: To prove the proposition, first of all we show that the right-hand-side of (5.3) can be achieved. Let Π be the orthogonal matrix that puts $\mathbf{F}_l\mathbf{F}_l^T$ into the following diagonal form:

$$\Pi^T\mathbf{F}_l\mathbf{F}_l^T\Pi = \begin{bmatrix} \sigma_1^2(\mathbf{F_l}) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_l^2(\mathbf{F}_l) \end{bmatrix}. \tag{5.4}$$

Define $\tilde{\mathbf{U}} = \Pi^T\mathbf{U}\Pi$ and $\tilde{\mathbf{V}} = \Pi^T\mathbf{V}\Pi$. Then $(\mathbf{U},\mathbf{V})$ is a solution to

$$\mathbf{H} = \mathbf{U}\mathbf{F}_l\mathbf{V}^T \tag{5.5}$$

if and only if $(\tilde{\mathbf{U}},\tilde{\mathbf{V}})$ is a solution to the equation

$$\Pi^T\mathbf{H}\Pi = \tilde{\mathbf{U}}\Pi^T\mathbf{F}_l\Pi\tilde{\mathbf{V}}^T. \tag{5.6}$$

Since the cost of $(\mathbf{U},\mathbf{V})$ and $(\tilde{\mathbf{U}},\tilde{\mathbf{V}})$ are identical, we can assume without loss of generality that $\mathbf{F}_l\mathbf{F}_l^T$ is in the diagonal form (5.4).

Let $\boldsymbol{\Theta}$ be the orthogonal matrix that diagonalizes $\mathbf{H}^T\mathbf{H}$ so that

$$\boldsymbol{\Theta}\mathbf{H}^T\mathbf{H}\boldsymbol{\Theta}^T = \begin{bmatrix} \sigma_1^2(\mathbf{H}) & & \dots & & 0 \\ & \ddots & & & \\ \vdots & & \sigma_r^2(\mathbf{H}) & & \vdots \\ & & & 0 & \\ & & & & \ddots \\ 0 & & \dots & & 0 \end{bmatrix}, \tag{5.7}$$

and let

$$\mathbf{R}_\delta = \left(\frac{q}{p}\right)^{1/4} \begin{bmatrix} \sqrt{\sigma_1(\mathbf{H})\sigma_1(\mathbf{F}_l)} & & \dots & & 0 \\ & \ddots & & & \\ \vdots & & \sqrt{\sigma_r(\mathbf{H})\sigma_r(\mathbf{F}_l)} & & \vdots \\ & & & \delta & \\ & & & & \ddots \\ 0 & & \dots & & \delta \end{bmatrix} \tag{5.8}$$

for a small $\delta > 0$. In equation (5.7) the lower right-hand-side zero-diagonal block is absent if $r = l$. Similarly for equation (5.8), the lower right-hand-side δ-diagonal block is absent if $r = l$. Define

$$\mathbf{U}_\delta = \mathbf{H}\boldsymbol{\Theta}\mathbf{R}_\delta{}^{-1},\ \mathbf{V}_\delta^T = \mathbf{F}_l^{-1}\mathbf{R}_\delta\boldsymbol{\Theta}^T, \tag{5.9}$$

we obtain a solution to (5.5). By direct computation, it follows that

$$\frac{1}{p}tr\mathbf{U}_\delta\mathbf{U}_\delta^T = \frac{1}{p}tr\mathbf{H}\boldsymbol{\Theta}\mathbf{R}_\delta^{-2}\boldsymbol{\Theta}^T\mathbf{H}^T = \frac{1}{p}tr\mathbf{R}_\delta^{-2}\boldsymbol{\Theta}^T\mathbf{H}^T\mathbf{H}\boldsymbol{\Theta} = \frac{1}{\sqrt{pq}}\sum_{k=1}^{r}\sigma_k(\mathbf{H})/\sigma_k(\mathbf{F}_l). \tag{5.10}$$

Since entries of $\mathbf{U}_\delta$ are affine functions of $1/\delta$, equation (5.10) implies $\mathbf{U}_\delta$ is independent of δ. We express this by writing $\mathbf{U}_\delta \equiv \mathbf{U}_0$. It is clear that $\lim_{\delta\to 0} \mathbf{V}_\delta$ exists. Denote it by $\mathbf{V}_0$. Then,

$$\begin{aligned} \tfrac{1}{q} tr \mathbf{V}_\delta \mathbf{V}_\delta^T &= \tfrac{1}{q} tr \mathbf{\Theta} \mathbf{R}_\delta (\mathbf{F}_l \mathbf{F}_l^T)^{-1} \mathbf{R}_\delta \mathbf{\Theta}^T = \tfrac{1}{q} tr \mathbf{R}_\delta (\mathbf{F}_l \mathbf{F}_l^T)^{-1} \mathbf{R}_\delta \\ &= \tfrac{1}{q} tr \mathbf{R}_\delta^2 (\mathbf{F}_l \mathbf{F}_l^T)^{-1} = \tfrac{1}{\sqrt{pq}} \textstyle\sum_{k=1}^r \sigma_k(\mathbf{H})/\sigma_k(\mathbf{F}_l) + \tfrac{\delta^2}{\sqrt{pq}} \textstyle\sum_{k=r+1}^l 1/\sigma_k(\mathbf{F}_l). \end{aligned} \tag{5.11}$$

Taking the limit as $\delta \to 0$,

$$\frac{1}{p} tr \mathbf{U}_0 \mathbf{U}_0^T + \frac{1}{q} tr \mathbf{V}_0 \mathbf{V}_0^T = \frac{2}{\sqrt{pq}} \sum_{k=1}^r \sigma_k(\mathbf{H})/\sigma_k(\mathbf{F}_l), \tag{5.12}$$

proving that the right-hand-side of (5.3) can be achieved.

To complete the proof, we want to show the right-hand-side of (5.3) is a lower bound for all single round protocols realizing $\mathbf{H}$. We show this first for invertible $\mathbf{H}$. In this case, all solutions $(\mathbf{U}, \mathbf{V})$ to the equation (5.5) consist of invertible matrices and for every invertible $\mathbf{V}$ there is a unique matrix $\mathbf{U}$ that solves (5.5). By applying polar decomposition to $\mathbf{F}_l \mathbf{V}^T$ we obtain

$$\mathbf{V}^T = \mathbf{F}_l^{-1} \mathbf{R} \mathbf{\Theta}^T, \tag{5.13}$$

for an orthogonal matrix, $\mathbf{\Theta}$, and a non-singular symmetric matrix, $\mathbf{R}$. These two matrices can be regarded as free variables and we can parametrize the solution space to (5.3) by $\mathbf{R}$ and $\mathbf{\Theta}$. Specifically, if we let $\mathbf{U} = \mathbf{H} \mathbf{\Theta} \mathbf{R}^{-1}$, then the control cost is

$$\frac{1}{p} tr \mathbf{U} \mathbf{U}^T + \frac{1}{q} tr \mathbf{V} \mathbf{V}^T = \frac{1}{p} tr \mathbf{R}^{-2} \mathbf{\Theta}^T \mathbf{H}^T \mathbf{H} \mathbf{\Theta} + \frac{1}{q} tr \mathbf{R}^2 (\mathbf{F}_l \mathbf{F}_l^T)^{-1}. \tag{5.14}$$

For any two l-by-l positive semidefinite symmetric matrices, $\mathbf{P}$ and $\mathbf{Q}$

$$tr \mathbf{P} \mathbf{Q} = \sum_{k=1}^l \lambda_k(\mathbf{P} \mathbf{Q}) \geq \sum_{k=1}^l \lambda_k(\mathbf{P}) \lambda_{l-k+1}(\mathbf{Q}). \tag{5.15}$$

(See, for instance, p.249 of [26].) Therefore,

$$tr \mathbf{R}^2 (\mathbf{F}_l \mathbf{F}_l^T)^{-1} \geq \sum_{k=1}^l \lambda_k^2(\mathbf{R}) \sigma_{l-k+1}^2(\mathbf{F}_l^{-1}) = \sum_{k=1}^l \lambda_k^2(\mathbf{R})/\sigma_k^2(\mathbf{F}_l). \tag{5.16}$$

Similarly,

$$tr \mathbf{R}^{-2} \mathbf{\Theta}^T \mathbf{H}^T \mathbf{H} \mathbf{\Theta} \geq \textstyle\sum_{k=1}^l \sigma_k^2(\mathbf{H})/\lambda_k^2(\mathbf{R}). \tag{5.17}$$

Hence,

$$\hat{C}_{\mathbf{F}_l}(\mathbf{H}) \geq \frac{1}{p} \sum_{k=1}^l \sigma_k^2(\mathbf{H})/\lambda_k^2(\mathbf{R}) + \frac{1}{q} \sum_{k=1}^l \lambda_k^2(\mathbf{R})/\sigma_k^2(\mathbf{F}_l). \tag{5.18}$$

One can minimize the right-hand-side of (5.18) by considering

$$\min \sum_{k=1}^{l} \left[\frac{1}{p} \frac{\sigma_k^2(\mathbf{H})}{t_k} + \frac{1}{q} \frac{t_k}{\sigma_k^2(\mathbf{F}_l)} \right], \tag{5.19}$$

subject to the constraint

$$t_1 \geq t_2 \ldots \geq t_l > 0. \tag{5.20}$$

Via calculus, the unconstrained optimal solution to (5.18) is shown to be

$$t_k = \sigma_k(\mathbf{H})\sigma_k(\mathbf{F}_l)\sqrt{\frac{q}{p}}, \tag{5.21}$$

which also satisfies the condition in (5.20). It follows that

$$\hat{C}_{\mathbf{F}_l}(\mathbf{H}) \geq \frac{2}{\sqrt{pq}} \sum_{k=1}^{l} \sigma_k(\mathbf{H})/\sigma_k(\mathbf{F}_l). \tag{5.22}$$

It follows that the proposition holds for invertible $\mathbf{H}$.

We want to show the inequality (5.22) also holds when $\mathbf{H}$ is not full rank. Suppose $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ give an optimal solution to $(\mathbf{H}, \mathrm{F}_l)$. Define

$$\mathbf{U}(\epsilon) = \hat{\mathbf{U}} + \epsilon\mathbf{I},\ \mathbf{V}(\epsilon) = \hat{\mathbf{V}} + \epsilon\mathbf{I}, \tag{5.23}$$

where $\mathbf{I}$ is the l-by-l identity matrix. Other than at most a finite set of values, all these matrices are invertible. Without loss of generality, assume that for some $a > 0$ and all $0 < \epsilon \leq a$, the matrices $\mathbf{U}(\epsilon)$ and $\mathbf{V}(\epsilon)$ are invertible and restrict ϵ in subsequent discussion to such an interval. Let $\mathcal{U}(\epsilon)$ denote the set of controls corresponding to $\mathbf{U}(\epsilon)$ and similarly, $\mathcal{V}(\epsilon)$ denote the sets of controls corresponding to $\mathbf{V}(\epsilon)$. Define

$$\mathbf{G}(\epsilon) = \mathbf{U}(\epsilon)\mathbf{F}_l\mathbf{V}(\epsilon)^T = \mathbf{H} + \epsilon(\hat{\mathbf{U}}\mathbf{F}_l + \mathbf{F}_l\hat{\mathbf{V}}^T) + \epsilon^2\mathbf{F}_l. \tag{5.24}$$

For $0 < \epsilon \leq a$, the matrices $\mathbf{G}(\epsilon)$ are invertible. Hence,

$$\hat{C}_{\mathbf{F}_l}(\mathbf{H}) = \frac{1}{p} tr\hat{\mathbf{U}}\hat{\mathbf{U}}^T + \frac{1}{q} tr\hat{\mathbf{V}}\hat{\mathbf{V}}^T = \lim_{\epsilon\to 0}\left(\frac{1}{p}tr\mathbf{U}(\epsilon)\mathbf{U}(\epsilon)^T + \frac{1}{q}tr\mathbf{V}(\epsilon)\mathbf{V}(\epsilon)^T\right) \tag{5.25}$$

$$\geq \lim_{\epsilon\to 0}\hat{C}_{\mathbf{F}_l}(\mathbf{G}(\epsilon)) = \lim_{\epsilon\to 0}\frac{2}{\sqrt{pq}}\sum_{k=1}^{l}\sigma_k(\mathbf{G}(\epsilon))/\sigma_k(\mathbf{F}_l) \tag{5.26}$$

$$= \frac{2}{\sqrt{pq}}\sum_{k=1}^{l}\sigma_k(\mathbf{H})/\sigma_k(\mathbf{F}_l) = \frac{2}{\sqrt{pq}}\sum_{k=1}^{r}\sigma_k(\mathbf{H})/\sigma_k(\mathbf{F}_l). \tag{5.27}$$

Note that the equality in (5.26) follows from the fact that the proposition holds for invertible target matrices, while the first equality in (5.27) follows from the continuity of the singular values as a function of the matrix coefficients. This proves the proposition. ■

Proof of Theorem 5.1: Let $\{e_1, e_2, \ldots\}$ and $\{f_1, f_2, \ldots\}$ be the bases corresponding to the representation $\mathbf{F}$. Let $\hat{\mathcal{U}} = \{\hat{u}_1, \ldots, \hat{u}_m\}$ and $\hat{\mathcal{V}} = \{\hat{v}_1, \ldots, \hat{v}_n\}$ be the controls in an optimal solution for the Optimization Problem $(\mathbf{H}, \mathbf{F})$, with corresponding matrices $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ respectively. Let

$$\hat{u}_i = \sum_{j=1}^{\infty} a_{ij} e_j, \ \hat{v}_i = \sum_{i=1}^{\infty} b_{ij} f_j. \tag{5.28}$$

Then

$$I(\hat{\mathcal{U}}, \hat{\mathcal{V}}) = \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{\infty} a_{ij}^2 + \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{\infty} b_{ij}^2, \tag{5.29}$$

and

$$\mathbf{H} = \hat{\mathbf{U}} \mathbf{F} \hat{\mathbf{V}}^T. \tag{5.30}$$

Define approximating control functions

$$\hat{u}_i^{(l)} = \sum_{j=1}^{l} a_{ij} e_j, \ \hat{v}_i^{(l)} = \sum_{i=1}^{l} b_{ij} f_j \tag{5.31}$$

with corresponding m-by-l and l-by-n matrices

$$\hat{\mathbf{U}}_l = \begin{bmatrix} a_{1,1} & \ldots & a_{1,l} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \ldots & a_{m,l} \end{bmatrix}, \ \hat{\mathbf{V}}_l = \begin{bmatrix} b_{1,1} & \ldots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{l,1} & \ldots & b_{l,n} \end{bmatrix}. \tag{5.32}$$

Then, equation (5.29) can be rewritten as

$$I(\hat{\mathcal{U}}, \hat{\mathcal{V}}) = \frac{1}{m} tr \hat{\mathbf{U}} \hat{\mathbf{U}}^T + \frac{1}{n} tr \hat{\mathbf{V}} \hat{\mathbf{V}}^T = \lim_{l \to \infty} \left(\frac{1}{m} tr \hat{\mathbf{U}}_l \hat{\mathbf{U}}_l^T + \frac{1}{n} tr \hat{\mathbf{V}}_l \hat{\mathbf{V}}_l^T \right). \tag{5.33}$$

Define the m-by-n approximate target matrix by

$$\mathbf{H}_l = \left(H_{ij}^{(l)} \right) = \hat{\mathbf{U}}_l \mathbf{F}_l \hat{\mathbf{V}}_l^T, \tag{5.34}$$

where $\mathbf{F}_l$ is the l-th principal minor of $\mathbf{F}$. Since F is bounded,

$$H_{ij} = F(\hat{u}_i, \hat{v}_j) = \lim_{l \to \infty} F(\hat{u}_i^{(l)}, \hat{v}_j^{(l)}) = \lim_{l \to \infty} H_{ij}^{(l)}. \tag{5.35}$$

In matrix form:

$$\lim_{l \to \infty} \hat{\mathbf{U}}_l \mathbf{F}_l \hat{\mathbf{V}}_l^T = \lim_{l \to \infty} \mathbf{H}_l = \mathbf{H}. \tag{5.36}$$

Since the singular value function is continuous on the space of m-by-n matrices, for any integer $k, 0 \leq k \leq m$, we have

$$\lim_{l\to\infty} \sigma_k(\mathbf{H}_l) = \sigma_k(\mathbf{H}). \tag{5.37}$$

For any integer $l \geqslant \max(m,n)$, define the l-by-l augmented matrix $\tilde{\mathbf{H}}_l$ by

$$\tilde{\mathbf{H}}_l = \begin{bmatrix} \mathbf{H}_l & \mathbf{0}_{m,l-n} \\ \mathbf{0}_{l-m,n} & \mathbf{0}_{l-m,l-n} \end{bmatrix}. \tag{5.38}$$

Here, $\mathbf{0}_{i,j}$ represents an i-by-j zero matrix, (empty when one of the dimensions is zero.) As argued before, the minimum control costs for the Optimization Problem $(\mathbf{H}_l, \mathbf{F}_l)$ and the Generalized Optimization Problem $(\tilde{\mathbf{H}}_l, \mathbf{F}_l, m, n)$ are identical. By Proposition 5.1,

$$\frac{1}{m} tr \hat{\mathbf{U}}_l \hat{\mathbf{U}}_l^T + \frac{1}{n} tr \hat{\mathbf{V}}_l \hat{\mathbf{V}}_l^T \geq \hat{C}_{\mathbf{F}_l}(\tilde{\mathbf{H}}_l) = \frac{2}{\sqrt{mn}} \sum_{k=1}^{l} \sigma_k(\tilde{\mathbf{H}}_l)/\sigma_k(\mathbf{F}_l). \tag{5.39}$$

Moreover, he last expression in (5.39) can be realized by some control functions.

The rank of $\tilde{\mathbf{H}}_l$ is equal to the rank of $\mathbf{H}_l$, which is at most $\min(m,n)$, moreover, the first $\min(m,n)$ singular values of the two matrices are identical. It follows that for $k > \min(m,n)$

$$\sigma_k(\tilde{\mathbf{H}}_l) = 0, \tag{5.40}$$

and the last expression in (5.39) is equal to

$$\frac{2}{\sqrt{mn}} \sum_{k=1}^{\min(m,n)} \sigma_k(\mathbf{H}_l)/\sigma_k(\mathbf{F}_l) \tag{5.41}$$

and can be realized. The first part of the theorem then follows from:

$$\frac{1}{m} tr \hat{\mathbf{U}} \hat{\mathbf{U}}^T + \frac{1}{n} tr \hat{\mathbf{V}} \hat{\mathbf{V}}^T \geq \lim_{l\to\infty} \frac{2}{\sqrt{mn}} \sum_{k=1}^{\min(m,n)} \sigma_k(\mathbf{H}_l)/\sigma_k(\mathbf{F}_l) = \frac{2}{\sqrt{mn}} \sum_{k=1}^{\min(m,n)} \sigma_k(\mathbf{H})/\sigma_k(\mathbf{F}). \tag{5.42}$$

For the strongly regular case, by definition for every k, there exists a finite l_k such that $\sigma_k(\mathbf{F}_{l_k}) = \sigma_k(\mathbf{F})$. Let

$$\bar{l} = \max_{k=1,\ldots,\min(m,n)} l_k. \tag{5.43}$$

For $k = 1, \ldots, \min(m,n)$,

$$\sigma_k(\mathbf{F}_{\bar{l}}) = \sigma_k(\mathbf{F}). \tag{5.44}$$

Thus the solution to the Optimization Problem $(\mathbf{H}_l, \mathbf{F}_{\bar{l}})$ achieves the control cost given by the lower bound in (5.42). This completes the proof of Theorem 5.1. ■

We conclude the discussion in this section by pointing out that there is an apparent non-symmetry in the solution constructed in the proof of Proposition 5.1. Even if $\mathbf{H}$ and $\mathbf{F}$ are symmetric, the optimal solution may not be symmetric in the sense that $\mathbf{U}$ and $\mathbf{V}$ may not be identical. This would appear less surprising if one considers the example where $\mathbf{H}$ is

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{5.45}$$

and $\mathbf{F}$ is the 2-by-2 identity matrix. Having a symmetric solution to the optimization problem would imply $\mathbf{H}$ is non-negative semidefinite which is obviously not the case.

6. COST OF DISTRIBUTED ACTION

Results in the preceding section provide an explicit formula for computing the control cost in the absence of any communication between the agents. However, if information can be shared between the agents, controls with lower average cost of distributed action can be designed. For the B-H system, if both agents have information on the other agent's choice, they can use the following control functions to realize the target, H_{ij}:

$$u_{ij}(t) = sgn(H_{ij})\sqrt{2\pi H_{ij}}\cos(2\pi t), \tag{6.1}$$

$$v_{ij}(t) = \sqrt{2\pi H_{ij}}\sin(2\pi t). \tag{6.2}$$

The control energy of such a protocol is $2\pi|H_{ij}|$. Using the isoperimetric inequality [30], one can show as in [35] that this is the minimum control cost to realize the single target, H_{ij}.

Hence, if both agents have complete information of the other agent's choice, the averaged control cost over all possible choices (m for Alice and n for Bob) is

$$J(\mathbf{H}) = \frac{2\pi}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}|H_{ij}|, \tag{6.3}$$

for the B-H system. One can compare this control cost with the control cost for a single round protocol defined in (2.9) and note that while the number of control input pairs available to Alice and Bob is the same in both cases, the number of distinct controls used by each is typically not.

In order to enable a concrete comparison between single round protocols and protocols based on completely shared information we will use the B-H system as an example. First of all, we consider the case that $\mathbf{H}$ is a Hadamard matrix of order n. Since the entries of such matrices are either 1 or -1, the averaged control cost with perfect information $J(\mathbf{H})$ is simply 2π. On the other hand, the singular values of $\mathbf{H}$ are all equal to $\sqrt{n}$. Thus, by Corollary 5.1 the optimal single round protocol cost $\hat{C}_F(\mathbf{H})$ is

$$\begin{cases} \frac{\pi\sqrt{n}}{2}(n+2) & \text{for even } n, \\ \frac{\pi\sqrt{n}}{2}(n+2+\frac{1}{n}) & \text{for odd } n. \end{cases} \tag{6.4}$$

Thus, the ratio of two control costs is asymptotically $(n+2)\sqrt{n}/4$.

We can perform a similar comparison for the case of orthogonal matrices. To do so, we need to quote the following bounds on the sum of the absolute value of the entries in an orthogonal matrix. These elementary results on matrices are provided here for completeness sake.

Proposition 6.1. *Consider B-H system. For an orthogonal* $\mathbf{H}$*,*

$$\frac{2\pi}{n} \le \frac{2\pi}{n^2}\sum_{i,j=1}^{n}|H_{ij}| \le \frac{2\pi}{\sqrt{n}}. \tag{6.5}$$

For a general $\mathbf{H}$*,*

$$\frac{2\pi}{n^2}\sum_{i,j=1}^{n}|H_{ij}| \le \frac{2\pi}{n}\sum_{k=1}^{n}\sigma_k(\mathbf{H}). \tag{6.6}$$

Proof: To prove the lower bound when $\mathbf{H}$ is an orthogonal matrix, use the inequality

$$\sum_{i,j=1}^{n}|H_{i,j}| \ge \sum_{i,j=1}^{n}H_{i,j}^2 = n. \tag{6.7}$$

To prove the upper bound, consider the optimization problem,

$$S = \max\sum_{i,j=1}^{n}x_{i,j} \tag{6.8}$$

subject to the constraint that for all j, $\sum_{i=1}^{n}x_{i,j}^2 = 1$. Since the constraint is weaker than the requirement that the $x_{i,j}$'s form an orthogonal matrix, it follows that

$$\sum_{i,j=1}^{n}|H_{i,j}| \le S = n\max_{z_1^2+\cdots+z_n^2=1}\sum_{i=1}^{n}z_i = n\sqrt{n}. \tag{6.9}$$

For the case where $\mathbf{H}$ is a general matrix, decompose it by SVD so that

$$\mathbf{H} = \mathbf{\Phi\Lambda\Theta}, \quad \mathbf{\Lambda} = \begin{bmatrix} \sigma_1(\mathbf{H}) & 0 & \cdots & 0 \\ 0 & \sigma_2(\mathbf{H}) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n(\mathbf{H}) \end{bmatrix}. \tag{6.10}$$

where $\mathbf{\Phi}$ and $\mathbf{\Theta}$ are orthogonal. For $k, 1 \le k \le n$, define an n-by-n matrix, $\mathbf{H}(k)$, by

$$\mathbf{H}(k) = \mathbf{\Phi}\mathbf{\Lambda}_k\mathbf{\Theta},\ \mathbf{\Lambda}_k = diag[\underbrace{0,\ldots,0}_{k-1},\sigma_k(\mathbf{H}),0,\ldots,0], \tag{6.11}$$

Then,

$$\begin{aligned} &\textstyle\sum_{i,j=1}^n |H(k)_{ij}| = \sigma_k(\mathbf{H})\sum_{i,j=1}^n |\Phi_{ik}\Theta_{kj}| \\ &\le \sigma_k(\mathbf{H})\textstyle\sum_{i=1}^n |\Phi_{ik}|\sum_{i=1}^n |\Theta_{kj}| \le n\sigma_k(\mathbf{H}). \end{aligned} \tag{6.12}$$

From this, it follows that

$$\frac{2\pi}{n^2}\sum_{i,j=1}^n |H_{ij}| \le \frac{2\pi}{n^2}\sum_{k=1}^n\sum_{i,j=1}^n |H(k)_{ij}| \le \frac{2\pi}{n}\sum_{k=1}^n \sigma_k(\mathbf{H}). \tag{6.13}$$

■

This result together with Corollary 5.1 imply the following

Theorem 6.1. *For the B-H system and a general n-by-n target matrix $\mathbf{H}$,*

$$\hat{C}_F(\mathbf{H}) - J(\mathbf{H}) \ge \sum_{k=1}^n (\lceil k/2 \rceil - 1)\sigma_k(\mathbf{H}) \ge 0. \tag{6.14}$$

The difference in the control energy is strictly positive if the rank of $\mathbf{H}$ is larger than 2.

If $\mathbf{H}$ is an orthogonal matrix, then the optimal single round cost $\hat{C}_F(\mathbf{H})$ is

$$\begin{cases} \frac{\pi}{2}(n+2) & \text{for even } n, \\ \frac{\pi}{2}(n+2+\frac{1}{n}) & \text{for odd } n. \end{cases} \tag{6.15}$$

By comparison,

$$\frac{2\pi}{n} \le J(\mathbf{H}) \le \frac{2\pi}{\sqrt{n}}. \tag{6.16}$$

For example, consider the case where $n = 2$. The identity matrix, $\mathbf{I}_2$, incurs an averaged cost of π under information sharing: there are two cases where control energy of 2π is needed for each case; for the other two cases, H_{12} and H_{21}, zero control can be used. However, for the case where the information on the choice is not shared, it is not possible to save control cost by setting any of the controls to zero, resulting in a cost increase by a factor of 2. In general, the control cost ratio grows super-linearly as a function of the dimension of target matrix.

7. MULTI-ROUND PROTOCOLS

Analysis from the previous section indicates that control cost can be substantially reduced by using protocols that allow communication between the agents. For the extreme scenario where the agents have complete prior information of each others inputs, optimization of the control cost can be reduced to solving a family of single output target optimization problems. For any m-by-n target matrix $\mathbf{H}$ and a general bilinear input-output mapping, $\mathbf{F}$, the averaged control cost for solving a family of single output target optimization problems is given by:

$$J(\mathbf{H}) = \frac{2}{mn\sigma_1(\mathbf{F})} \sum_{i=1}^{m} \sum_{j=1}^{n} |H_{i,j}|. \tag{7.1}$$

This is a generalization of equation (6.3).

If the agents do not have prior information on the inputs to be selected, they can communicate their input choice to each other in a multi-round protocol. Detailed analysis of general multi-round protocols lies beyond the scope of this paper. We shall briefly consider two-phase protocols, however, in which one phase allows partial information to be shared at negligible cost. The main result here is that if the cost of signaling certain partial information is negligible, the control cost can be made arbitrarily close to the lower bound, $J(\mathbf{H})$ that was determined in the previous section. It is importzant to note that the protocols approaching $J(\mathbf{H})$ do not necessarily require the agents to communicate their choices completely to each other. It will be shown that the number of bits communicated is related to the classical communication complexity of computing $\mathbf{H}$. This result provides insight into the *value of information* in terms of the control energy savings that can be achieved through communication between the agents as they evaluate $\mathbf{H}$.

Various concepts of communication *at negligible cost* can be considered. One possible approach to partial information exchange in a two-phase protocol is to use a *side channel* in an initiation phase in which the cost of transmitting some partial information between the agents in the first phase. It is in the second phase that this information is used by the agents to select controls that effect the computation specified in the single-round protocols of Section 5. Another approach to information sharing at negligible communication cost is to assume that under certain circumstances, very low cost control signals can be used. Formally, for the model (2.3) described in section 2, we introduce the following concept of *ϵ-signaling capability*.

Definition 7.1. *Alice possesses ϵ-signaling capability around the initial state $\mathbf{x}_0$ if for any ϵ*

there exist times, $0 < t_1 < t_2$ *and controls* u_1 *and* u_2 *for Alice,* v_1 *for Bob, so that*

$$\int_0^{t_2} (u_1^2 + v_1^2)dt < \epsilon, \ \int_0^{t_2} (u_2^2 + v_1^2)dt < \epsilon. \tag{7.2}$$

Moreover, when u *is set to* u_1 *or* u_2 *and* v *is set to* v_1*, the following conditions hold:*
1) $\mathbf{x}_{u_1,v_1}(t_1) \neq \mathbf{x}_{u_2,v_1}(t_1)$, 2) $\mathbf{x}_{u_1,v_1}(t_2) = \mathbf{x}_{u_2,v_1}(t_2) = \mathbf{x}_0$.

One can define similar capability for Bob. For example, for the B-H system both agents possess ϵ-signaling capability as the loop controls can enclose arbitrarily small areas. For systems in which both agents have ϵ-signaling capability one can design multi-round protocols for realizing $\mathbf{H}$ with control cost arbitrarily close to $J(\mathbf{H})$. These protocols consist of two phases. In the first phase, the agents communicate their choices of inputs to each other. Based on the information received, the original target matrix is decomposed into a finite number of submatrices and controls can then be applied to realize the submatrix that corresponds to the choices of the agents. We call such a protocol a *two-phase protocol.* To describe the detail, we recall a definition from communication complexity theory, (see for example [19]).

Definition 7.2. *A submatrix is said to be monochromatic if all its entries have the same value.*

Given an m-by-n target matrix, $\mathbf{H}$, we can define a set of submatrices, $\{\mathbf{H}_1, \ldots, \mathbf{H}_K\}$, so that $\mathbf{H}_k$ is an m_k-by-n_k submatrix with its (i, j) entry defined by

$$H_k(i, j) = H_{t_{k,i}, s_{k,j}} \tag{7.3}$$

where $t_{k,i}$ lies in an index set $\mathcal{M}_k \subset \{1, \ldots, m\}$ and $s_{k,j}$ lies in an index set $\mathcal{N}_k \subset \{1, \ldots, n\}$. Note that by definition, $|\mathcal{M}_k| = m_k$, $|\mathcal{N}_k| = n_k$. Define $l_k = min(m_k, n_k)$.

The set of submatrices forms a matrix partition for $\mathbf{H}$ if the following holds:

1) For any i and j with $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$, there exist k, α, and β such that $t_{k,\alpha} = i, \ s_{k,\beta} = j$.
2) If $t_{k,\alpha} = t_{k',\alpha'}, \ s_{k,\beta} = s_{k',\beta'}$, then $k = k'$, $\alpha = \alpha'$, and $\beta = \beta'$.

For example, the following figure shows a submatrix partition involving five submatrices:

$$\left[\begin{array}{cccc} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{31} & H_{33} & H_{34} \\ H_{41} & H_{41} & H_{43} & H_{44} \end{array}\right] \tag{7.4}$$

The matrices in this partition are sub-blocks, but in general matrix partitions need not be entirely comprised of a set of sub-blocks.

It follows from direct verification that for all k

$$\sum_{l=1}^{l_k} \sigma_l(\mathbf{H}_k)^2 = tr\mathbf{H}_k\mathbf{H}_k^T = \sum_{i=1}^{m_k}\sum_{j=1}^{n_k} H_k^2(i,j) = \sum_{t_{k,i}\in\mathcal{M}_k}\sum_{s_{k,j}\in\mathcal{N}_k} H^2_{t_{k,i},s_{k,j}}. \tag{7.5}$$

Thus, for a submatrix partition into K submatrices,

$$\sum_{k=1}^{K}\sum_{l=1}^{l_k} \sigma_l(\mathbf{H}_k)^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} H_{i,j}^2 = tr\mathbf{H}\mathbf{H}^T. \tag{7.6}$$

A sub-matrix partition can be regarded as a decomposition of a complex distributed control problem into simpler sub-problems. We can estimate the effectiveness of a decomposition by calculating the control cost averaged over the decomposed sub-problems, namely,

$$A = \frac{2}{mn}\sum_{k=1}^{K} m_k n_k \hat{C}_{\mathbf{F}}(\mathbf{H}_k). \tag{7.7}$$

The following result provides a lower bound for this averaged control cost.

Theorem 7.1. *Consider a bounded, bilinear input-output mapping, F, with a regular matrix representation $\mathbf{F}$. The average control cost, A, for applying single round protocols to the submatrices in a submatrix partition $\{\mathbf{H}_1,\ldots,\mathbf{H}_K\}$ of $\mathbf{H}$ satisfies the lower bound:*

$$A = \frac{1}{mn}\sum_{k=1}^{K} m_k n_k \hat{C}_{\mathbf{F}}(H_k) \geq \frac{2}{mn\sigma_1(\mathbf{F})}\sum_{i=1}^{m}\sum_{j=1}^{n}|H_{i,j}|. \tag{7.8}$$

If all the submatrices are monochromatic, this lower bound is the infimum value of A.

Proof:

$$A = \frac{1}{mn}\sum_{k=1}^{K} m_k n_k \hat{C}_{\mathbf{F}}(H_k) \geq \frac{2}{mn}\sum_{k=1}^{K}\sqrt{m_k n_k}\sum_{l=1}^{l_k}\sigma_l(\mathbf{H}_k)/\sigma_l(\mathbf{F}) \tag{7.9}$$

$$\geq \frac{2}{mn\sigma_1(\mathbf{F})}\sum_{k=1}^{K}\sqrt{m_k n_k}\sum_{l=1}^{l_k}\sigma_l(\mathbf{H}_k) \tag{7.10}$$

$$\geq \frac{2}{mn\sigma_1(\mathbf{F})}\sum_{k=1}^{K}\left(m_k n_k \sum_{l=1}^{l_k}\sigma_l^2(\mathbf{H}_k)\right)^{1/2} = \frac{2}{mn\sigma_1(\mathbf{F})}\sum_{k=1}^{K}\left(m_k n_k \sum_{i=1}^{m_k}\sum_{j=1}^{n_k} H_k^2(i,j)\right)^{1/2} \tag{7.11}$$

$$\geq \frac{2}{mn\sigma_1(\mathbf{F})}\sum_{k=1}^{K}\sum_{i=1}^{m_k}\sum_{j=1}^{n_k}|H_k(i,j)| = \frac{2}{mn\sigma_1(\mathbf{F})}\sum_{i=1}^{m}\sum_{j=1}^{n}|H_{i,j}|. \tag{7.12}$$

(7.12) follows from the well-known inequality that for any real numbers, $(x_1, \dots, x_p)$

$$p\sum_{i=1}^{p} x_i^2 \geq \left(\sum_{i=1}^{p} x_i\right)^2 \tag{7.13}$$

with equality holding if and only if all the x_i's equal to each other.

According to Theorem 5.1 the infimum value of A is given by (7.9). If all the submatrices are monochromatic, the other inequalities, namely (7.10), (7.11), and (7.12) all become equality, and the last expression in (7.12) is the infimum value. ■

In the first phase of a two-phase protocol, the agents communicate with each other via the dynamic system by means of ϵ signals. The bit sequence defined by the communication complexity protocol can be regarded as an algorithm to identify the chosen submatrix in a given partition. We can visualize the algorithm by moving down a binary tree, so that depending on the value of the bit sent by either one of the agents in the communication protocol, we descend from a given node to its left or right child. For additional details about communication protocol,we refer to [19] or [34]. To make explicit contact with [19], suppose that each submatrix in the given matrix partition is monochormatic. The number of leaves in the binary tree that defines the protocol is equal to the number of submatrices defining the partition, and each of these submatrices is mapped to one of the leaves of the binary tree. The maximum number of bits communicated in the protocol is equal to two times the depth of the tree. (Since the communication has to pass through the dynamic system, if Alice wants to send one bit of information to Bob, the bit has to pass from Alice to the dynamic system and from the dynamic system to Bob, leading to two communication bits being exchanged.) The *protocol complexity* is thus defined as two times the height the height of the binary tree and can provide an upper bound for control communication complexity.

For illustration, consider a target function,

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 5 & 5 \\ 2 & 3 & 5 & 5 \\ 2 & 3 & 5 & 5 \end{bmatrix} \tag{7.14}$$

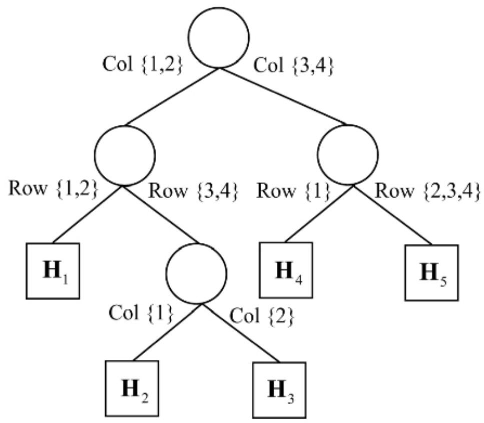
The submatrix partition shown in (7.4) represents a decomposition of the target matrix into the minimum number of monochromatic blocks. (Such minimal decomposition is not unique.) One can define a communication protocol to identify the components of the matrix partition (monochromatic blocks in this case) as follows.

Protocol to communicate the structure of a matrix partition: Assume that Alice controls the choice of columns and Bob controls the choice of rows.

1) Alice sends a bit to Bob with value 0 if she chooses the first 2 columns, otherwise she sends 1.
2) Upon receiving a bit of value 1 from Alice, Bob sends a bit to Alice with value 0 if he chooses the first row, otherwise, he sends a bit of value 1. No further bit needs to be sent.
3) On the other hand if a bit of value 0 is received from Alice, Bob sends a bit to Alice with value 0 if he chooses the first two rows, otherwise he sends a bit 1. Only in the latter case, Alice sends one more bit, with value 0 if she chooses the first column and 1 if she chooses the second column.

The communication protocol can be represented by the binary tree shown in Figure 2. A maximum of six bits (counting bits sent by the dynamic system) are needed in this protocol in order to guarantee all submatrices can be identified.

In our two-phase protocol, once the communication phase is completed—which is to say that the phase-one protocol has run to completion and a leaf node identifying a submatrix has been reached, the second phase of the protocol starts. The target matrix that is collaboratively evaluated by the agents in phase two is the submatrix that had been selected in phase one. It is assumed that in phase one, communication of negligible cost (e.g. ϵ-signaling) occurs, but that in phase two, open-loop controls of the form described in Section 5 realize the output specified in the chosen submatrix. It is clear that one can construct two-phase protocols with total control cost arbitrarily close to (7.1), the lower bound for all protocols realizing $\mathbf{H}$. We summarize the results in this section in the following theorem.



$$\mathbf{H}_1 = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}, \mathbf{H}_2 = \begin{bmatrix} H_{31} \\ H_{32} \end{bmatrix}, \mathbf{H}_3 = \begin{bmatrix} H_{41} \\ H_{42} \end{bmatrix}, \mathbf{H}_4 = \begin{bmatrix} H_{13} & H_{14} \end{bmatrix}, \mathbf{H}_5 = \begin{bmatrix} H_{23} & H_{24} \\ H_{33} & H_{34} \\ H_{43} & H_{44} \end{bmatrix}.$$

Fig. 2. Binary tree of a communication protocol to realize the partition defined by the matrix in equation (7.4).

Theorem 7.2. *Consider a bounded, bilinear input-output mapping, F, with a regular matrix representation $\mathbf{F}$. Suppose both agents have ϵ-signaling capability around the initial state. Let $\mathbf{H}$ be an m-by-n target matrix. The infimum control cost of any multi-round protocol that realizes $\mathbf{H}$ is given by:*

$$J(\mathbf{H}) = \frac{2}{mn\sigma_1(\mathbf{F})} \sum_{i=1}^{m} \sum_{j=1}^{n} |H_{i,j}|. \tag{7.15}$$

By comparing the control cost of an optimal single round protocol as given by (5.1) with that of the multi-round protocol (7.15), one can estimate the value of communicated bits in reducing the control energy cost. Using the target matrix $\mathbf{H}$ of (7.14) as an example, for the B-H system the minimal control cost without using any communication is defined in equation (5.2) and has a value of 7.6795π. If communication cost is negligible then the minimal control cost regardless of the number of bits communicated is given by right-hand-side of (7.8) and has a value of 2.875π. This can be achieved by the information sharing protocol shown in Figure 2, which has a protocol complexity of 6 (= two times the length of the binary tree). The protocol complexity as we have defined it is a “worst case” metric, reflecting the maximum number of bits that might need to be communicated. Thus, the value of a single communication bit in reducing the control energy cost for this problem is at least 0.8π with the units being control energy (as defined by (2.8)) per bit.

8. CONCLUSION

This paper has continued our study of problems in control communication complexity, which may be viewed as an extension of classical communication complexity with the additional focus on control cost. There are several important application contexts in which the optimization problems of the type we have considered seem to arise naturally. The single round protocols for steering the B-H system realizes the solution to a problem in distributed computing where independent agents act to evaluate a function without foreknowledge of each other's choices. As was noted in [2], the problems also arise naturally in what we have called the *standard parts optimal control problem* in which it is desired to find a specific number, m, of control inputs to a given input-output system that can be used in different combinations to attain a certain number, n, of output objectives so as to minimize the cost of control averaged over the different objectives. As noted above and more explicitly in [35], such optimization problems are of interest in the context of quantum computing, and more recently similar problems have been discussed in connection with protocols for assembly of molecular components in synthetic biology ([20]). The connection with quantum control, and in particular, the control of quantum spin systems (see for example the references to quantum control and computation in [35]) is another interesting line of investigation that is under way and will be treated elsewhere.

REFERENCES

[1] E. J. Anderson and R. R. Weber, "The Rendezvous Problem on Discrete Locations," *J. Appl. Prob.*, 28 pp. 839-851, 1990.

[2] J. Baillieul and W.-S. Wong, "The Standard Parts Problem and the Complexity of Control Communication," *Proc. of the 48-th IEEE Conf. on Decision and Control*, Shanghai, pp. 2723-2728, December 16-18, 2009.

[3] T. Basar and G.J. Olsder, *Dynamic Noncooperative Game Theory,* SIAM, Philadelphia, 1995.

[4] R.W. Brockett, "Control theory and singular Riemannian geometry," in *New Directions in Applied Mathematics*, P.S. Hilton and G.S Young, Eds., Springer, Berlin, 1981, pp. 11-27.

[5] R.W. Brockett, "Asymptotic Stability and Feedback Stabilization," in *Differential Geometric Control Theory*, R. W. Brockett, R. S. Millman, and H. J. Sussman, Eds. Boston, Birkhauser, pp. 181-191, 1983.

[6] R.W. Brockett, "On the Rectification of Vibrating Motion," *Sensors and Actuators*, 20 pp. 91-96, 1989.

[7] R.W. Brockett, "Cycles that Effect Change," in *Motion, Control, and Geometry: Proceedings of a Symposium* (1997), National Academies Press, pp. 20-31, 1997.

[8] R.W. Brockett, "Pattern Generation and the Control of Nonlinear Systems," *IEEE Trans. on Automatic Control*, 48:10, pp. 1699-1711, 2003.

[9] R. Carli and F. Bullo, "Quantized Coordination Algorithms for Rendezvous and Deployment", *SIAM Journal on Control and Optimization*, 48(3), pp. 1251–1274, 2009.

[10] J. Cortés, S. Martínez, and F. Bullo, "Robust Rendezvous for Mobile Autonomous Agents via Proximity Graphs in Arbitrary Dimensions", *IEEE Trans. Automat. Contr.*, 51, pp. 1289–1298, 2006.

[11] D. D'Alessandro and M. Dahleh, "Optimal Control of Two-Level Quantum Systems," *IEEE Trans. on Automatic Control*, 46:6 pp. 866-876, 2001.

[12] D.F. Delchamps, "Stabilizing a Linear System with Quantized State Feedback", *IEEE Trans. Automat. Contr.*, 35, pp. 916–924, 1990.

[13] W.B. Dunbar and R. Murray, "Distributed Receding Horizon Control for Multi-Vehicle Formation Stabilization", *Automatica*, 42, pp. 549–558, 2006.

[14] N. Elia and S.K. Mitter "Stabilization of Linear Systems with Limited Information", *IEEE Trans. Automat. Contr.*, 49, pp. 1384–1400, 2001.

[15] F. Fagnani and S. Zampieri, "Stability Analysis and Synthesis for Scalar Linear Systems with a Quantized Feedback", *IEEE Trans. Automat. Contr.*, 48, pp. 1569–1584, 2003.

[16] J.A. Fax and R. Murray, "Information Flow and Cooperative Control of Vehicle Formations", *IEEE Trans. Automat. Contr.*, 49, pp. 1465–1476, 2004.

[17] G. Guo and W.S. Wong, "Optimal Control of Multi-Agent Systems Allowing Input Choices", Preprint.

[18] Ho, Yu-Chi, 1980. "Team Decision Theory and Information Structures," *Proceedings of the IEEE*, v. 68:6, pp. 644-654.

[19] E. Kushilevitz & N. Nisan, *Communication Complexity*, Cambridge University Press, Cambridge, UK, 1997.

[20] T. Knight,"Idempotent Vector Design for Standard Assembly of Biobricks," MIT Synthetic Biology Working Group, 2006. [Online]. Available: http://dspace.mit.edu/handle/1721.1/21168

[21] N.E. Leonard and P.S. Krishnaprasad, "Motion Control of Drift-Free, Left-Invariant Systems on Lie Groups," *IEEE Trans. on Automatic Control*, 40:9, pp. 1539-1554, 1995.

[22] K. Li and J. Baillieul, "Robust Quantization for Digital Finite Communication Bandwidth (DFCB) Control", *IEEE Trans. Automat. Contr.*, 49, pp. 1573–1584, 2004.

[23] J.-S. Li and N. Khaneja, "Ensemble Controllability of the Bloch Equations," in *Proc. of the 45-th IEEE Conference on Decision and Control*, San Diego, pp. 2483-2487, December 13-15, 2006.

[24] J.-S. Li and N. Khaneja, "Ensemble Control of the Bloch Equations," *IEEE Trans. on Automatic Control*, 54:3, pp. 528-536, 2009.

[25] J.-S. Li, J. Ruths, T.-Y. Yu, H. Arthanari, and G. Wagner, "Optimal Pulse Design in Quantum Control: A Unified Computational Method," To appear in *PNAS*.

[26] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, Orlando, 1979.

[27] R. Montgomery, *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, (Mathematical Surveys and Monographs, Volume 91), American Mathematical Society, ISBN 0-8218-1391-9, 2002.

[28] G.N. Nair and R.J. Evans, "Exponential Stability of Finite-Dimensional Linear Systems with Limited Data Rates", *Automatica*, 39 (2002), pp. 585–593.

[29] R. Olfati-Saber and R. Murray, "Consensus Problems in Networks of Agents With Switching Topology and Time-Delays", *IEEE Trans. Automat. Contr.*, 49, pp. 1520–1533, 2004.

[30] R. Osserman, “The Isoperimetric Inequality”, *Bull. of American Math. Soc.*, 84:6 pp. 1182–1238, 1978.

[31] D.V. Raghunathan and J. Baillieul, “Communicating Information in the Null-Space of Linear and Nonlinear Systems,” Preprint.

[32] S. Tatikonda and S.K. Mitter, “Control Under Communication Constraints”, *IEEE Trans. Automat. Contr.*, 49, pp. 1056-1068, 2004.

[33] H. Weyl, *Gruppentheorie und Quantenmechanik*, Zweite Auflage, 366 S. S. Hirzel, Leipzig 1931.

[34] W.S. Wong, “Control Communication Complexity of Distributed Control Systems,” *SIAM Journal of Control and Optimization*, 48:3 pp. 1722–1742, 2009.

[35] W.S. Wong and J. Baillieul, “Control Communication Complexity of Nonlinear Systems,” *Communications in Information and Systems*, 9:1, pp. 103-140, 2009. http://people.bu.edu/johnb/CIS-9-1-A5-wong.pdf

[36] W.S. Wong and R.W. Brockett, “Systems with Finite Communication Bandwidth Constraints I: State Estimation Problems,” *IEEE Trans. Automat. Contr.*, 42, pp. 1294–1298, 1997.

[37] W.S. Wong and R.W. Brockett, “Systems with Finite Communication Bandwidth Constraints II: Stabilization with Limited Information Feedback”, *IEEE Trans. Automat. Contr.*, 44, pp. 1049–1053, 1999.

[38] A.C.-C. Yao, “Some Complexity Questions Related to Distributive Computing,” *Proc. of the 11-th Annual ACM Symposium on Theory of Computing (STOC)*, 1979.

APPENDIX A

To prove the rank of a matrix representation of a bilinear input-output mapping is well-defined, let $\mathbf{F}$ be a representation corresponding to the bases $\{e_1, e_2, \ldots\}$ and $\{f_1, f_2, \ldots\}$. Let $\mathbf{F}'$ be a representation corresponding to the bases $\{e_1, e_2, \ldots\}$ and $\{f'_1, f'_2, \ldots\}$. Assume first that $\mathbf{F}$ has finite rank equal to r. By relabeling the indices, we can assume without loss of generality that the first r columns of $\mathbf{F}$ are independent. Hence, there exists a matrix $\mathbf{C}$ with r rows and infinitely many columns such that

$$\mathbf{F} = \begin{bmatrix} F_{11} & F_{12} & F_{13} & \ldots \\ F_{21} & F_{22} & F_{23} & \ldots \\ F_{31} & F_{32} & F_{33} & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} & \ldots & F_{1r} \\ F_{21} & F_{22} & \ldots & F_{2r} \\ F_{31} & F_{32} & \ldots & F_{3r} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \mathbf{C} \equiv \bar{\mathbf{F}}\mathbf{C}. \tag{8.1}$$

By relabeling the indices of the basis $\{e_1, e_2, \ldots\}$ if necessary, we can assume that the first r rows $\bar{\mathbf{F}}$ are independent. Label the submatrix of $\bar{\mathbf{F}}$ formed by the first r rows by $\mathbf{D}$. Note that $\mathbf{D}$ is an invertible r-by-r matrix and

$$\mathbf{C} = \mathbf{D}^{-1} \begin{bmatrix} F_{11} & F_{12} & F_{13} & \ldots \\ F_{21} & F_{22} & F_{23} & \ldots \\ \vdots & \vdots & \vdots & \vdots \\ F_{r1} & F_{r2} & F_{r3} & \ldots \end{bmatrix}. \tag{8.2}$$

That is, for $1 \leq k \leq r$ and $1 \leq l$

$$C_{kl} = \sum_{i=1}^{r} D_{ki}^{-1} F_{il}. \tag{8.3}$$

Let C_0 be the upper bound of the entries in $\mathbf{D}^{-1}$, so that for all $1 \leq i, j \leq r$,

$$D_{ij}^{-1} \leq C_0 < \infty. \tag{8.4}$$

For any positive integer j, the base vector, f'_j, has a representation of the form

$$f'_j = \sum_{l=1}^{\infty} \alpha_{jl} f_l. \tag{8.5}$$

By definition, the coefficients are square summable so that

$$\sum_{l=1}^{\infty} \alpha_{jl}^2 < \infty. \tag{8.6}$$

For any $1 \leq i \leq r$, define

$$g_{ij} = \sum_{l=1}^{\infty} \alpha_{jl} sgn(\alpha_{jl} F_{il}) f_l. \tag{8.7}$$

Since the coefficients for g_{ij} are square summable, this is a well-defined element in $\mathcal{L}$. Moreover, F is bounded implies

$$F(e_i, g_{ij}) = \sum_{l=1}^{\infty} |\alpha_{jl} F_{il}| < \infty. \tag{8.8}$$

On the other hand,

$$F'_{ij} = F(e_i, f'_j) = \sum_{l=1}^{\infty} \alpha_{jl} F_{il} = \sum_{l=1}^{\infty} \alpha_{jl} \sum_{k=1}^{r} F_{ik} C_{kl}. \tag{8.9}$$

Formally, by exchange the summation order, the right-hand-side of equation (8.9) is equal to

$$\sum_{k=1}^{r} F_{ik} \sum_{l=1}^{\infty} \alpha_{jl} C_{kl} = \sum_{k=1}^{r} F_{ik} E_{jk}, \quad E_{jk} \equiv \sum_{l=1}^{\infty} \alpha_{jl} C_{kl}. \tag{8.10}$$

The equality holds if the expression in (8.10) is convergent. By means of equation (8.3),

$$|E_{jk}| \leq \sum_{l=1}^{\infty} \sum_{i=1}^{r} |\alpha_{jl} D_{ki}^{-1} F_{il}| \leq C_0 \sum_{i=1}^{r} \sum_{l=1}^{\infty} |\alpha_{jl} F_{il}|. \tag{8.11}$$

It follows that the expression in (8.10) is convergent and

$$F'_{ij} = \sum_{l=1}^{\infty} \alpha_{jl} \sum_{k=1}^{r} F_{ik} C_{kl} = \sum_{k=1}^{r} F_{ik} \sum_{l=1}^{\infty} \alpha_{jl} C_{kl}. \tag{8.12}$$

Then equation (8.12) can be rewritten as

$$\mathbf{F}' = \bar{\mathbf{F}}\mathbf{E},\ \mathbf{E} \equiv \begin{bmatrix} E_{11} & E_{21} & E_{31} & \dots \\ E_{12} & E_{22} & E_{32} & \dots \\ \vdots & \vdots & \vdots & \vdots \\ E_{1r} & E_{2r} & E_{3r} & \dots \end{bmatrix}. \tag{8.13}$$

Therefore, the rank of $\mathbf{F}'$ is equal to or less than r. By reversing the roles of $\mathbf{F}$ and $\mathbf{F}'$, the ranks of the two matrices must be identical if one of them has finite rank. Thus, if the basis $\{e_1, e_2, \dots\}$ is fixed the rank is independent of the other basis. Similarly, one can argue that if $\{f_1, f_2, \dots\}$ is fixed, the rank is independent of the other basis.

In the general case, consider matrix representation, $\mathbf{F}$, with respect to bases $\{e_1, e_2, \dots\}$ and $\{f_1, f_2, \dots\}$, and matrix representation, $\mathbf{F}''$, with respect to bases $\{e_1', e_2', \dots\}$ and $\{f_1', f_2', \dots\}$. If the rank of $\mathbf{F}$ is finite, it is equal to the rank of the matrix representation with respect to the bases $\{e_1, e_2, \dots\}$ and $\{f_1', f_2', \dots\}$, which in turn is equal to the rank $\mathbf{F}''$. It is easy to see that it is not possible for one representation to have finite rank while the other has an infinite rank. Thus the rank is independent of the choice of the bases.

APPENDIX B

Let $\{e_1, e_2, \dots\}$ and $\{f_1, f_2, \dots\}$ be orthonormal bases corresponding to the representation of $\mathbf{F}$. Let v be an element of $\mathcal{L}$ which is contained completely in the subspace spanned by $\{f_1, \dots, f_l\}$. Let $\mathbf{v}$ represent the corresponding l-dimensional column vector. The vector $\mathbf{F}_l\mathbf{v}$ can be regarded as an element contained in the subspace spanned by $\{e_1, \dots, e_l\}$, denote the corresponding element in $\mathcal{L}$ by u. Then,

$$F(u, v) = F(\sum_{i=1}^{l} u_i e_i, \sum_{j=1}^{l} v_j f_j) = \sum_{i=1}^{l}\sum_{j=1}^{l} u_i v_j F_{ij} = \mathbf{v}^T \mathbf{F}_l^T \mathbf{F}_l \mathbf{v}. \tag{9.1}$$

Using the assumption that F is bounded and inequality (2.6), it follows that

$$\|\mathbf{F}_l\mathbf{v}\|^2 = |\mathbf{v}^T \mathbf{F}_l^T \mathbf{F}_l \mathbf{v}| = \|F(u, v)\| \le \|F\|\|\mathbf{v}\|\|\mathbf{F}_l\mathbf{v}\|. \tag{9.2}$$

Hence,

$$\|\mathbf{F}_l\mathbf{v}\| \le \|F\|\|\mathbf{v}\|. \tag{9.3}$$

Let $\mathbf{w}_i$ be the eigenvector of $\mathbf{F}_l^T\mathbf{F}_l$ with unit norm that corresponds to $\sigma_i(\mathbf{F}_l)$. Then,

$$\sigma_i^2(\mathbf{F}_l) = \mathbf{w}_i^T \mathbf{F}_l^T \mathbf{F}_l \mathbf{w}_i = \|\mathbf{F}_l\mathbf{w}_i\|^2 \le \|F\|\|\mathbf{F}_l\mathbf{w}_i\| \le \|F\|^2. \tag{9.4}$$