Cooperative Control of Linear Systems with Choice Actions*

Zhongchang Liu[1], Wing Shing Wong[1] and Ge Guo[2]

***Abstract*— In this paper, an estimate-inference-feedback control methodology is proposed for affine systems involving two agents executing cooperative control based on individual choices. No explicit communication channel exists between the two agents. The system state is estimated independently with an unbiased minimum variance (UMV) estimator by each agent and knowledge about their choice actions are updated in discrete time steps. Based on the estimated system state and probability distributions of the choices, sub-optimal controllers are designed iteratively. Analysis and simulation results show that the proposed control law is robust to disturbances and more energy-efficient than strategies ignoring choice information.**

I. INTRODUCTION

Cooperative decision making involving a group of participants is an interesting and thriving topic for researchers from various disciplines. Typically, the studied problem is formulated as a team task with some uncertainty elements, which cannot be accomplished by a single controlling participant, also known as agent. Thus, each agent has to take into consideration of one's own decision along with the information obtained on the other agents' decisions. The objective is to find strategies for each agent such that a given objective goal is achieved. The information structure, that is, "who knows what and when do they know it", plays a crucial role in the design problem of strategies, or equivalently, the problem of "what should they do with that information". These problems have long been considered in the literatures, for example [1]–[6], [16] to name a few.

One scenario falling into the category of team decision problem is the choice-based actions of multiple agents, where agents cooperate to realize a target selected from a common target list according to the joint but independent choices of the agents. Application examples include controlling the position of a wireless sensor according to the monitoring regions of interest to the non-communicating users, replenishing the stock level of a warehouse according to the demand levels of multiple retailers, determining to rendezvous or not based on people's moods (refer to [10] for the detailed description of the rendezvous problem), among others. This line of investigation was first drawn into attention in [7] and extended in successive literatures [8]–[11]. In [7] and [8], the authors investigated the interplay of control and communication and formally introduced the concept of control communication complexity and control energy complexity of implementing such tasks. In [9], this problem was further connected with the standard parts optimal control problem as suggested by Brockett in [12]. The distributed realization of a target matrix was extended to bilinear input-output mapping systems in [10]. In general, it is possible for the agents to lower the control cost by signaling information of their choices to each other. Two bounding extreme cases were analyzed in [10], one without any communication and the other with full communication, in order to determine inherent value of a communication bit. These research works lead to new perspectives on distributed control and at the same time raise numerous challenging questions, such as the feasibility of the set of targets, the optimality of control protocols etc.

In this paper, we study the distributed action affine systems in details, noting that the rendezvous problem [10] and many agent related problems [14]–[16] can be described in terms of affine models. Unlike previous works [7]–[10] which focus on the complexity analysis, the design of distributed control schemes is considered in this paper. Even if one could design open-loop controls for such problems for deterministic systems as in [11], these solutions usually do not work well under state uncertainties. Therefore, our focus in this paper is on feedback schemes that allow control decisions to be based on information obtained through partial observations. Due to the lack knowledge of each other's choice and hence control input, a special type of Kalman filter is introduced to handle uncertainties in the system state and observations. Mutual online inference of each other's selected choice is embedded in the distributed feedback control law. The resultant algorithm improves overall performance accuracy while lowering the control cost for target realization without requiring explicit communication between agents on their selected choices.

This paper is organized as follows. In Section II, we provide the description of the basic model and the main problem to be solved. In Section III, the UMV state estimator and the choice information inference procedure are presented. In Section IV, both open-loop control and estimate-infer-feedback control (EIFC) algorithms are presented. Simulation results in Section V provide partial confirmation of our claims along with comparison between different methods. Conclusions and future work are provided in Section VI.

II. PROBLEM STATEMENT

While [11] investigated the controller design methodology for deterministic continuous-time systems, in the present paper we focus on controlled stochastic systems by two-agent, Alice and Bob. Suppose the choices of Alice are labeled by a

*This work is supported by the National Natural Science Foundation of China under grant number 61174060.

[1]Z. C. Liu and W. S. Wong are with Department of Information Engineering, the Chinese University of Hong Kong, Hong Kong, China. zcliu@ie.cuhk.edu.hk, wswong@ie.cuhk.edu.hk

[2]G. Guo is with School of Control Science and Engineering, Dalian University of Technology, Dalian, China. geguo@yeah.net

finite set $\mathscr{A} = \{1,2,\ldots,N_A\}$, and $\mathscr{B} = \{1,2,\ldots,N_B\}$ is used to label Bob's choices. When $i \in \mathscr{A}$ is selected by Alice and $j \in \mathscr{B}$ is selected by Bob, the discrete-time model of affine systems can be described by

$$\mathbf{x}_{k+1}^{ij} = \mathbf{A}\mathbf{x}_k^{ij} + \mathbf{B}\mathbf{u}_k^i + \mathbf{C}\mathbf{v}_k^j + \xi_k, \tag{1}$$
$$\mathbf{z}_k^{A,ij} = \mathbf{F}^A\mathbf{x}_k^{ij} + \eta_k^A, \tag{2}$$
$$\mathbf{z}_k^{B,ij} = \mathbf{F}^B\mathbf{x}_k^{ij} + \eta_k^B, \tag{3}$$

where $\mathbf{x}_k^{ij} \in \mathbb{R}^n$ and $\mathbf{z}_k^{ij} \in \mathbb{R}^p$ (in the sequel, we omit the superscripts ij where there is no ambiguity) are respectively the system state and the agents' measurements of the state at time k, $k = 0,\ldots,N-1$. $\mathbf{x}_0$ is assumed a random vector with mean $\hat{\mathbf{x}}_0$ and variance $\mathbf{Q}_0$. The control input of Alice corresponding to her choice i is represented by $\mathbf{u}_k^i \in \mathbb{R}^{m_u}$ and the control input of Bob corresponding to his choice j by $\mathbf{v}_k^j \in \mathbb{R}^{m_v}$. The noises $\xi \in \mathbb{R}^n$ and $\eta \in \mathbb{R}^p$ are white Gaussian with covariances $E\{\xi_k\xi_l^T\} = \mathbf{Q}_k\delta_{kl}$, $E\{\eta_k\eta_l^T\} = \mathbf{R}_k\delta_{kl}$, $E\{\xi_k\eta_l^T\} = \mathbf{0}$, and are uncorrelated with $\mathbf{x}_k$.

Both agents select their choices at initial time and the choices remain unchanged. At terminal time, the target state to be realized is represented by $\mathbf{H}_{ij}$ when Alice selects choice i and Bob choice j. Hence, we can compactly summarize the targets via a matrix $\mathbf{H}$ which is known as the *target matrix*. The target matrix is said to be *compatible* if for any indices $i,i' \in \mathscr{A}$, and $j,j' \in \mathscr{B}$, its entries satisfy

$$\mathbf{H}_{ij} - \mathbf{H}_{i'j} = \mathbf{H}_{ij'} - \mathbf{H}_{i'j'}, \tag{4}$$

otherwise, it is called *incompatible*. In this paper, we only consider compatible targets, further discussions on how to realize a incompatible target matrix can be found in [11] for multi-agent affine systems.

It is assumed that the entire choice set of Alice $\mathscr{A}$ and the corresponding control inputs are known to Bob, but the choice she makes at the initial time is not directly disclosed to him. Similar assumption applies to Bob's choice decision. Therefore, the joint state maneuvering problem (or standard parts problem stated in [9]) is distributed and in general may require the agents to gain at least partial knowledge of each other's choice for final solution. For affine systems, the exceptional cases where no direct communication on the selected choices is required for solution are exactly those with compatible target matrices. In [11], one can find solution to the design problem of deriving optimal distributed control inputs $\mathbf{u}^i$, $\mathbf{v}^j$ for all $i \in \mathscr{A}$ and $j \in \mathscr{B}$ such that all targets in a given compatible matrix $\mathbf{H}$ can be achieved when chosen by the agents, subject to the cost functional

$$J^o(\kappa) = E\{\sum_{k=\kappa}^{N-1}[\frac{1}{N_A}\sum_{i=1}^{N_A}(\mathbf{u}_k^i)^T\mathbf{u}_k^i + \frac{1}{N_B}\sum_{j=1}^{N_B}(\mathbf{v}_k^j)^T\mathbf{v}_k^j]\}. \tag{5}$$

Note that this cost functional highlights the fact that the control cost is averaged over all possible event outcomes when different combinations of choice actions are taken. If the particular choice pair (i,j) can be made known to both agents, then the above cost function will be reduced to a centralized one

$$J^c(\kappa) = E\{\sum_{k=\kappa}^{N-1}[(\mathbf{u}_k^i)^T\mathbf{u}_k^i + (\mathbf{v}_k^j)^T\mathbf{v}_k^j]\},$$

which requires less control cost. This motivates us to solve the problem by allowing the agents to iteratively estimate each other's choice while solving a series of related optimization problems by replacing the initial choice estimates in (5) with the latest choice estimates.

The following technical assumptions are required for our proposed solution approach.

Assumption 1: For the system (1)-(3), we assume that:

1) $\mathbf{A}$ is nonsingular.
2) The system is controllable by each individual agent, i.e., $(\mathbf{A},\mathbf{B})$ and $(\mathbf{A},\mathbf{C})$ are controllable pairs.
3) The system is observable by each individual agent, i.e., $(\mathbf{A},\mathbf{F}^A)$ and $(\mathbf{A},\mathbf{F}^B)$ are observable pairs.

Assumption 2: The two agents have analogous input dynamics, i.e., $\mathbf{C} = \gamma\mathbf{B}$, $\gamma \in \mathbb{R}, \gamma \neq 0$, and the same measurement $\mathbf{z}_k$, i.e. $\mathbf{z}_k^A = \mathbf{z}_k^B \equiv \mathbf{z}_k = \mathbf{F}\mathbf{x}_k + \eta_k$.

Remark 1: The first part of the assumption 2 arises naturally from scenarios where agents are peers in the system [14]; heterogeneous cases [15] are not considered in this paper. The latter part of the assumption is necessitated by our restriction that there is no communication among agents, unlike traditional models which assume information flows either all-to-all or among neighboring agents ([16], [17]). This information structure can substantially mitigate the difficulties introduced by the noise η, so that we can concentrate on the uncertainties caused by the distributed choices. Scenarios in which this assumption holds include those cases where agents share data from a common observer.

III. Estimation-Based Choice Discrimination

In a choice-based system, the target matrix is assumed to be known to both agents. If the entries $\mathbf{H}_{ij}$ are distinct, it is possible to develop estimators for the agents to accurately determine the choice selected by the other agent. As a consequence, although no explicit communication is allowed the system dynamic can convey enough information to the agents even under state and observation noises.

In this section, we present a two-stage choice determination procedure. In the first stage of a time interval $[k,k+1]$, agents estimate the state of the system separately. In the second stage, each agent makes an inference on the other agent's choice based on the estimated state. These procedures are called the *estimate and infer* approach which has similar flavor with the communication phase of a two-phase protocols depicted in [10], in which partial information is permitted to be shared with negligible cost.

A. State Estimators

Conventionally, with multiple agents contributing to a common dynamical platform, inputs from the other agents can be viewed as unknown exogenous inputs from the viewpoint of a particular observer. As a result, the Kitanidis' estimator [18] [19] can be applied to estimate the system state. This works because the Kitanidis' estimator is a type of augmented state Kalman filter that offers unbiased minimum variance (UMV) state estimation by bypassing the effect of the unknown input. In this subsection, we show that a feasible solution can be obtained by requiring the agents to use this type of estimators provided that the following assumption is satisfied.

Assumption 3: ([18], [19]) $rank\mathbf{FB} = rank\mathbf{B} = m_u = m_v$.

Remark 2: This assumption requires $p \geq m_u$, but is not limiting in engineering applications as discussed in [18]. To give an example, this assumption as well as Assumption 1 can be both satisfied for integrators which are used frequently as point models in the multi-agent literatures [14]. Another claim is that this estimator is not optimal in the sense of mean square error for our problem, since partial information about each other's control input is available for both agents.

Alice's state estimate: From the viewpoint of Alice, suppose she selects component $i \in \mathscr{A}$, the UMV estimation of $\mathbf{x}_k$ is given by [18]:

$$\hat{\mathbf{x}}^i_{k|k-1} = \mathbf{A}\hat{\mathbf{x}}^i_{k-1|k-1} + \mathbf{B}\mathbf{u}^i_{k-1} \tag{6}$$

$$\hat{\mathbf{x}}^i_{k|k} = \hat{\mathbf{x}}^i_{k|k-1} + \mathbf{L}^A_k(\mathbf{z}_k - \mathbf{F}\hat{\mathbf{x}}^i_{k|k-1}) \tag{7}$$

$$\begin{aligned}\mathbf{L}^A_k = \mathbf{K}_k + (\mathbf{I} - \mathbf{K}_k\mathbf{F})\mathbf{C}[\mathbf{C}^T\mathbf{F}^T(\mathbf{F}\hat{\mathbf{P}}_k\mathbf{F}^T \\ + \mathbf{R}_k)^{-1}\mathbf{F}\mathbf{C}]^{-1}\mathbf{C}^T\mathbf{F}^T(\mathbf{F}\hat{\mathbf{P}}_k\mathbf{F}^T + \mathbf{R}_k)^{-1}\end{aligned} \tag{8}$$

$$\mathbf{K}_k = \hat{\mathbf{P}}_k\mathbf{F}^T(\mathbf{F}\hat{\mathbf{P}}_k\mathbf{F}^T + \mathbf{R}_k)^{-1} \tag{9}$$

$$\hat{\mathbf{P}}_k = \mathbf{A}\mathbf{T}_{k-1}\mathbf{A}^T + \mathbf{Q}_{k-1} \tag{10}$$

$$\mathbf{T}_k = (\mathbf{I} - \mathbf{L}^A_k\mathbf{F})\hat{\mathbf{P}}_k(\mathbf{I} - \mathbf{L}^A_k\mathbf{F})^T + \mathbf{L}^A_k\mathbf{R}_k(\mathbf{L}^A_k)^T. \tag{11}$$

This estimator satisfies $E\{\mathbf{x}_k - \hat{\mathbf{x}}^i_{k|k}\} = \mathbf{0}$, $E\{(\mathbf{x}_k - \hat{\mathbf{x}}^i_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}^i_{k|k})^T\} = \mathbf{T}_k$, and $(\mathbf{L}^A_k\mathbf{F} - \mathbf{I})\mathbf{C} = \mathbf{0}$.

Bob's state estimate: On the side of Bob, if $j \in \mathscr{B}$ is chosen, the formulas are analogous except the a priori estimation $\hat{\mathbf{x}}_{k|k-1}$ and the gain $\mathbf{L}^B_k$ of the UMV estimator should be re-formulated as

$$\hat{\mathbf{x}}^j_{k|k-1} = \mathbf{A}\hat{\mathbf{x}}^j_{k-1|k-1} + \mathbf{C}\mathbf{v}^j_{k-1}, \tag{12}$$

$$\begin{aligned}\mathbf{L}^B_k = \mathbf{K}_k + (\mathbf{I} - \mathbf{K}_k\mathbf{F})\mathbf{B}[\mathbf{B}^T\mathbf{F}^T(\mathbf{F}\hat{\mathbf{P}}_k\mathbf{F}^T \\ + \mathbf{R}_k)^{-1}\mathbf{F}\mathbf{B}]^{-1}\mathbf{B}^T\mathbf{F}^T(\mathbf{F}\hat{\mathbf{P}}_k\mathbf{F}^T + \mathbf{R}_k)^{-1}.\end{aligned} \tag{13}$$

It follows that the two estimators satisfy the following property:

Lemma 1: Under Assumption 2, the two agents, Alice and Bob, have an identical UMV estimate $\hat{\mathbf{x}}_{k|k}$ of the system state $\mathbf{x}_k$ (with the superscript ij omitted) if $\hat{\mathbf{x}}^i_0 = \hat{\mathbf{x}}^j_0$.

Proof: Suppose that $\hat{\mathbf{x}}^i_{k-1|k-1} = \hat{\mathbf{x}}^j_{k-1|k-1} \equiv \hat{\mathbf{x}}_{k-1|k-1}$. Then

$$\begin{aligned}&\hat{\mathbf{x}}^i_{k|k} - \hat{\mathbf{x}}^j_{k|k} \\ =&(\mathbf{I} - \mathbf{L}^A_k\mathbf{F})(\mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}\mathbf{u}^i_{k-1}) + \mathbf{L}^A_k\mathbf{z}_k \\ &- [(\mathbf{I} - \mathbf{L}^B_k\mathbf{F})(\mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{C}\mathbf{v}^j_{k-1}) + \mathbf{L}^B_k\mathbf{z}_k] \\ =&(\mathbf{L}^A_k - \mathbf{L}^B_k)(\mathbf{z}_k - \mathbf{F}\mathbf{A}\hat{\mathbf{x}}_{k-1|k-1}) \\ &+ (\mathbf{I} - \mathbf{L}^A_k\mathbf{F})\mathbf{B}\mathbf{u}^i_{k-1} - (\mathbf{I} - \mathbf{L}^B_k\mathbf{F})\mathbf{C}\mathbf{v}^j_{k-1}.\end{aligned}$$

Noting that $\mathbf{C} = \gamma\mathbf{B}$ in Assumption 2, it follows from (8) and (13) that $\mathbf{L}^A_k = \mathbf{L}^B_k \equiv \mathbf{L}_k$. Also, $(\mathbf{I} - \mathbf{L}^A_k\mathbf{F})\mathbf{B} = (\mathbf{I} - \mathbf{L}^B_k\mathbf{F})\mathbf{B} = \mathbf{0}$, and $(\mathbf{I} - \mathbf{L}^B_k\mathbf{F})\mathbf{C} = (\mathbf{I} - \mathbf{L}^A_k\mathbf{F})\mathbf{C} = \mathbf{0}$ out of the unbiased property of the estimators. It is now clear that $\hat{\mathbf{x}}^i_{k|k} = \hat{\mathbf{x}}^j_{k|k} \equiv \hat{\mathbf{x}}_{k|k}$. ■

In the sequel, we use $\hat{\mathbf{x}}_{k|k}$ to denote the common state estimation and $\mathbf{L}_k$ to denote the common estimator gain used by both agents at step $0 \leq k \leq N$. Denote by $\mathscr{X}_k \triangleq \{\hat{\mathbf{x}}_0, \hat{\mathbf{x}}_{1|1}, \ldots, \hat{\mathbf{x}}_{k|k}\}$ the sequence of estimation up to time k. It can be shown that $\hat{\mathbf{x}}_{k|k}$ is a Gaussian random variable with mean

$$\bar{\mathbf{x}}^{ij}_{k|k} \triangleq E\{\hat{\mathbf{x}}_{k|k}|\mathscr{X}_{k-1}\} = \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}\mathbf{u}^i_{k-1} + \mathbf{C}\mathbf{v}^j_{k-1}, \tag{14}$$

and variance

$$\bar{\mathbf{P}}_k \triangleq var\{\hat{\mathbf{x}}_{k|k}|\mathscr{X}_{k-1}\} = \mathbf{L}_k(\mathbf{F}\hat{\mathbf{P}}_k\mathbf{F}^T + \mathbf{R}_k)\mathbf{L}^T_k. \tag{15}$$

B. Choice Inference

Once the system state is estimated, the agents can use the information to estimate the choice selected by the other agents. Before we explain how this can be done, we need to introduce more notation: denote by α^l when Alice's choice is $l \in \mathscr{A}$, that is, $\alpha = l$, and by β^m when $\beta = m \in \mathscr{B}$.

Alice's inference: Suppose the choice of Alice is i. Then, for all $m \in \mathscr{B}$ the probability density function (PDF) of $\hat{\mathbf{x}}_{k|k}$ can be computed by

$$\begin{aligned}&p(\hat{\mathbf{x}}_{k|k}|\beta^m, \alpha^i, \mathscr{X}_{k-1}) \\ =&\frac{1}{(2\pi)^{\frac{r}{2}}(\det^*\bar{\mathbf{P}}_k)^{\frac{1}{2}}}\exp\{-\frac{1}{2}(\hat{\mathbf{x}}_{k|k} - \bar{\mathbf{x}}^{im}_{k|k})^T\bar{\mathbf{P}}^+_k(\hat{\mathbf{x}}_{k|k} - \bar{\mathbf{x}}^{im}_{k|k})\}.\end{aligned} \tag{16}$$

where $r = rankP$, $\det^*\mathbf{M}$ and $\mathbf{M}^+$ are respectively the pseudo-determinant and pseudo-inverse of a singular matrix $\mathbf{M}$. Given that at the $(k-1)$-th time step the probability $P(\beta^m|\alpha^i, \mathscr{X}_{k-1})$ of hypothesis β^m is known, this probability can be updated at the k-th step by using all the above newly calculated PDFs $p(\hat{\mathbf{x}}_{k|k}|\beta^m, \alpha^i, \mathscr{X}_{k-1})$, $m = 1, \ldots, N_B$:

$$\begin{aligned}P(\beta^m|\alpha^i, \mathscr{X}_k) &= \frac{p(\beta^m, \hat{\mathbf{x}}_{k|k}|\alpha^i, \mathscr{X}_{k-1})}{p(\hat{\mathbf{x}}_{k|k}|\alpha^i, \mathscr{X}_{k-1})} \\ =&\frac{p(\hat{\mathbf{x}}_{k|k}|\beta^m, \alpha^i, \mathscr{X}_{k-1})P(\beta^m|\alpha^i, \mathscr{X}_{k-1})}{\sum_{m=1}^{N_B} p(\hat{\mathbf{x}}_{k|k}|\beta^m, \alpha^i, \mathscr{X}_{k-1})P(\beta^m|\alpha^i, \mathscr{X}_{k-1})}.\end{aligned} \tag{17}$$

Bob's inference: Assume j is selected by Bob, and then the likelihood of $\hat{\mathbf{x}}_{k|k}$ hypothesized on Alice's action α^l, $l = 1, \ldots, N_A$ is given by

$$\begin{aligned}&p(\hat{\mathbf{x}}_{k|k}|\alpha^l, \beta^j, \mathscr{X}_{k-1}) \\ =&\frac{1}{(2\pi)^{\frac{r}{2}}(\det^*\bar{\mathbf{P}}_k)^{\frac{1}{2}}}\exp\{-\frac{1}{2}(\hat{\mathbf{x}}_{k|k} - \bar{\mathbf{x}}^{lj}_{k|k})^T\bar{\mathbf{P}}^+_k(\hat{\mathbf{x}}_{k|k} - \bar{\mathbf{x}}^{lj}_{k|k})\},\end{aligned} \tag{18}$$

and, the probability of hypothesis α^l at the k-th step is updated by Bayes' rule

$$\begin{aligned}&P(\alpha^l|\beta^j, \mathscr{X}_k) \\ =&\frac{p(\hat{\mathbf{x}}_{k|k}|\alpha^l, \beta^j, \mathscr{X}_{k-1})P(\alpha^l|\beta^j, \mathscr{X}_{k-1})}{\sum_{l=1}^{N_A} p(\hat{\mathbf{x}}_{k|k}|\alpha^l, \beta^j, \mathscr{X}_{k-1})P(\alpha^l|\beta^j, \mathscr{X}_{k-1})}.\end{aligned} \tag{19}$$

Remark 3: Let (i,j) represents the actual choice pair selected. For a properly designed algorithm, $P(\beta^j|\alpha^i,\mathscr{X}_k)$ will gradually dominate $P(\beta^m|\alpha^i,\mathscr{X}_k)$ for all $m\neq j$ and converge to 1 eventually after a number of observations. So is the case for $P(\alpha^i|\beta^j,\mathscr{X}_k)$ which will gradually dominate $P(\alpha^l|\beta^j,\mathscr{X}_k)$ for all $l\neq i$.

To start the above estimate and infer procedures, all the probabilities of an agent's choices can be assumed equal at the initial time, i.e., $P(\alpha^l|\beta^j,\hat{\mathbf{x}}_0)=1/N_A$, $\forall l\in\mathscr{A}$ and $P(\beta^m|\alpha^i,\hat{\mathbf{x}}_0)=1/N_B$, $\forall m\in\mathscr{B}$. Also, both agents should be informed of a common initial state $\mathbf{x}_0$.

IV. Controller Design

A. Optimal Open-loop Control

Denote p^i as the probability of Alice's choice α^i; similarly, denote q^j as the probability of Bob's choice β^j. In [11], the deterministic model containing no state or observation noise was investigated under the assumption that the choices of the agents are uniformly distributed, i.e. $p^i=1/N_A$ for all $i\in\mathscr{A}$, and $q^j=1/N_B$ for all $j\in\mathscr{B}$. Open-loop controls based on the initial state $\mathbf{x}_\kappa$ and a target matrix $\mathbf{H}$ that minimize a deterministic analogue of the cost functional (5) were derived. As an extension, for stochastic systems with an optimal state estimation $\hat{\mathbf{x}}_\kappa$ and arbitrary given choice distributions $\{p^i:i=1,2,\ldots,N_A\}$, $\{q^j:j=1,2,\ldots,N_B\}$, open-loop controls that minimize the cost functional

$$J^o(\kappa)=E\{\sum_{k=\kappa}^{N-1}[\sum_{i=1}^{N_A}p^i(\mathbf{u}_k^i)^T\mathbf{u}_k^i+\sum_{j=1}^{N_B}q^j(\mathbf{v}_k^j)^T\mathbf{v}_k^j]\}, \tag{20}$$

can be found following the separation principle (cf.[20]).

Proposition 1: Considering the system (1) with optimal initial state estimation $\hat{\mathbf{x}}_\kappa$, the optimal open-loop control inputs minimizing (20) subject to a compatible target matrix $\mathbf{H}$ are given as follows: for $i=1,2,\ldots,N_A$ and $j=1,2,\ldots,N_B$,

$$\begin{aligned}\hat{\mathbf{u}}_k^i=&\mathbf{B}^T(\mathbf{A}^T)^{N-1-k}(\mathbf{G}_B+\mathbf{G}_C)^{-1}\mathbf{d}_\kappa^{u,i},\\ \hat{\mathbf{v}}_k^j=&\mathbf{C}^T(\mathbf{A}^T)^{N-1-k}(\mathbf{G}_B+\mathbf{G}_C)^{-1}\mathbf{d}_\kappa^{v,j},\end{aligned}$$

where

$$\begin{aligned}\mathbf{d}_\kappa^{u,i}=&-\mathbf{A}^N\hat{\mathbf{x}}_\kappa+(\mathbf{I}+\mathbf{G}_C\mathbf{G}_B^{-1})\mathbf{H}_{i1}-\mathbf{H}_{11}\\&+\sum_{m=1}^{N_B}q^m\mathbf{H}_{1m}-\mathbf{G}_C\mathbf{G}_B^{-1}\sum_{l=1}^{N_A}p^l\mathbf{H}_{l1},\\ \mathbf{d}_\kappa^{v,j}=&-\mathbf{A}^N\hat{\mathbf{x}}_\kappa+(\mathbf{I}+\mathbf{G}_B\mathbf{G}_C^{-1})\mathbf{H}_{1j}-\mathbf{H}_{11}\\&+\sum_{l=1}^{N_A}p^l\mathbf{H}_{l1}-\mathbf{G}_B\mathbf{G}_C^{-1}\sum_{m=1}^{N_B}q^m\mathbf{H}_{1m},\end{aligned}$$

$$\mathbf{G}_B=\sum_{k=0}^{N-1-\kappa}\mathbf{A}^k\mathbf{B}\mathbf{B}^T(\mathbf{A}^T)^k,\ \mathbf{G}_C=\sum_{k=0}^{N-1-\kappa}\mathbf{A}^k\mathbf{C}\mathbf{C}^T(\mathbf{A}^T)^k.$$

The proof is similar to that for deterministic systems in [11] thus is omitted.

This open-loop control result provides the basic step of the new feedback control algorithm proposed in this paper, which aims to handle the uncertainties created by the state and observation noises.

B. Feedback Control

In conventional single-choice linear stochastic systems, feedback controllers can be designed from classical LQG optimal control theory [20]. Unfortunately, this method does not apply to systems with distributed choices. An alternate way is to use closed-loop forms of minimum energy control [21] by replacing $\hat{\mathbf{x}}_\kappa$ in the open-loop control law in Proposition 1 with $\hat{\mathbf{x}}_{k|k}$ for $k=\kappa,\ldots,N-1$. This yields a distributed feedback control law that relies only on the observed information $\mathbf{z}$. However, this method, which is referred to as the *estimate-feedback control* (EFC) approach, requires high control cost if the choice information is not dynamically updated as the system approaches the terminal time. This is because we have to solve essentially a series of optimization problems with the cost functionals, $J^o(0),J^o(1),\ldots,J^o(N-1)$ of the form (20) which are defined over shorter and shorter execution time.

To alleviate this undesirable situation, feedback controls need to contain information about the agents' selected choices, that is, we need to minimize a series of new cost functionals, $J(0),J(1),\ldots,J(N-1)$ of the form

$$J(\kappa)=E\{\sum_{k=\kappa}^{N-1}[\sum_{i=1}^{N_A}p_\kappa^i(\mathbf{u}_k^i)^T\mathbf{u}_k^i+\sum_{j=1}^{N_B}q_\kappa^j(\mathbf{v}_k^j)^T\mathbf{v}_k^j]\}, \tag{21}$$

where p_κ^i can be derived from Bob's estimate of Alice's choice α^i at time κ; similarly, q_κ^j is Alice's estimate of Bob's choice β^j. It is noticed that for our basic step as in Proposition 1 to work, both p_k^i and q_k^j should be known to both agents.

This requirement prevents the probabilities $P(\alpha^i|\beta^m,\mathscr{X}_k)$ and $P(\beta^j|\alpha^l,\mathscr{X}_k)$ derived in section III-B from being used directly in the controllers, since they are conditioned on one agent's private choice which is unknown to another agent. Instead, the agents need to compromise by using an averaged value, i.e., for all $i\in\mathscr{A}$, $j\in\mathscr{B}$, and $k=0$, they use $p_0^i=1/N_A$, and $q_0^j=1/N_B$, and for $k=1,\ldots,N-1$, they apply

$$p_k^i=\sum_{m=1}^{N_B}P(\alpha^i|\beta^m,\mathscr{X}_k)q_{k-1}^m, \tag{22}$$

$$q_k^j=\sum_{l=1}^{N_A}P(\beta^j|\alpha^l,\mathscr{X}_k)p_{k-1}^l. \tag{23}$$

Now, in consideration of Assumption 1 to Assumption 3, the feedback control law utilizing $\hat{\mathbf{x}}_{k|k}$ can be described by: For $i=1,2,\ldots,N_A$ and $j=1,2,\ldots,N_B$,

$$\hat{\mathbf{u}}_k^i=-\mathbf{K}_k^u(\mathbf{A}^{N-k}\hat{\mathbf{x}}_{k|k}-\mathbf{h}_k^{u,i}), \tag{24}$$

$$\hat{\mathbf{v}}_k^j=-\mathbf{K}_k^v(\mathbf{A}^{N-k}\hat{\mathbf{x}}_{k|k}-\mathbf{h}_k^{v,j}), \tag{25}$$

where $\hat{\mathbf{x}}_{k|k}$ is the UMV estimate of $\mathbf{x}_k$,

$$\mathbf{K}_k^u=\frac{1}{1+\gamma^2}\mathbf{B}^T(\mathbf{A}^T)^{N-1-k}\mathbf{G}_k^{-1}, \tag{26}$$

$$\mathbf{K}_k^v=\frac{1}{1+\gamma^2}\mathbf{C}^T(\mathbf{A}^T)^{N-1-k}\mathbf{G}_k^{-1}, \tag{27}$$

$$\mathbf{G}_k = \sum_{r=0}^{N-1-k} \mathbf{A}^r \mathbf{B}\mathbf{B}^T (\mathbf{A}^T)^r, \tag{28}$$

$$\mathbf{h}_k^{u,i} = (1+\gamma^2)\mathbf{H}_{i1} - \mathbf{H}_{11} + \sum_{m=1}^{N_B} q_k^m \mathbf{H}_{1m} - \gamma^2 \sum_{l=1}^{N_A} p_k^l \mathbf{H}_{l1}, \tag{29}$$

$$\mathbf{h}_k^{v,j} = (1+\frac{1}{\gamma^2})\mathbf{H}_{1j} - \mathbf{H}_{11} + \sum_{l=1}^{N_A} p_k^l \mathbf{H}_{l1} - \frac{1}{\gamma^2} \sum_{m=1}^{N_B} q_k^m \mathbf{H}_{1m}. \tag{30}$$

One can see that the state-estimation, choice probabilities and the control inputs are updated iteratively which engenders an *estimate-infer-feedback control* (EIFC) algorithm as summarized in Algorithm 1. This control approach provides more performance accuracy than the open-loop method, since state perturbation caused by the disturbance process ξ_k is eliminated in every step.

Algorithm 1 Estimate-Infer-Feedback Control Algorithm

Initialization: $\hat{\mathbf{x}}_0$: initial state estimate; $\mathbf{H}$: target matrix; $p_0^l = 1/N_A$, for all $l \in \mathscr{A}$: initial distribution of Alice's choice set; $q_0^m = 1/N_B$, for all $m \in \mathscr{B}$: initial distribution of Bob's choice set; $\alpha = i \in \mathscr{A}$: Alice's choice; $\beta = j \in \mathscr{B}$: Bob's choice; $k = 0$; $N \geq n \in \mathbb{Z}_+$: terminal time;

Steps:

1: compute $\hat{\mathbf{u}}_k^i$ and $\hat{\mathbf{v}}_k^j$ according to (24)-(30) and run the system (1);

2: $k = k+1$; if $k = N$, stop;

3: acquire $\mathbf{z}_k$, and compute $\hat{\mathbf{x}}_{k|k}$ according to (6)-(13);

4: compute $\{p_k^l : l \in \mathscr{A}\}$ and $\{q_k^m : m \in \mathscr{B}\}$ according to (16)-(19) and (22)-(23); go back to step 1;

Remark 4: The formulation of Alice's control law $\hat{\mathbf{u}}_k^i$ relies on Bob's estimates p_k^m, similarly this holds for Bob's control law $\hat{\mathbf{v}}_k^j$. In addition to their own estimate-infer procedures presented in section III, each agent needs to conduct additional estimate-infer procedures from the perspective of the other agent. This implies that the computation work of each agent has to be doubled, which may be considered as the price to pay for the lack of communication.

Remark 5: When Alice chooses i and Bob chooses j, it can be proved that the expected system state $\mathbf{x}_k$ converges to the target state $\mathbf{H}_{ij}$. Therefore, if the target states are all distinct, the probabilities p_k^i and q_k^j of the selected choice pair (i,j) will tend to 1 while p_k^m for all $m \neq i$ and q_k^l for all $l \neq j$ will approach 0, as $k \to N$. Then, the following control law follows directly from (24)-(30),

$$\hat{\mathbf{u}}_k^i = -\mathbf{K}_k^u(\mathbf{A}^{N-k}\hat{\mathbf{x}}_{k|k} - \mathbf{H}_{ij}), \hat{\mathbf{v}}_k^j = -\mathbf{K}_k^v(\mathbf{A}^{N-k}\hat{\mathbf{x}}_{k|k} - \mathbf{H}_{ij}).$$

From this one can deduce that the cost of the EIFCs, $\hat{\mathbf{u}}_k^i$ and $\hat{\mathbf{v}}_k^j$, proposed in this paper is lower than the EFC approach which does not update about the choice probabilities. Details will be provided elsewhere due to space limitation.

V. Simulation Results

Simulations are conducted for agents described by second-order integrators, where

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}, \mathbf{C} = -\mathbf{B}, \mathbf{F} = [1,0], \mathbf{Q} = \begin{bmatrix} 0.02 & 0 \\ 0 & 0.2 \end{bmatrix}$$

and $\mathbf{R} = 0.1$, $\mathbf{x}_0 = [0,1]^T$, $N = 10$. Then, $\mathbf{z}_k$ may be interpreted as the distance between two agents. Each agent has two possible choices indexed by $\{1,2\}$. The second coordinate of all target values are set to 0, while the projection of the first coordinate of the target matrix is given by

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} = \begin{bmatrix} 40 & 20 \\ -20 & -40 \end{bmatrix}.$$

It follows that the target matrix is compatible.

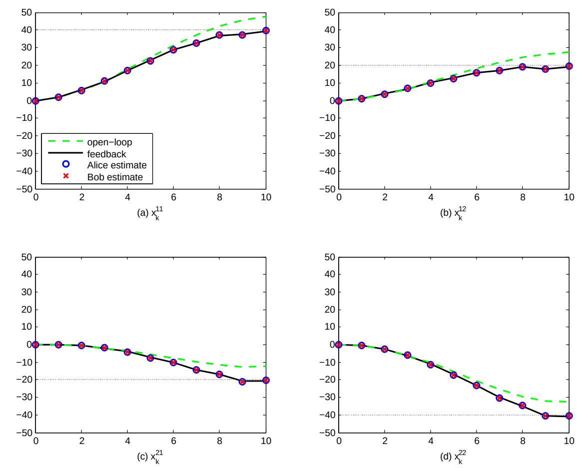

Fig. 1. State trajectories and state estimation for the first coordinate of $\mathbf{x}_k$. Green-dashed lines: open-loop control cannot achieve the targets; Black-solid lines: feedback control achieves the targets with high accuracy; Circles: state estimation by Alice; Cross points: state estimation by Bob.

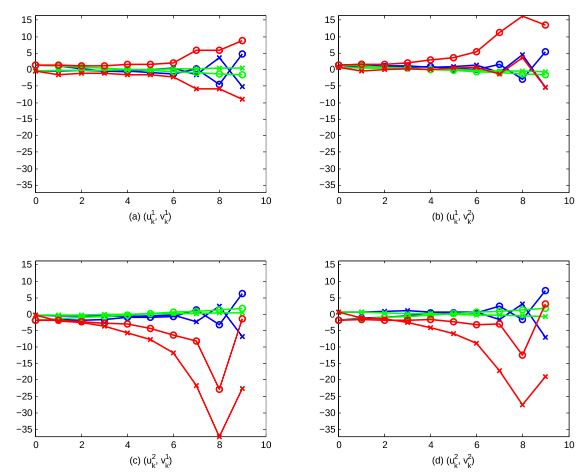

Fig. 2. Control inputs pairs (circle-lines for u_k^i, cross-lines for v_k^j). Green: open-loop control; Blue: EIFC; Red: EFC with fixed probability 0.5.

Simulation results for three types of control methods, open-loop control, EFC with fixed choice probability 0.5 and EIFC, are shown in Fig.1-Fig.2. Fig.1 shows the superiority of feedback control in the presence of disturbances and demonstrates the states estimated by both agents are equal. Fig.2 indicates that the EIFC approach with updated choice probabilities can substantially reduce control cost in contrast to feedback control laws with unchanged probabilities. Fig.3

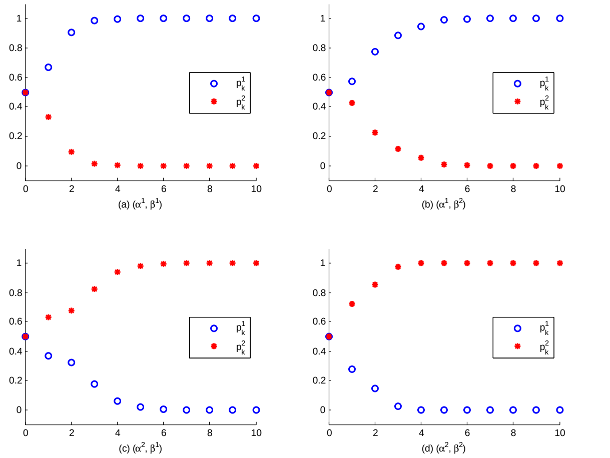

Fig. 3. Time-varying probability distributions $\{p_k^i, i = 1,2\}$ of Alice's choice set. The actual control input of Alice is: (a) u_k^1 (b) u_k^1 (c) u_k^2 (d) u_k^2.

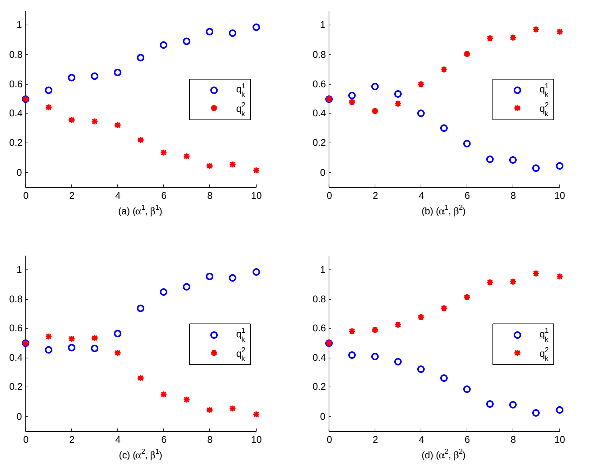

Fig. 4. Time-varying probability distributions $\{q_k^j, j = 1,2\}$ of Bob's choice set. The actual control input of Bob is: (a) v_k^1 (b) v_k^2 (c) v_k^1 (d) v_k^2.

and Fig.4 show that as more information are observed, the probability of an agent's selected choice judged by the other agent becomes dominant, so both agents gradually build up confidence on each other's choice.

VI. Conclusions

In this paper, we show how to address a cooperative control problem that allows agents to have distributed choice. Under the premise of no explicit communication channels between agents, distributed feedback controllers are devised by using state estimations from UMV estimators and applying inference mechanism to gain choice probabilities of the other party. This approach is shown to have lower energy consumption as the centralized controllers. But the proposed approach is sub-optimal due to the sub-optimality of the UMV estimator for our problem. So, an optimal estimator suitable for this type of new problems needs to be designed. For choice-based systems with more than two agents, controller design problems are more complicated and remains open. Analysis of control cost that can be saved by introducing explicit communication, in other words, the value of information [10] is another issue to be investigated in the future.

References

[1] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM J. Control Optim.* vol.6, no.1, pp. 131-147, 1968.

[2] Y. C. Ho, "Team decision theory and information structures," *Proc. IEEE*, vol. 68, no. 6, pp. 644-654, 1980.

[3] S. K. Mitter and A. Sahai, "Information and control: Witsenhausen revisited," in *Learning, Control and Hybrid Systems: Lecture Notes in Control and Information Sciences 241*, Y. Yamamoto and S. Hara, Eds. New York, NY: Springer, pp. 281-293, 1999.

[4] D. V. Raghunathan and J. Baillieul, "Exploiting information content in relative motion," in *Proc. of the American Control Conference*, St. Louis, pp. 2166-2171, June 10-12, 2009.

[5] A. Arsie, K. Savla, and E. Frazzoli, "Efficient routing algorithms for multiple vehicles with no explicit communications," *IEEE Trans. Automat. Contr.*, vol. 54, no. 19, pp. 2302-2317, 2009.

[6] P. Grover and A. Sahai, "Implicit and explicit communication in decentralized control," in *48th Annual Allerton Conference on Communication, Control, and Computing*, Illinois, USA, pp. 278-285, 2010.

[7] W. S. Wong, "Control communication complexity of distributed control systems," *SIAM J. Control Optim.*, vol. 48, no. 3, pp. 1722-1742, 2009.

[8] W. S. Wong and J. Baillieul, "Control communication complexity of nonlinear systems," *Communications in Information and Systems*, vol. 9, no. 1, pp. 103-140, 2009.

[9] J. Baillieul and W. S. Wong, "The standard part problem and the complexity of control communication," in *Proc. IEEE Conference on Decision and Control*, Shanghai, pp.2723-2728, December 16-18, 2009.

[10] W. S. Wong and J. Baillieul, "Control communication complexity of distributed actions," *IEEE Trans. Automat. Contr.*, vol. 57, no. 11, pp. 2731-2745, 2012.

[11] G. Guo, W. S. Wong and Z. C. Liu, "Cooperative target realization in multi-agent systems allowing choice-based actions," *preprint, arXiv: 1207.0097*, 2012.

[12] R. W. Brockett, "Minimizing attention in a motion control context," in *Proc. IEEE Conference on Decision and Control*, pp. 3349-3352, 2003.

[13] S. Yüksel, "Decentralized computation and communication in stabilization of distributed control systems," *Information Theory and Applications (ITA)*, Feb. 2009.

[14] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: algorithms and theory," *IEEE Trans. Automat. Contr.*, vol. 51, no. 3, pp. 401-420, 2006.

[15] H. Kim, H. Shim, J. H. Seo, "Output consensus of heterogeneous uncertain linear multi-agent systems," *IEEE Trans. Automat. Contr.*, vol. 56, no.1, pp. 200-206, 2011.

[16] J. A. Fax, and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1465-1476, 2004.

[17] J. Sijs, M. Lazar P. P. J. van den Bosch, Z. Papp, "An overview of non-centralized Kalman filters." in *IEEE Multi-conference on Systems and Control*, San Antonio, USA, pp. 739-744, September 3-5, 2008.

[18] P. K. Kitanidis, "Unbiased minimum-variance linear state estimation," *Automatica*, vol. 23, pp. 775-778, 1987.

[19] S. Gillijns, B. D. Moor, "Unbiased minimum-variance input and state estimation for linear discrete-time systems," *Automatica*, vol. 43, pp. 111-116, 2007.

[20] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, New York: Wiley-Interscience, 1972.

[21] E. Kondo, T. Kiyota, and T. Sunaga, "Closed-loop expressions of minimum energy control for single-input linear digital systems," *IEEE Trans. Automat. Contr.*, vol. 38, no. 4, pp. 603-606, 1993.