2016 IEEE 55th Conference on Decision and Control (CDC)
ARIA Resort & Casino
December 12-14, 2016, Las Vegas, USA

Strategic Control of a Tracking System

Muhammed O. Sayin, Emrah Akyol, Tamer Başar

***Abstract*— We consider stochastic dynamic game problems where a trajectory controller takes an action to construct an information bearing signal, namely the control input, and subsequently a tracking system takes an action, i.e., constructs a tracking output, based on the control input. The trajectory controller has access to two Gaussian processes evolving according to first-order autoregressive models, e.g., desired and private states. Different from the design of a measurement or sensing scheme for a tracking system, here the trajectory controller and the tracker have different objectives. Particularly, the trajectory controller aims to drive the tracking system to a desired path, different from the tracker's actual intent, by constructing the measurement signal. For finite horizon problems involving two different quadratic cost functions, we show that the optimal control input policies are linear functions of the current states when the states evolve in parallel. We then extend this result for the general case when the trajectory controller has a myopic objective and show that the optimal control input policies are also linear functions of the current states. Finally, we restrict the policy space for the control input to the set of all linear mappings of the current states and convert the finite horizon stochastic game problem into a discrete time deterministic optimal control problem. We also include some illustrative numerical examples for different strategic control scenarios.**

I. Introduction

Along with the increased efficiency in the data transmission and the enhanced processing performance of sensors, there is an increasing demand for simultaneous design of measurement, estimation, and control architectures in decentralized systems [1]–[5]. In such problems, at each instant, decisions, e.g., estimation or control, are made based on the received information from the measurement devices. Due to the improved processing capabilities of these devices, the measurements can also be processed with respect to the objective of the decision maker rather than a direct transmission, which improves the performance in decentralized systems [4]. Additionally, due to time delay and processing power related concerns, low complexity, such as linear, strategies play an important role in simultaneous measurement and decision making. As an example, for decentralized control with feedback, reference [1] has studied a stochastic dynamic decision problem in which both the transmitted information from the measurement device and the consecutive decision made by the controller are designed together. For stochastic control problems with Gaussian states evolving according to first order autoregressive models, it was shown that the optimal measurement strategy is a linear function of the innovation in the measurement, i.e., new information given the previous measurements. Correspondingly, due to the linear measurement policy, the optimal control policy is a linear function of the measurement signal. Continuous-time version of this joint design problem was also addressed, in [6]. In [5], the authors study real-time tracking of Gaussian processes that evolve according to first order autoregressive models over an additive white Gaussian noise channel, without feedback, and also show the optimality of the innovation encoder. However, in all these studies, the measurement and the decision devices cooperate to achieve a common goal.

Departing from the paradigm of cooperative design, we consider in this paper a strategic environment where the information provider and the decision maker have different objectives, and therefore take actions in a non-cooperative manner. Originally, such a scheme, namely strategic information transmission, has been introduced in a seminal Econometrica paper by V. Crawford and J. Sobel [7], and attracted significant attention in the economics literature due to the wide range applications from advertising to expert advise sharing problems [8]–[10]. Here, the transmitter has access to a source output and a private bias information, and sends a message to the receiver to maximize a utility function depending on the receiver's action and the information regarding the source output and private bias. On the other side, the receiver takes an action based on the transmitter's message in order to maximize a different utility function that is independent of the private information. Under Nash equilibrium, in which the players announce their strategies simultaneously, the authors have shown that a quantization-based mapping of the source and the private information achieves the equilibrium [7]. In [11], the authors extend the one-shot game of strategic information transmission to a multi-stage one with a finite horizon, such that the information provider and decision makers interact several times regarding a constant unknown state of the world. Recently, references [12], [13] have studied strategic information transmission strategies under a different equilibrium concept, the Stackelberg equilibrium, where there is a hierarchy in the announcement of the strategies, and have shown that for the quadratic-Gaussian case, equilibrium achieving strategies are linear functions of the source output and the private information, in contrast to the quantized schemes that emerge under Nash equilibrium.

In this paper, we consider a strategic environment in which the measurement device, i.e., a trajectory controller, and the decision device, i.e., a tracking system, have dif-

This research was supported in part by the U.S. Air Force Office of Scientific Research (AFOSR) MURI grant FA9550-10-1-0573, in part by the Office of Naval Research (ONR) MURI grant N00014-16-1-2710, and in part by NSF under grant CCF 11-11342.

M. O. Sayin, E. Akyol, and T. Başar are with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL. 61801; emails:{sayin2, akyol, basar1}@illinois.edu

978-1-5090-1837-6/16/$31.00 ©2016 IEEE

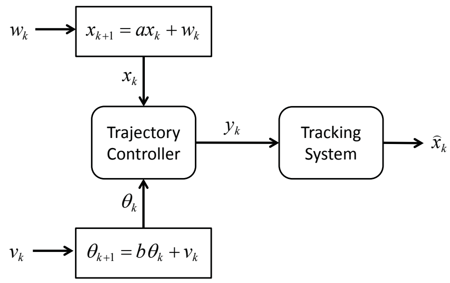

Fig. 1: The trajectory controller and the tracking system at time instant k.

ferent objectives, i.e., they do not cooperate. As seen in Fig. 1, the trajectory controller has access to two different Gaussian processes, i.e., desired and private states, that evolve according to first order autoregressive models. And the tracking system tracks the desired state through the information provided by the trajectory controller. However, the controller has a different objective, such that the output of the tracking system tracks the sum of the desired and the private states, and hence it constructs the transmitted information, i.e., the control input, accordingly. We assume that the tracking system has complete knowledge about the control input as in the general case [14]. We further assume that there is a hierarchy between the trajectory controller and the tracking system such that the trajectory controller announces the policies in the construction of the control input beforehand. The difference between the objectives and the hierarchy between the systems correspond to a dynamic Stackelberg game between the trajectory controller and the tracking system. We point out that different from [11], here not only the players interact with each other several times but also the underlying states evolve in time, i.e., they are not constants.

For finite horizon problems with two different quadratic cost functions, we show that the optimal control policies are linear functions of the current states when the states evolve in parallel and correspondingly, the output of the tracking system is a linear function of the received control inputs. For the general case, we prove that the optimal control input policies are also linear functions of the current states if the trajectory controller has a myopic objective rather than a finite horizon goal. Finally, restricting the policy space for the control input into the set of all linear mappings of the current states, we show that the dynamic game problem can be considered as a discrete time optimal control problem.

We can list the main features of this paper as follows: (1) We study a dynamic Stackelberg game of strategic information transmission where the players interact several times regarding unknown, *time-variant* states of the world. (2) We show that when the states evolve in parallel, linear policies can lead to a Stackelberg equilibrium with a *finite horizon* as in the one-shot case [12], [13]. (3) For myopic objectives, we show that optimal control input policies are linear functions of the current states within the *general* class of the policies, i.e., without an a priori linearity restriction as in [15].

We organize the paper as follows. In Section II, we describe the problem. In Section III, we analyze the optimality of linear policies. We formulate the myopically optimal policies in Section IV. In Section V, we analyze the optimal linear policies for the finite horizon problem. In Section VI, we provide numerical examples for different strategic control scenarios. We conclude the paper in Section VII.

II. Problem Formulation

Consider two discrete-time, scalar, stationary processes $\{x_k\}$, the desired state, and $\{\theta_k\}$, the private state, evolving according to the following autoregressive model

$$\begin{bmatrix} x_{k+1} \\ \theta_{k+1} \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} x_k \\ \theta_k \end{bmatrix} + \begin{bmatrix} w_k \\ v_k \end{bmatrix},\ k = 1, 2, \cdots, \tag{1}$$

where $|a| < 1$, $|b| < 1$, $x_1 \sim \mathcal{N}(0, \sigma_x^2)$, $\theta_1 \sim \mathcal{N}(0, \sigma_\theta^2)$, and[1] $x_1 \perp\!\!\!\perp \theta_1$. The noise processes $\{w_k\}, \{v_k\}$ are white Gaussian such that $w_k \sim \mathcal{N}(0, \sigma_w^2)$, $v_k \sim \mathcal{N}(0, \sigma_v^2)$ are independent of each other, and of the current states x_k and θ_k.

We consider a strategic environment where the trajectory controller and the tracking system have different objectives. As seen in Fig. 1, at each instant k, the trajectory controller has access to the desired and the private states, i.e., $x_{[1,k]} := x_1, \cdots, x_k$ and $\theta_{[1,k]} := \theta_1, \cdots, \theta_k$, and constructs $y_k \in \mathbb{R}$, namely the control input, which is fully determined by the conditional distribution $p(\cdot | x_{[1,k]}, \theta_{[1,k]})$ as

$$\mathbb{P}(y_k \in \mathscr{Y}) = \int_{y \in \mathscr{Y}} p(y | x_{[1,k]}, \theta_{[1,k]}) dy\ \forall \mathscr{Y} \subseteq \mathbb{R}$$

holds almost everywhere in $x_{[1,k]}$ and $\theta_{[1,k]}$. With an abuse of notation, we denote this function by $y_k = y_k(x_{[1,k]}, \theta_{[1,k]})$ and the set of all such Lebesgue-measurable functions from $\mathbb{R}^k \times \mathbb{R}^k$ to $\mathbb{R}$ by Γ_k. Correspondingly, the tracking system receives the control input y_k and has a memory such that the output of the tracking system $\hat{x}_k$ is fully determined by the conditional distribution $p(\cdot | y_{[1,k]})$ as

$$\mathbb{P}(\hat{x}_k \in \mathscr{X}) = \int_{\hat{x} \in \mathscr{X}} p(\hat{x} | y_{[1,k]}) d\hat{x}\ \forall \mathscr{X} \subseteq \mathbb{R}$$

holds almost everywhere in $y_{[1,k]}$. We denote this function by $\hat{x}_k = \hat{x}_k(y_{[1,k]})$ and the set of all such Lebesgue-measurable functions from $\mathbb{R}^k$ to $\mathbb{R}$ by Ω_k.

The tracking system aims to track the desired state x_k through a finite horizon objective

$$\min_{\substack{\hat{x}_k(\cdot) \in \Omega_k, \\ \forall k \in \{1, \cdots, n\}}} \sum_{k=1}^{n} \mathbb{E}\left\{ \left(x_k - \hat{x}_k\left(y_{[1,k]}\right) \right)^2 \right\}. \tag{2}$$

whereas the trajectory controller aims $\hat{x}_k$ to track $x_k + \theta_k$ through

$$\min_{\substack{y_k(\cdot,\cdot) \in \Gamma_k, \\ \forall k \in \{1, \cdots, n\}}} \sum_{k=1}^{n} \mathbb{E}\left\{ \left(x_k + \theta_k - \hat{x}_k\left(y_{[1,k]}\right) \right)^2 \right\}. \tag{3}$$

[1]As notation, $\perp\!\!\!\perp$ denotes independence and $\mathcal{N}(0, \cdot)$ denotes the Gaussian distribution with zero mean and designated variance.

Note that even though the tracking system has a finite horizon objective, as seen in Fig. 1, the scheme involves no feedback channel from the tracking system to the trajectory controller. Hence, the tracking outputs have no impact on the forthcoming control inputs, which yields that we can consider the finite horizon objective (2) as a myopic objective

$$\min_{\hat{x}_k(\cdot)\in\Omega_k} \mathbb{E}\left\{(x_k-\hat{x}_k(y_{[1,k]}))^2\right\}. \tag{4}$$

On the contrary, while constructing the control inputs, the trajectory controller should consider the impact of y_k in the forthcoming steps. Here, we assume that the systems are aware of the differences between their objectives (3) and (4), and all the statistics are common knowledge. Since the objectives (3) and (4) differ, this scheme corresponds to a game problem between the players: the trajectory controller and the tracking system. Mainly, there are two approaches to such game problems [16]. In the Nash equilibrium, the players simultaneously take actions. Particularly, the trajectory controller chooses a policy for the control input y_k from the policy space Γ_k and the tracking system chooses a policy for the output $\hat{x}_k$ from the policy space Ω_k. Once the policies are announced, the players cannot change their actions. Hence, the Nash equilibrium has the players adopt their policies such that they would have no incentive to change their actions unilaterally. In the Stackelberg equilibrium, there is a hierarchy between the players such that one of the players, namely the leading player, takes action beforehand knowing that the other player will take the best response against his/her action [17].

In this paper, we focus on the Stackelberg equilibrium in which the trajectory controller leads the game by announcing the policy in the construction of the control inputs. We point out that given the transmitted control inputs, the reaction set of the tracker is a singleton. Hence, for any given control inputs $y_1,\cdots,y_n$, the best reaction of the tracker is given by

$$\hat{x}_k^*(y_{[1,k]}) = \arg\min_{\hat{x}(\cdot)\in\Omega_k} \mathbb{E}\{(x_k-\hat{x})^2\}, \tag{5}$$

for each $k=1,2,\cdots,n$. Correspondingly, the optimal control inputs $y_1^*,\cdots,y_n^*$, i.e., the policies yielding the Stackelberg equilibrium (5), are constructed according to the best reaction of the tracker (5). Further, in the construction of the optimal control inputs, the controller faces the following minimization problem with finite horizon:

$$\min_{\substack{y_k(\cdot,\cdot)\in\Gamma_k,\\ \forall k\in\{1,\cdots,n\}}} \sum_{k=1}^{n} \mathbb{E}\left\{\left(x_k+\theta_k-\hat{x}_k^*(y_{[1,k]})\right)^2\right\}. \tag{6}$$

We also study the scenarios when the trajectory controller has a myopic objective, which results in the following minimization problem for the controller:

$$y_k^* = \arg\min_{y_k(\cdot,\cdot)\in\Gamma_k} \mathbb{E}\left\{(x_k+\theta_k-\hat{x}_k^*(y_{[1,k]}))^2\right\}. \tag{7}$$

In the myopic scenario (7), the controller takes action by considering only the current stage irrespective of the actions' impact on future stages.

In the next section, we discuss the scenarios in which linear control strategies achieve the equilibrium with respect to (5) and (6).

III. Optimality of the Linear Control for Finite Horizon Costs

We point out that for $n=1$, the controller has access to a single desired source x and a single private source θ. The tracking system, or the estimator, since the states do not evolve in time, aims to minimize

$$\mathbb{E}\{(x-\hat{x}(y))^2\}, \tag{8}$$

through $\hat{x}(y)$ over all Lebesgue measurable functions from $\mathbb{R}$ to $\mathbb{R}$. In particular, $\hat{x}(y)$ denotes the estimate obtained through the control input y while the controller constructs that information such that

$$\mathbb{E}\{(x+\theta-\hat{x}(y(x,\theta)))^2\} \tag{9}$$

is minimized by $y(x,\theta)$ over all Lebesgue measurable functions from $\mathbb{R}\times\mathbb{R}$ to $\mathbb{R}$. This corresponds to a strategic version of the information transmission scenario [7], where the trajectory controller is the transmitter, the estimator is the receiver, and the control input y is the transmitted message. The transmitter and the receiver have different distortion measures (9) and (8), respectively, rather than a common goal, e.g., (8) solely, as in the conventional communication models. Recently, in [12], [13], for the quadratic-Gaussian case, the authors have shown that the Stackelberg equilibrium can be achieved through a linear mapping of the source and the private information, and the following theorem from [12] provides the corresponding optimal policies.

Theorem 1. *For the Stackelberg equilibrium in a strategic environment, where* $\left[\begin{smallmatrix} x\\ \theta\end{smallmatrix}\right] \sim \mathcal{N}\left(\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right], \left[\begin{smallmatrix}\sigma_x^2 & \rho\\ \rho & \sigma_{\theta^2}\end{smallmatrix}\right]\right)$ *and the players have objectives* (8) *and* (9)*, the optimal policy of the leader is given by* $y(x,\theta)=x+\alpha\theta$*, where*

$$\alpha = \frac{-\sigma_x^2+\sigma_x\sqrt{\sigma_x^2+4(\sigma_\theta^2+\rho)}}{2(\sigma_\theta^2+\rho)} \tag{10}$$

and that of the follower is $\hat{x}(y)=\frac{\sigma_x^2+\alpha\rho}{\sigma_x^2+\sigma_\theta^2+2\alpha\rho}y$*.*

We use Thm. 1 to formulate optimal linear policies for the Stackelberg game (5) and (6). The following theorem shows that there exists a linear policy for the control input that achieves the equilibrium, i.e., optimal with respect to (6), if the desired and the private states evolve in parallel.

Theorem 2. *For* $a=b$ *in* (1)*, the optimal control policies with respect to* (6) *are linear mappings of the current desired state* x_k *and the private state* θ_k *as* $y_k^*=x_k+\alpha\theta_k$*, where* α *is time invariant and defined in* (10)*.*

Proof. Note that by (5), the best response of the tracking system is given by

$$\hat{x}_k = \arg\min_{\hat{x}} \mathbb{E}\{(x_k-\hat{x})^2|y_{[1,k]}\} = \mathbb{E}\{x_k|y_{[1,k]}\}.$$

Hence the trajectory controller aims to solve

$$\min_{\substack{y_k(\cdot,\cdot)\in\Gamma_k,\\ \forall k\in\{1,\cdots,n\}}} \sum_{k=1}^{n} \mathbb{E}\left\{\left(x_k+\theta_k-\mathbb{E}\{x_k|y_{[1,k]}\}\right)^2\right\}.$$

Here, we consider the case where $a = b$. Assume that the first control input is given by $y_1(x_1, \theta_1) = x_1 + \alpha\theta_1$, where α is defined in (10). At the next stage, the controller aims to solve the following myopic minimization problem:

$$\min_{y(\cdot,\cdot)\in\Gamma_2} \mathbb{E}\left\{(x_2 + \theta_2 - \mathbb{E}\{x_2|y_1, y\})^2\right\}, \tag{11}$$

which lower bounded by the global minimum at $k = 2$ as

$$\min_{y'(\cdot,\cdot)\in\Gamma_{x_2,\theta_2}} \mathbb{E}\left\{(x_2 + \theta_2 - \mathbb{E}\{x_2|y'\})^2\right\}, \tag{12}$$

where Γ_{x_2,θ_2} denotes the set of all Lebesgue measurable mappings of x_2 and θ_2 to $\mathbb{R}$, and (12) bounds (11) from below since the previous control inputs restrict the performance of the trajectory controller. Thm. 1 implies that $y'(x_2, \theta_2) = x_2 + \alpha\theta_2$ minimizes (12).

Next, consider $\mathbb{E}\{x_2|y_1, y'\}$, which can be written as

$$\begin{aligned}\mathbb{E}\{x_2|y_1, y'\} &= \mathbb{E}\{x_2 - \mathbb{E}\{x_2|y'\}|y_1, y'\} + \mathbb{E}\{x_2|y'\}\\ &= \mathbb{E}\{x_2 - \mathbb{E}\{x_2|y'\}|y_1 - \mathbb{E}\{y_1|y'\}, y'\} + \mathbb{E}\{x_2|y'\},\end{aligned}$$

where y' is uncorrelated with $x_2 - \mathbb{E}\{x_2|y'\}$ and $y_1 - \mathbb{E}\{y_1|y'\}$. Since they are all Gaussian, $y' \perp\!\!\!\perp x_2 - \mathbb{E}\{x_2|y'\}$ and $y' \perp\!\!\!\perp y_1 - \mathbb{E}\{y_1|y'\}$. Then, we obtain

$$\mathbb{E}\{x_2|y_1, y'\} = \mathbb{E}\{x_2 - \mathbb{E}\{x_2|y'\}|y_1 - \mathbb{E}\{y_1|y'\}\} + \mathbb{E}\{x_2|y'\}.$$

Note that we have

$$\begin{aligned}&\mathbb{E}\{(x_2 - \mathbb{E}\{x_2|y'\})(y_1 - \mathbb{E}\{y_1|y'\})\}\\ &\quad = \mathbb{E}\{(x_2 - \mathbb{E}\{x_2|y'\})y_1\} - \mathbb{E}\{(x_2 - \mathbb{E}\{x_2|y'\})\mathbb{E}\{y_1|y'\}\},\end{aligned}$$

where the second term on the right-hand-side is 0 since $y' \perp\!\!\!\perp x_2 - \mathbb{E}\{x_2|y'\}$. Additionally, we have

$$\begin{aligned}\mathbb{E}\{(x_2 - \mathbb{E}\{x_2|y'\})y_1\} &= \mathbb{E}\left\{\left(x_2 - \frac{\sigma_x^2}{\sigma_x^2 + \alpha^2\sigma_\theta^2}y'\right)y_1\right\},\\ &= \mathbb{E}\left\{\frac{a\alpha^2\sigma_\theta^2 x_1 - a\alpha\sigma_x^2\theta_1}{\sigma_x^2 + \alpha^2\sigma_\theta^2}(x_1 + \alpha\theta_1)\right\} = 0,\end{aligned} \tag{13}$$

where the parallel evolution of the states plays a significant role on the uncorrelatedness of $x_2 - \mathbb{E}\{x_2|y'\}$ and y_1. Hence, we have $\mathbb{E}\{x_2 - \mathbb{E}\{x_2|y'\}|y_1 - \mathbb{E}\{y_1|y'\}\} = 0$. Eventually we obtain $\mathbb{E}\{x_2|y_1, y'\} = \mathbb{E}\{x_2|y'\}$ and $y(x_2, \theta_2) = x_2 + \alpha\theta_2$ in (11) minimizes the lower bound (12). Hence, $y_2(x_2, \theta_2) = x_2 + \alpha\theta_2$ is not only myopically optimal but also globally optimal. Following identical steps, we can also show that if $y_1, \cdots, y_{m-1}$ are globally optimal control inputs then $y_m(x_m, \theta_m) = x_m + \alpha\theta_m$ is also globally optimal. By induction, we then conclude that $\forall k \in \{1, \cdots, n\}$, $y_k(x_k, \theta_k) = x_k + \alpha\theta_k$ is the globally optimal control input.

Note that the minimum value of the finite horizon cost (6) is bounded by the myopic minimum and the sum of the global minima as

$$\begin{aligned}&\sum_{k=1}^{n} \min_{y_k(\cdot,\cdot)\in\Gamma_k} \mathbb{E}\{(x_k + \theta_k - \mathbb{E}\{x_k|y_{[1,k]}\})^2\}\\ &\quad \geq \min_{\substack{y_k(\cdot,\cdot)\in\Gamma_k,\\ \forall k\in\{1,\cdots,n\}}} \sum_{k=1}^{n} \mathbb{E}\left\{\left(x_k + \theta_k - \mathbb{E}\{x_k|y_{[1,k]}\}\right)^2\right\}\\ &\quad \geq \sum_{k=1}^{n} \min_{y'(\cdot,\cdot)\in\Gamma_{x_k,\theta_k}} \mathbb{E}\{(x_k + \theta_k - \mathbb{E}\{x_k|y'\})^2\},\end{aligned}$$

yet the myopic minimum achieves the sum of the global minima and the corresponding optimal control input is $y_k = x_k + \alpha\theta_k$. Then y_k is also optimal with respect to (6) and the proof is concluded. $\square$

Remark 1. *By Thm. 2, the optimal control policies are* $y_{k+1}(x_{k+1}, \theta_{k+1}) = x_{k+1} + \alpha\theta_{k+1} = a(x_k + \alpha\theta_k) + w_k + \alpha v_k$. *Particularly, at time* $k+1$, *the estimate can be written as* $\mathbb{E}\{x_{k+1}|y_{[1,k+1]}\} = \mathbb{E}\{x_{k+1}|x_1 + \alpha\theta_1, w_1 + \alpha v_1, ..., w_k + \alpha v_k\}$. *We note that since* x_k *and* θ_k *are stationary processes,* $a = b$ *yields* $\sigma_w^2 = \frac{\sigma_x^2}{\sigma_\theta^2}\sigma_v^2$. *Consequently,* $w_k + \alpha v_k$ *is obtained as*

$$w_k + \alpha v_k = \arg \min_{w'(\cdot,\cdot)\in\Gamma_{w_k,v_k}} \mathbb{E}\{(w_k + v_k - \mathbb{E}\{w_k|w\})^2\}.$$

This implies that for $a = b$, *the construction of the optimal controls can be considered as the strategic transmission of independent and identically distributed desired and private source sequences. Particularly, in a strategic environment the whitening of the processes that are correlated in time is possible if the desired and the private states evolve in parallel.*

IV. Myopically Optimal Control

Here, we consider the scenario where the trajectory controller has a myopic objective (7) and show that the optimal control input policies provided later in Thm. 3 are linear functions of the current states. Thm. 1 yields that at $k = 1$ the myopically optimal control input is a linear function of the desired state x_1 and the private state θ_1. At $k = 2$, since all the parameters in (7) are jointly Gaussian due to $y_1 = x_1 + \alpha\theta_1$, the myopic objective, which is minimized with respect to $y_2(\cdot,\cdot)$, is $\mathbb{E}\{(x_2 + \theta_2 - \mathbb{E}\{x_2|y_1, y_2\})^2\} = \mathbb{E}\{(\tilde{x}_2 + \theta_2 - \mathbb{E}\{\tilde{x}_2|\tilde{y}_2\})^2\}$, where $\tilde{x}_2 = x_2 - \mathbb{E}\{x_2|y_1\}$ and $\tilde{y}_2 = y_2 - \mathbb{E}\{y_2|\mathbb{E}\{x_2|y_1\}\}$. Note that $\tilde{x}_2 \perp\!\!\!\perp \mathbb{E}\{x_2|y_1\}$ and $\tilde{y}_2 \perp\!\!\!\perp \mathbb{E}\{x_2|y_1\}$, yet θ_2 and $\mathbb{E}\{x_2|y_1\}$ are not independent. However, we can write $\mathbb{E}\{(x_2 + \theta_2 - \mathbb{E}\{x_2|y_1, y_2\})^2\} = \mathbb{E}\{(\tilde{x}_2 + \tilde{\theta}_2 - \mathbb{E}\{\tilde{x}_2|\tilde{y}_2\})^2\} + \mathbb{E}\{(\mathbb{E}\{\theta_2|\mathbb{E}\{x_2|y_1\}\})^2\}$, where $\tilde{\theta}_2 = \theta_2 - \mathbb{E}\{\theta_2|\mathbb{E}\{x_2|y_1\}\}$ and the second term on the right-hand-side does not depend on the minimization argument $y_2(\cdot,\cdot)$. Then the minimization objective can be written as

$$\min_{y_2(\cdot,\cdot)\in\Gamma_2} \mathbb{E}\{(\tilde{x}_2 + \tilde{\theta}_2 - \mathbb{E}\{\tilde{x}_2|y_2\})^2\} \tag{14}$$

subject to $y_2 \perp\!\!\!\perp \mathbb{E}\{x_2|y_1\}$, and the following theorem shows that the policy space Γ_2 can be narrowed.

Blackwell's Irrelevant Information Theorem [18]–[20]. *Consider standard Borel spaces [21]* $\mathbb{X}$, $\mathbb{Y}$, *and* $\mathbb{U}$ *such that there is a probability measure* P *on* $\mathbb{X}\times\mathbb{Y}$, *and let* $c : \mathbb{X}\times\mathbb{U} \to \mathbb{R}$ *be a bounded Borel-measurable cost function. Then, for any Borel measurable function* $\gamma : \mathbb{X}\times\mathbb{Y} \to \mathbb{U}$, *i.e., any policy based on both* $x \in \mathbb{X}$ *and* $y \in \mathbb{Y}$, *there exists another Borel measurable function* $\gamma^\circ : \mathbb{X} \to \mathbb{U}$, *i.e., a policy depending on x only, such that*

$$\int_{\mathbb{X}} c(x, \gamma^\circ(x))P(dx) \leq \int_{\mathbb{X}\times\mathbb{Y}} c(x, \gamma(x, y))P(dx, dy).$$

We invoke Blackwell's Theorem in the following:

$$\min_{y_2(\cdot,\cdot)\in\Gamma_2} \mathbb{E}\{(\tilde{x}_2+\tilde{\theta}_2-\mathbb{E}\{\tilde{x}_2|y_2\})^2\} \quad (15)$$

$$\geq \min_{y_2(\cdot,\cdot)\in\Gamma_{\tilde{x}_2,\tilde{\theta}_2}} \mathbb{E}\{(\tilde{x}_2+\tilde{\theta}_2-\mathbb{E}\{\tilde{x}_2|y_2\})^2\}, \quad (16)$$

where $\Gamma_{\tilde{x}_2,\tilde{\theta}_2}$ denotes the set of all Lebesgue measurable mappings of $\tilde{x}_2$ and $\tilde{\theta}_2$ since the bounded Borel measurable cost function $\mathbb{E}\{(\tilde{x}_2+\tilde{\theta}_2-\mathbb{E}\{\tilde{x}_2|y_2\})^2\}$ in (15) depends on only $\tilde{x}_2$, $\tilde{\theta}_2$ and the policy y_2. Note that since the independence constraint restricts the policy space for $y_2(\cdot,\cdot)$, (16) bounds (14) from below and $\tilde{y}(\cdot,\cdot)\in\Gamma_{\tilde{x}_2,\tilde{\theta}_2}$ already implies independence from $\mathbb{E}\{x_2|y_1\}$ since $\tilde{x}_2 \perp\!\!\!\perp \mathbb{E}\{x_2|y_1\}$ and $\tilde{\theta}_2 \perp\!\!\!\perp \mathbb{E}\{x_2|y_1\}$. Hence, the myopic objective at $k=2$ can be written as (16) and Thm. 1 says that the corresponding optimal control input is also a linear function and given by $y_2(\tilde{x}_2,\tilde{\theta}_2)=\tilde{x}_2+\alpha_2\tilde{\theta}_2$, where

$$\alpha_2=\frac{-\tilde{\sigma}_{x,2}^2+\tilde{\sigma}_{x,2}\sqrt{\tilde{\sigma}_{x,2}^2+4(\tilde{\sigma}_{\theta,2}^2+\tilde{\rho}_2)}}{2(\tilde{\sigma}_{\theta,2}^2+\tilde{\rho}_2)}, \quad (17)$$

$$\tilde{\sigma}_{x,2}^2=\mathbb{E}\{\tilde{x}_2^2\},\tilde{\sigma}_{\theta,2}^2=\mathbb{E}\{\tilde{\theta}_2^2\},\tilde{\rho}_2=\mathbb{E}\{\tilde{x}_2\tilde{\theta}_2\}.$$

Suppose that until $k=m-1$, the optimal control inputs in terms of the myopic objective (7) are all linear-in-parameters. This yields that $y_{[1,m-1]}$ are jointly Gaussian with the current states x_m and θ_m. Then, we obtain

$$\mathbb{E}\{(x_m+\theta_m-\mathbb{E}\{x_m|y_{[1,m]}\})^2\}=\mathbb{E}\{(\tilde{x}_m+\theta_m-\mathbb{E}\{\tilde{x}_m|\tilde{y}_m\})^2\},$$

where $\tilde{x}_m = x_m - \mathbb{E}\{x_m|y_{[1,m-1]}\}$ and $\tilde{y}_m = y_m - \mathbb{E}\{y_m|\mathbb{E}\{x_m|y_{[1,m-1]}\}\}$. Since $\tilde{x}_m-\mathbb{E}\{\tilde{x}_m|\tilde{y}_m\}$ is uncorrelated with $\mathbb{E}\{x_m|y_1,\cdots,y_{m-1}\}$, we have $\tilde{x}_m-\mathbb{E}\{\tilde{x}_m|\tilde{y}_m\} \perp\!\!\!\perp \mathbb{E}\{\theta_m|\mathbb{E}\{x_m|y_{[1,m-1]}\}\}$ and the myopic objective can be written as

$$\min_{y(\cdot,\cdot)\in\Gamma_{\tilde{x}_m,\tilde{\theta}_m}} \mathbb{E}\{(\tilde{x}_m+\tilde{\theta}_m-\mathbb{E}\{\tilde{x}_m|\tilde{y}\})^2\},$$

where $\tilde{\theta}_m=\theta_m-\mathbb{E}\{\theta_m|\mathbb{E}\{x_m|y_{[1,m-1]}\}\}$. Hence the optimal control at $k=m$ is also linear-in-parameters $\tilde{x}_m$ and $\tilde{\theta}_m$. By induction, we conclude that the optimal control inputs with respect to the myopic objective are all linear and given by $y_k(\tilde{x}_k,\tilde{\theta}_k)=\tilde{x}_k+\alpha_k\tilde{\theta}_k$ for certain α_k, $k=1,\cdots,n$. However, $y_k=x_k+\alpha_k\theta_k$ can also achieve the myopic objective (7) since $\mathbb{E}\{x_k|y_{[1,k-1]},\tilde{x}_k+\alpha_k\tilde{\theta}_k\}=\mathbb{E}\{x_k|y_{[1,k-1]},x_k+\alpha_k\theta_k\}$.

Next, we formulate a recursive technique to calculate α_k. To this end, by (17), $\tilde{\sigma}_{x,k}^2$, $\tilde{\sigma}_{\theta,k}^2$ and $\tilde{\rho}_k$ should be calculated and they are given by

$$\tilde{\sigma}_{x,k}^2=\sigma_x^2-\frac{(\mathbb{E}\{x_k\bar{x}_k\})^2}{\mathbb{E}\{\bar{x}_k^2\}},\ \tilde{\sigma}_{\theta,k}^2=\sigma_\theta^2-\frac{(\mathbb{E}\{\theta_k\bar{x}_k\})^2}{\mathbb{E}\{\bar{x}_k^2\}}, \quad (18)$$

$$\tilde{\rho}_k=-\frac{\mathbb{E}\{\theta_k\bar{x}_k\}\mathbb{E}\{x_k\bar{x}_k\}}{\mathbb{E}\{\bar{x}_k^2\}}, \quad (19)$$

for $k>1$, where $\bar{x}_k=\mathbb{E}\{x_k|y_{[1,k-1]}\}$. We let

$$\Sigma_k=\mathbb{E}\left\{\begin{bmatrix} x_k-\bar{x}_k \\ \theta_k-\bar{\theta}_k\end{bmatrix}\begin{bmatrix} x_k-\bar{x}_k \\ \theta_k-\bar{\theta}_k\end{bmatrix}^T\right\},$$

where $\bar{\theta}_k=\mathbb{E}\{\theta_k|y_{[1,k-1]}\}$ and Σ_k can be calculated recursively as

$$\Sigma_k=M\Sigma_{k-1}M-\frac{M\Sigma_{k-1}\bar{\alpha}_{k-1}\bar{\alpha}_{k-1}^T\Sigma_{k-1}M}{\bar{\alpha}_{k-1}^T\Sigma_{k-1}\bar{\alpha}_{k-1}}+\Lambda \quad (20)$$

for $k>1$ with $\Sigma_1=\begin{bmatrix}\sigma_x^2 & 0\\ 0 & \sigma_\theta^2\end{bmatrix}$, where $M=\begin{bmatrix}a & 0\\ 0 & b\end{bmatrix}$, $\Lambda=\begin{bmatrix}\sigma_w^2 & 0\\ 0 & \sigma_v^2\end{bmatrix}$, and $\bar{\alpha}_k=\begin{bmatrix}1 & \alpha_k\end{bmatrix}^T$. Note that (18) and (19) include $\mathbb{E}\{x_k\bar{x}_k\}$ and $\mathbb{E}\{\theta_k\bar{x}_k\}$. After some algebra, $\mathbb{E}\{x_k\bar{x}_k\}$ and $\mathbb{E}\{\theta_k\bar{x}_k\}$ can be calculated recursively as in (22) for $k=2,\cdots,n$, where $\mu_k\triangleq[\mathbb{E}\{x_k\bar{x}_k\},\mathbb{E}\{x_k\bar{\theta}_k\},\mathbb{E}\{\theta_k\bar{x}_k\},\mathbb{E}\{\theta_k\bar{\theta}_k\}]^T$, $\mu_1=[0,0,0,0]^T$, and

$$\begin{bmatrix}\rho_{x,k}\\ \rho_{\theta,k}\end{bmatrix}=\frac{1}{\bar{\alpha}_{k-1}^T\Sigma_{k-1}\bar{\alpha}_{k-1}}M\Sigma_{k-1}\bar{\alpha}_{k-1}.$$

Additionally, $\mathbb{E}\{\bar{x}_k^2\}$ can be calculated through $E\{(x_k-\bar{x}_k)^2\}=\sigma_x^2-2\mathbb{E}\{x_k\bar{x}_k\}+\mathbb{E}\{\bar{x}_k^2\}$ and $E\{(x_k-\bar{x}_k)^2\}$ is calculated in Σ_k, i.e., the top left element of Σ_k.

On the other side, the tracking system can calculate the output, i.e., $\hat{x}_k=\mathbb{E}\{x_k|y_{[1,k]}\}$, recursively through

$$\begin{bmatrix}\hat{x}_k\\ \hat{\theta}_k\end{bmatrix}=\left(I-\frac{\bar{\Sigma}_{k-1}\bar{\alpha}_k\bar{\alpha}_k^T}{\bar{\alpha}_k^T\bar{\Sigma}_{k-1}\bar{\alpha}_k}\right)M\begin{bmatrix}\hat{x}_{k-1}\\ \hat{\theta}_{k-1}\end{bmatrix}+\frac{\bar{\Sigma}_{k-1}\bar{\alpha}_k}{\bar{\alpha}_k^T\bar{\Sigma}_{k-1}\bar{\alpha}_k}y_k \quad (22)$$

and

$$\bar{\Sigma}_k=M\bar{\Sigma}_{k-1}M-\frac{M\bar{\Sigma}_{k-1}\bar{\alpha}_k\bar{\alpha}_k^T\bar{\Sigma}_{k-1}M}{\bar{\alpha}_k^T\bar{\Sigma}_{k-1}\bar{\alpha}_k}+\Lambda, \quad (23)$$

for $k=1,\cdots,n$, where $\begin{bmatrix}\hat{x}_0 & \hat{\theta}_0\end{bmatrix}=\begin{bmatrix}0 & 0\end{bmatrix}$ and $\bar{\Sigma}_0=\begin{bmatrix}\sigma_x^2 & 0\\ 0 & \sigma_\theta^2\end{bmatrix}$. The following theorem captures these results.

Theorem 3. *The optimal control policies with respect to the myopic objective* (7) *are linear mappings of the current desired state* x_k *and the private state* θ_k *as* $y_k=x_k+\alpha_k\theta_k$, *where* $\alpha_k=\frac{-\tilde{\sigma}_{x,k}^2+\tilde{\sigma}_{x,k}\sqrt{\tilde{\sigma}_{x,k}^2+4(\tilde{\sigma}_{\theta,k}^2+\tilde{\rho}_k)}}{2(\tilde{\sigma}_{\theta,k}^2+\tilde{\rho}_k)}$ *and* $\tilde{\sigma}_{x,k}^2$, $\tilde{\sigma}_{\theta,k}^2$, *and* $\tilde{\rho}_k$ *can be calculated recursively through* (18)-(22). *Furthermore, the optimal policies for the tracking output are linear functions of the accessed control inputs and can be calculated recursively through* (22) *and* (23).

We note that in [15], the authors study a similar problem with the trajectory controller having a myopic objective. Different from our objective here, however, they consider the more general multi-dimensional case, but under an a priori linearity restriction on the receiver (the tracking system in our setting). They show that under such a restriction, there exists a linear policy that achieves the equilibrium. Here,

$$\mu_k=\begin{bmatrix} a^2-a\rho_{x,k} & -a\rho_{x,k}\alpha_{k-1} & 0 & 0\\ -a\rho_{\theta,k} & ab-a\rho_{\theta,k}\alpha_{k-1} & 0 & 0\\ 0 & 0 & ab-b\rho_{x,k} & -b\rho_{x,k}\alpha_{k-1}\\ 0 & 0 & -b\rho_{\theta,k} & b^2-b\rho_{\theta,k}\alpha_{k-1}\end{bmatrix}\mu_{k-1}+\begin{bmatrix} a\rho_{x,k} & 0\\ a\rho_{\theta,k} & 0\\ 0 & b\rho_{x,k}\alpha_{k-1}\\ 0 & b\rho_{\theta,k}\alpha_{k-1}\end{bmatrix}\begin{bmatrix}\sigma_x^2\\ \sigma_\theta^2\end{bmatrix}, \quad (22)$$

however, we show that for the case of scalar states, linear strategies are optimal within the general class.

In the next section, we study the linear strategies that minimize (6).

V. Optimal Linear Control

To recapitulate, what we have shown above is that if the trajectory controller has a myopic objective as in (7), then the optimal control policies are linear functions of the current states, i.e., x_k and θ_k. Correspondingly, in order to formulate policies for the finite horizon cost function (6), we restrict the policy space for the control inputs to the set of all linear mappings of the current states, i.e., the control input y_k is constructed as[2] $y_k = x_k + \alpha_k \theta_k$, where $\alpha_k \in \mathbb{R}$ for $k = 1, \cdots, n$ are certain weights. We denote this function by $y_k(x_k, \theta_k)$ and the set of all such linear mappings by Υ_k, i.e., $y_k(\cdot,\cdot) \in \Upsilon_k$. More precisely, here the objective of the trajectory controller is given by

$$\min_{\substack{y_k(\cdot,\cdot)\in\Upsilon_k,\\ \forall k\in\{1,\dots,n\}}} \sum_{k=1}^{n} \mathbb{E}\left\{ \left(x_k + \theta_k - \mathbb{E}\{x_k | y_{[1,k-1]}, y_k\}\right)^2 \right\}. \quad (24)$$

Let c_k be the incurred cost at instant k such that $c_k \triangleq \mathbb{E}\{(x_k + \theta_k - \mathbb{E}\{x_k|y_{[1,k]}\})\}$. As derived in the Section IV, c_k can be written as

$$c_k = \mathbb{E}\{(\tilde{x}_k + \tilde{\theta}_k - \mathbb{E}\{\tilde{x}_k|\tilde{y}_k\})^2\} + \mathbb{E}\{(\mathbb{E}\{\theta_k|\bar{x}_k\})^2\}. \quad (25)$$

After some algebra, we obtain

$$\mathbb{E}\{(\tilde{x}_k + \tilde{\theta}_k - \mathbb{E}\{\tilde{x}_k|\tilde{y}_k\})^2\} = \tilde{\sigma}_{x,k}^2 + 2\tilde{\rho}_k + \tilde{\sigma}_{\theta,k}^2 - \frac{(\tilde{\sigma}_{x,k}^2 + \alpha_k\tilde{\rho}_k)(\tilde{\sigma}_{x,k}^2 + (\alpha_k + 2)\tilde{\rho}_k + 2\alpha_k\tilde{\sigma}_{\theta,k}^2)}{\tilde{\sigma}_{x,k}^2 + 2\alpha_k\tilde{\rho}_k + \alpha_k^2\tilde{\sigma}_{\theta,k}^2}$$

and the second term on the right hand side of (25) is given by (using (18)):

$$\mathbb{E}\{(\mathbb{E}\{\theta_k|\bar{x}_k\})^2\} = \mathbb{E}\left\{ \left(\frac{\mathbb{E}\{\theta_k\bar{x}_k\}}{\mathbb{E}\{\bar{x}_k^2\}} \bar{x}_k \right)^2 \right\} = \sigma_\theta^2 - \tilde{\sigma}_{\theta,k}^2.$$

In particular, c_k depends on $\tilde{\sigma}_k^2$, $\tilde{\rho}_k$, $\tilde{\sigma}_{\theta,k}^2$, and α_k only. Then, we introduce a state vector $z_k = \left[\sigma_{x,k}^2 \; \rho_k \; \sigma_{\theta,k}^2 \; \mathbb{E}\{x_k\bar{x}_k\} \; \mathbb{E}\{x_k\bar{\theta}_k\} \; \mathbb{E}\{\theta_k\bar{x}_k\} \; \mathbb{E}\{\theta_k\bar{\theta}_k\}\right]^T$ such that by (20) and (22), there is a nonlinear recursive relation between z_k and z_{k-1} as $z_k = f(z_{k-1}, \alpha_{k-1})$. Additionally, by (18) and (19), c_k depends on z_k and α_k in a time-invariant way and we denote the corresponding relation as $c_k = c(z_k, \alpha_k)$. Hence, we can convert problem (24) to a finite-horizon discrete time optimal control problem:

$$\min_{\substack{\alpha_k\in\mathbb{R},\\ \forall k\in\{1,\dots,n\}}} \sum_{k=1}^{n} c(z_k, \alpha_k) \text{ subject to } z_{k+1} = f(z_k, \alpha_k).$$

Then we can impose necessary conditions on the optimal control inputs through the Minimum Principle as in [22], [23] or approach the problem numerically as a nonlinear program.

Next, we analyze the strategic control of a tracking system through numerical examples.

[2]Note that having no weight assigned to x_k does not lead to any loss of generality, since y_k can always be scaled.

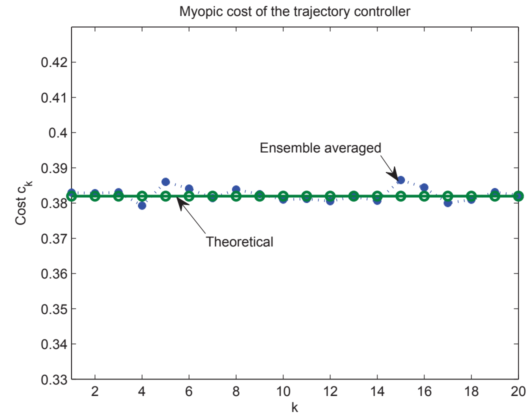

Fig. 2: Myopic cost for the trajectory controller when states evolve in parallel, i.e., $a = b$.

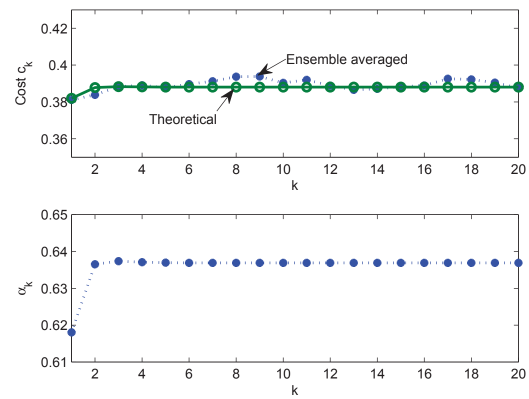

Fig. 3: Myopic cost for the trajectory controller when the private state is a more colored process, i.e., relatively more correlated in time.

VI. Numerical Examples

In this section we provide some numerical examples which serve to illustrate the results of the previous sections. To this end, we set the initial states x_1 and θ_1 as standard normal random variables, i.e., $\sigma_x^2 = \sigma_\theta^2 = 1$, and first consider the case when the states evolve in parallel, i.e., $a = b$, and the trajectory controller employs myopic policies, i.e., aims to minimize (7). Thm. 2 shows that if $a = b$, the myopic policies are linear functions of the current states and they can achieve the global minimum. In Fig. 2, we plot both theoretical and ensemble averaged (over 10^5 independent trials) incurred cost of the trajectory controller, i.e., $c_k = \mathbb{E}\{(x_k + \theta_k - \hat{x}_k)^2\}$, when the desired and the private states evolve in parallel, e.g., $a = b = 0.4$. Note that at $k = 1$, the trajectory controller aims to minimize $\mathbb{E}\{(x_1 + \theta_1 - \mathbb{E}\{x_1|y_1\})^2\}$ through y_1 over all Lebesgue measurable functions. In particular, the controller aims to achieve the global minimum at $k = 1$. As seen in Fig. 2, the controller also obtains the same minimum incurred cost for $k > 1$, which yields that the myopic policies also achieve the global minimum as expected.

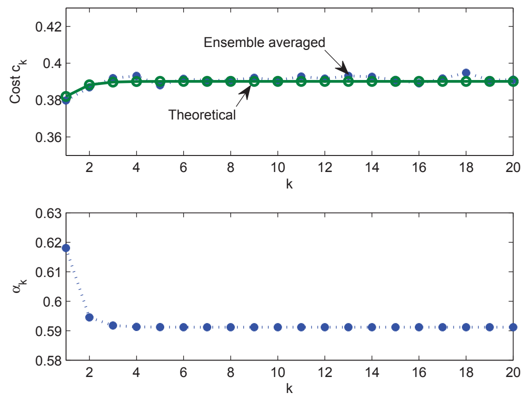

Fig. 4: Myopic cost for the trajectory controller when the desired state is a more colored process.

Next, we examine the scenarios when the states evolve differently. The Fig. 3 plots the incurred cost of the trajectory controller for myopic policies, i.e., $y_k = x_k + \alpha_k \theta_k$, when the private state θ_k is more colored, i.e., the process is more correlated in time, since $b = 0.6 > a = 0.4$. In this case, while the myopic objective y_1 achieves the global minimum at $k = 1$, the incurred cost increases at $k > 1$ and then reaches a steady-state value. We observe that α_k increases after $k = 1$ and takes values around a limit point, e.g., 0.64. In Fig. 4, we analyze the case when the desired state is a relatively more colorful process, i.e., $a = 0.6$ and $b = 0.4$. In that case, different from the Fig. 3, we see that α_k decreases in time. Particularly, in the myopic policies, weight of the relatively colorful state processes increases in time and in both cases, weights reach a steady-state value.

Finally, we study the optimal linear control for the objective (24). In the same framework with the previous examples, we have $n = 10$, $a = 0.6$, and $b = 0.4$. We restrict the policy space of the trajectory controller such that $\alpha_k \in \{0.58, 0.59, 0.6, 0.61, 0.62\}$. We have chosen these values since in Fig. 4, we observe that in the class of myopic policies for the same setup, $\alpha_k \in [0.58, 0.62]$. Similar to the myopic policies, we obtain that the non-increasing weight sequence $\{0.62, 0.61, ..., 0.61, 0.59\}$ achieves the equilibrium with respect to (24). We point out that the corresponding cost is 3.887 which is slightly less than the cost achieved by the myopic policies, i.e., 3.891. This also implies that the myopic policies are not optimal with respect to the finite horizon cost (6) in general.

VII. Conclusion

In this paper, we have addressed the strategic control of a tracking system in the context of dynamic Stackelberg equilibrium. With finite horizon objectives, we have shown that a linear function of the current states can lead to a Stackelberg equilibrium if the states evolve in parallel. Furthermore, if the trajectory controller has a myopic objective, there also exists a linear function of the current states that leads to equilibrium in the general case. Finally, we have shown that restricting the policy space of the trajectory controller to the set of all linear mappings of the current states, the stochastic dynamic game problem with finite horizon cost can be viewed as a discrete time optimal control problem. Some future directions of research on this topic include the analysis of the policies when there is a channel between the trajectory controller and the tracking system, and the strategic control of a control system, extending [1] to the strategic setting.

References

[1] R. Bansal and T. Başar, "Simultaneous design of measurement and control strategies for stochastic systems with feedback," *Automatica*, vol. 25, no. 5, pp. 679–694, 1989.

[2] S. Tatikonda, A. Sahai, and S. K. Mitter, "Stochastic linear control over a communication channel," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1549–1561, 2004.

[3] D. Teneketzis, "On the structure of optimal real time encoders and decoders in noisy communication," *IEEE Trans. Info. Theory*, vol. 52, no. 9, pp. 4017–4035, 2006.

[4] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proc. IEEE*, vol. 95, no. 1, pp. 138–162, 2005.

[5] C.-L. Huang, R. Sengupta, and J. Huang, "Real-time tracking of linear processes over an AWGN channel without feedback: the MMSE optimality of the innovation encoder," in *Proc. 50th IEEE Conf. on Decision and Control and Europen Control Conf. (CDC-ECC)*, 2011.

[6] T. Başar and R. Bansal, "Optimum design of measurement channels and control policies for linear-quadratic stochastic systems," *European Journal of Operations Research*, vol. 73, pp. 226–236, 1994.

[7] V. Crawford and J. Sobel, "Strategic information transmission," *Econometrica*, pp. 1431–1451, 1982.

[8] J. Farrell and M. Rabin, "Cheap talk," *J. Econ. Pers.*, pp. 103–118, 1996.

[9] M. Battaglini, "Multiple referrals and multidimensional cheap talk," *Econometrica*, vol. 70, no. 4, pp. 1379–1401, 2002.

[10] M. Ottaviani and P. N. Sorensen, "Professional advice," *J. Econ. Theory*, vol. 126, pp. 120–142, 2006.

[11] M. Golosov, V. Skreta, A. Tsyvinski, and A. Wilson, "Dynamic strategic information transmission," *Journal of Economic Theory*, vol. 151, pp. 304–341, 2014.

[12] E. Akyol, C. Langbort, and T. Başar, "Strategic compression and transmission of information," in *Proc. IEEE Info. Theory Workshop*, 2015.

[13] ——, "Information-theoretic approach to strategic communication as a hierarchical game," *Proc. IEEE*, 2017, in press.

[14] T. Kailath, "An innovations approach to least-squares estimation part I: Linear filtering in additive white noise," *IEEE Transactions on Automatic Control*, vol. AC-13, no. 6, pp. 646–655, 1968.

[15] F. Farokhi, A. Teixeira, and C. Langbort, "Estimation with strategic sensors," *IEEE Trans. Automatic Control*, 2016, in press.

[16] T. Başar and G. Olsder, *Dynamic Noncooperative Game Theory*. Society for Industrial Mathematics (SIAM) Series in Classics in Applied Mathematics, 1999.

[17] H. von Stackelberg, *Market Structure and Equilibrium*. Springer-Verlag Wien, 2013.

[18] D. Blackwell and C. Ryll-Nardzewski, "Non-existence of everywhere proper conditional distributions," *The Annals of Mathematical Statistics*, vol. 34, pp. 223–225, 1962.

[19] D. Blackwell, "Memoryless strategies in finite-stage dynamic programming," *Annals Math. Statistics*, vol. 35, pp. 863–865, 1963.

[20] S. Yüksel and T. Başar, *Stochastic Network Control Systems: Stabilization and Optimization Under Information Constraints*. Birkhäuser, 2013.

[21] C. Preston, "A note on standard Borel and related spaces," *Journal of Contemporary Mathematical Analysis*, vol. 44, no. 1, 2009.

[22] V. Guibout and A. M. Bloch, "A discrete maximum principle for solving optimal control problems," in *Proc. 43rd IEEE Conf. on Decision Control*, 2004, pp. 1806–1811.

[23] T. Ohsawa, A. M. Bloch, and M. Leok, "Discrete Hamiltonian-Jacobi Theory and discrete optimal control," in *Proc. 43rd IEEE Conf. on Decision Control*, 2010, pp. 5438–5443.