Sequential Empirical Coordination Under an Output Entropy Constraint

Ehsan Shafieepoorfard and Maxim Raginsky

Abstract— **This paper considers the problem of sequential empirical coordination, where the objective is to achieve a given value of the expected uniform deviation between state-action empirical averages and statistical expectations under a given strategic probability measure, with respect to a universal Glivenko-Cantelli class of test functions. A communication constraint is imposed on the Shannon entropy of the resulting action sequence. It is shown that the fundamental limit on the output entropy is given by the minimum of the mutual information between the state and the action processes under all strategic measures that have the same marginal state process as the target measure and approximate the target measure to desired accuracy with respect to the underlying Glivenko-Cantelli seminorm. The fundamental limit is shown to be asymptotically achievable by tree-structured codes.**

I. Introduction

Rate-distortion theory [1] is a branch of information theory concerned with fundamental limits of compressed representations of stochastic processes subject to reconstruction fidelity criteria. The issue of efficient compressed representations arises in the context of networked control systems as well [2], [3]. For instance, consider a network involving a large number of decision-makers (DM's), each of whom can only observe the state of the plant through a noiseless digital channel with finite capacity. In such situations, we must design not only the control policies for the DM's, but also the *encoders* (or quantizers) for transmitting suitably compressed representations of the state process to the DM's. However, in the context of control, the relevant operational criterion is not the fidelity of reconstruction of the state process at the DM's, but rather the performance of control policies that govern the implementation of actions on the basis of the compressed representation of the state of the plant.

The book of Yüksel and Başar [3] contains numerous references to the vast literature on this problem of *control under communication constraints*. In most of these works, a suitable form of the principle of separation between estimation and control is established for the information structure at hand, and then the techniques of rate-distortion theory are brought to bear on the problem of quantizer design. The main difference between the control-oriented use of rate-distortion theory and the traditional communication-oriented use is that, in the former, one has to take into account the issues of delay and real-time implementability of control actions. To handle these issues, Tatikonda has developed a *sequential* generalization of rate-distortion theory [4, Ch. 5]; a recent preprint of Tanaka et al. [5] presents a nice treatment of LQG control under information constraints in the context of sequential rate-distortion theory.

However, there is an alternative perspective on the problem of compressed representations in networked control systems – that of *empirical coordination* under communication constraints. The problem of coordination, first introduced in the information theory literature by Cuff et al. [6], can be stated as follows: Consider a finite collection of DM's, who wish to generate actions in response to a random state variable according to some prescribed policy, but can only receive information about the state over finite-capacity noiseless digital links. Suppose that we have a large number of independent and identically distributed (i.i.d.) copies of the state, and let the DM's generate a sequence of actions based on the information they receive about this state sequence. What are the minimal communication requirements (in bits per copy), to guarantee that the long-term empirical frequencies of realized states and actions approximate, to desired accuracy, the ideal joint probability law of states and actions induced by the marginal law of the state and the policy?

Cuff et al. [6] assume that both the state and the actions take values in finite sets, and measure the quality of approximation by the total variation distance between the empirical distribution of states and actions and the target joint distribution. However, this criterion is inapplicable to continuous-valued states and/or actions with nonatomic probability laws because the total variation distance between any nonatomic probability measure and any discrete probability measure is identically equal to one. To resolve this issue, Raginsky [7] proposed a relaxed approximation criterion: Fix a suitable class of bounded real-valued test functions on the space of all state-action pairs and consider the worst-case deviation between their empirical averages and their expectations with respect to the target measure. Under the regularity assumption that the class of test functions has the so-called *universal Glivenko–Cantelli property* (cf. [8] and references therein, as well as Section II for definitions), Raginsky [7] obtained a full information-theoretic characterization of the minimal communication requirements for empirical coordination. Since any uniformly bounded class of real-valued functions on a finite set is universal Glivenko–Cantelli, the framework of [6] emerges as a special case.

In this paper, we present an extension of the empirical coordination framework of [7] to the sequential setting: We consider a network comprising an Information Source (IS) and a large number N of DM's, designated as

Research supported in part by the NSF under CAREER award no. CCF-1254041, and by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-0939370.

E. Shafieepoorfard and M. Raginsky are with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL, 61801 USA. Emails: {shafiee1,maxim}@illinois.edu.

$\mathrm{DM}_1, \dots, \mathrm{DM}_N$. The IS observes N independent copies of a discrete-time state process of fixed finite duration T. There is a fixed causal policy for generating actions contingent on the states, and each DM must implement this policy on its own copy of the state process. However, no DM has direct access to the IS. Instead, the IS can broadcast public signals to all DM's over a finite-capacity noiseless digital channel. We are interested in the minimal communication requirements that are needed to guarantee that, in the limit as $N \to \infty$, the empirical distribution of states and actions across all DM's at each time $t \in [T]$ can approximate the state-action distribution induced by the state process law and by the policy specification. The coding scheme employed by the IS must satisfy the sequentiality constraint: The signal transmitted by the IS to the DM's at time t may only depend on the realizations of the state processes up to time t. Following Tatikonda [4] (see also [9]), we quantify the communication resources by the Shannon entropy of the signal process.

Our main contribution is a full information-theoretic characterization of the fundamental limit on the amount of communication from the IS to the DM's in the setting of sequential empirical coordination. We refer to this fundamental limit as the *sequential rate-distortion function for empirical coordination*. Specifically, we show that, for all large enough N, this fundamental limit can be achieved by means of tree-structured codes of the kind employed by Tatikonda [4], and that no sequential scheme for empirical coordination can beat this fundamental limit. While we do not make any structural assumptions on the state process (e.g., it is not assumed to be memoryless, Markov, ergodic, etc.), we assume that the target policy is feedforward (i.e., there is no functional dependence of future states on current and past actions).

II. Preliminaries and notation

All spaces are assumed to be standard Borel (i.e., isomorphic to a Borel subset of a Polish space) [10], and will be equipped with their Borel σ-fields. Given such a space $(\mathsf{X}, \mathcal{B}(\mathsf{X}))$, we will denote by $\mathcal{S}(\mathsf{X})$ the space of all finite signed Borel measures on X, by $\mathcal{P}(\mathsf{X}) \subset \mathcal{S}(\mathsf{X})$ the space of all Borel probability measures on X, by $M(\mathsf{X})$ the space of all measurable functions $\mathsf{X} \to \mathbb{R}$, and by $M_b(\mathsf{X})$ the subspace of all $f \in M(\mathsf{X})$ with $\|f\|_\infty \triangleq \sup_{x \in \mathsf{X}} |f(x)| < \infty$. We will often use the inner-product notation for integrals: for any $\nu \in \mathcal{S}(\mathsf{X})$ and any $f \in M(\mathsf{X})$,

$$\langle \nu, f \rangle \triangleq \int_{\mathsf{X}} f \, \mathrm{d}\nu.$$

A Markov kernel from X into another space U is a mapping $K \colon \mathsf{X} \times \mathcal{B}(\mathsf{U}) \to [0,1]$, such that $K(x, \cdot) \in \mathcal{P}(\mathsf{U})$ for all $x \in \mathsf{X}$ and $K(\cdot, B) \in M(\mathsf{X})$ for all $B \in \mathcal{B}(\mathsf{U})$. We denote the space of all such Markov kernels by $\mathcal{M}(\mathsf{U}|\mathsf{X})$. Any $\mu \in \mathcal{P}(\mathsf{X})$ and $K \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ induce a probability measure $\mu \otimes K \in \mathcal{P}(\mathsf{X} \times \mathsf{U})$, defined by

$$\mu \otimes K(A \otimes B) \triangleq \int_A \mu(\mathrm{d}x) K(x, B), \ \forall A \in \mathcal{B}(\mathsf{X}),\ B \in \mathcal{B}(\mathsf{U});$$

then $\mu K(B) \triangleq \mu \otimes (\mathsf{X} \times B)$ denotes the induced marginal law on U. Since any $\nu \in \mathcal{P}(\mathsf{U})$ can be viewed as a Markov kernel $(x, B) \mapsto \nu(B)$, all product measures $\mu \otimes \nu \in \mathcal{P}(\mathsf{X} \times \mathsf{U})$ are a special case of this construction.

Let $M_b^1(\mathsf{X}) \triangleq \{f \in M_b(\mathsf{X}) : \|f\|_\infty \le 1\}$. Given any $\mathcal{F} \subseteq M_b^1(\mathsf{X})$, we define the following seminorm on $\mathcal{S}(\mathsf{X})$:

$$\|\nu\|_{\mathcal{F}} \triangleq \sup_{f \in \mathcal{F}} |\langle \nu, f \rangle|.$$

The *empirical distribution* of $x^n = (x_1, \dots, x_n) \in \mathsf{X}^n$ is

$$\mathsf{P}_{x^n} \triangleq \frac{1}{n} \sum_{i=1}^n \delta_{x_i},$$

where δ_x denotes the Dirac measure centered at x. We say that $\mathcal{F} \subseteq M_b^1(\mathsf{X})$ is a universal Glivenko–Cantelli class (or uGC class, for short) if, for any $\mu \in \mathcal{P}(\mathsf{X})$,

$$\|\mathsf{P}_{X^n} - \mu\|_{\mathcal{F}} \xrightarrow{\text{a.s.}} 0 \qquad \text{as } n \to \infty,$$

where $X_1, X_2, \dots$ is a sequence of independent and identically distributed (i.i.d.) random elements of X with common marginal probability law μ. The requirement of niform boundedness of $\mathcal{F}$ cannot be dropped, since it can be shown that if $\mathcal{F}$ is uGC, then the class $\mathcal{F}_0 \triangleq \{f - \inf f : f \in \mathcal{F}\}$ must be uniformly bounded [11, Sec. 6.6]. It should be pointed out that, unless the underlying space is discrete, $\mathcal{F} \equiv M_b^1(\mathsf{X})$ is not uGC [7, Sec. IV].

We will also need some basic notions and definitions from information theory [12], [13]. The *relative entropy* (or information divergence) between $\mu, \nu \in \mathcal{P}(\mathsf{X})$ is

$$D(\mu\|\nu) \triangleq \begin{cases} \left\langle \mu, \log \dfrac{\mathrm{d}\mu}{\mathrm{d}\nu} \right\rangle, & \text{if } \mu \ll \nu \\ +\infty, & \text{otherwise} \end{cases},$$

where $\mu \ll \nu$ denotes absolute continuity of μ w.r.t. ν. The *Shannon mutual information* in a pair $(\mu, K) \in \mathcal{P}(\mathsf{X}) \times \mathcal{M}(\mathsf{U}|\mathsf{X})$ is then given by

$$I(\mu, K) \triangleq D(\mu \otimes K \| \mu \otimes \mu K);$$

if (X, U) is a random element of $\mathsf{X} \times \mathsf{U}$ with law $\mu \otimes K$, then we will also write $I(X; U)$ or $I_{\mu,K}(X; U)$ for $I(\mu, K)$ and call it the mutual information between X and U. Given a random triple (X, U, Z) taking values in $\mathsf{X} \times \mathsf{U} \times \mathsf{Z}$, the *conditional mutual information* between X and U given Z is defined as

$$I(X; U|Z) \triangleq \int_{\mathsf{Z}} P_Z(\mathrm{d}z) D(P_{XU|Z=z} \| P_{X|Z=z} \otimes P_{U|Z=z}),$$

where P_Z denotes the marginal law of Z, $P_{XU|Z=z}$ denotes the regular conditional probability distribution of (X, U) given $Z = z$, etc. If U is a random element of U with a purely atomic probability law ν, then its *Shannon entropy* is

$$H(U) \triangleq -\sum_u \nu[U = u] \log \nu[U = u],$$

where the sum is over the atoms of ν. For an arbitrary random element Z jointly distributed with U, the *conditional entropy* of U given Z is given by

$$H(U|Z) \triangleq \int_{\mathsf{Z}} P_Z(\mathrm{d}z) H(U|Z = z),$$

where $H(U|Z = z)$ is the Shannon entropy of U w.r.t. the regular conditional probability distribution $P_{U|Z=z}$. We work with natural logarithms throughout, so all entropies and mutual informations are measured in nats.

III. Problem formulation

We now give the precise formulation of the sequential empirical coordination problem, which was informally stated in the Introduction. Fix a state space X, an action space U, and a finite time horizon $T \in \mathbb{N}$. Let $\mu \in \mathcal{P}(\mathsf{X}^T)$ be a given state process law, and let $\pi = (\pi^{(t)})_{t=1}^T$ be a given sequence of Markov kernels $\pi^{(t)} \in \mathcal{M}(\mathsf{U}|\mathsf{X}^t \times \mathsf{U}^{t-1})$. Together, these objects define the *strategic measure* $\mathbf{P}^\pi_\mu \in \mathcal{P}(\mathsf{X}^T \times \mathsf{U}^T)$ by

$$\mathbf{P}^\pi_\mu(\mathrm{d}x^T, \mathrm{d}u^T) \triangleq \mu(\mathrm{d}x^T) \otimes \bigotimes_{t=1}^T \pi_t(\mathrm{d}u_t|x^t, u^{t-1})$$
$$\equiv \bigotimes_{t=1}^T \mu^{(t)}(\mathrm{d}x_t|x^{t-1}) \otimes \pi^{(t)}(\mathrm{d}u_t|x^t, u^{t-1}), \qquad (1)$$

where, for each t, $\mu^{(t)} \in \mathcal{M}(\mathsf{X}_t|\mathsf{X}^{t-1})$ is the regular conditional probability distribution of X_t given X^{t-1} induced by μ. In operational terms, π prescribes a causal policy for a DM to take actions in U based on the past history of states and actions. As evident from (1), we assume that π is a feedforward policy, i.e., there is no functional dependence of future states on current and past actions.

Consider the situation when N DM's, designated as $\mathrm{DM}_1, \dots, \mathrm{DM}_N$, observe independent copies of the state process with law μ and implement the policy π. Formally, let $\boldsymbol{X} = (X_{t,n} : t \in [T], n \in [N])$ be a $T \times N$ array of random elements of X. For all $A \subseteq [T]$ and $B \subseteq [N]$, we will denote by $X_{A,B}$ the subarray $(X_{t,n} : t \in A, n \in B)$. Then the columns $X_{[T],n} = (X_{t,n})_{t\in[T]}$, for $n \in [N]$, are i.i.d. copies of the state process with law μ. Similarly, let $\boldsymbol{U} = (U_{t,n} : t \in [T], n \in [N])$ be an array of random elements of U, such that the pair processes $(X_{[T],n}, U_{[T],n})$, $n \in [N]$, are i.i.d. copies of the state-action process with law $\mathbf{P}^\pi_\mu$. We will abuse the notation and also write $\mathbf{P}^\pi_\mu$ for the probability law of the random couple $(\boldsymbol{X}, \boldsymbol{U})$. Operationally speaking, DM_n receives sequential observations of the state process $X_{[T],n}$ and generates the actions $U_{[T],n}$ using the policy π. Let μ_t denote the marginal law of X_t under $\mathbf{P}^\pi_\mu$, and let π_t denote the conditional law of U_t given X_t under $\mathbf{P}^\pi_\mu$. Then, for any uGC class $\mathcal{F}$ of functions on $\mathsf{X} \times \mathsf{U}$,

$$\left\|\mathsf{P}_{X_{t,[N]},U_{t,[N]}} - \mu_t \otimes \pi_t\right\|_{\mathcal{F}} \xrightarrow{\text{a.s.}} 0 \quad \text{as } N \to \infty \qquad (2)$$

where

$$\mathsf{P}_{X_{t,[N]},U_{t,[N]}} \triangleq \frac{1}{N}\sum_{n=1}^N \delta_{(X_{t,n},U_{t,n})}$$

is the empirical distribution of states and actions at time t. Since (2) holds for every $t \in [T]$, we have

$$\frac{1}{T}\sum_{t=1}^T \left\|\mathsf{P}_{X_{t,[N]},U_{t,[N]}} - \mu_t \otimes \pi_t\right\|_{\mathcal{F}} \xrightarrow{\text{a.s.}} 0 \quad \text{as } N \to \infty.$$

In other words, if a large number of independent DM's implement the same policy π on their private copies of the state process, then the state-action empirical distributions at each time t will closely resemble the true marginal distributions $\mu_t \otimes \pi_t$ of (X_t, U_t), with high probability.

We are interested in the following problem: Suppose that an Information Source (IS) receives causal observations of $\boldsymbol{X}$, i.e., at each time $t \in [T]$ the IS observes $X_{[t],[N]}$. The IS can transmit information to the DM's via a common capacity-limited noiseless digital channel. A *sequential N-code* is a collection $\gamma = (\gamma_t)_{t\in[T]}$ of measurable mappings $\gamma_t : \mathsf{X}_{[t],[N]} \to \mathsf{U}_{t,[N]}$ with countable ranges. The state process law $\mu \in \mathcal{P}(\mathsf{X}^T)$ and the code γ specify a unique[1] probability law $\mathbf{P}^\gamma_\mu$ for the random couple $(\boldsymbol{X}, \boldsymbol{U})$ via

$$\mathbf{P}^\gamma_\mu(\mathrm{d}x_{[T],[N]}, \mathrm{d}u_{[T],[N]})$$
$$\triangleq \bigotimes_{n=1}^N \mu(\mathrm{d}x_{[T],n}) \otimes \bigotimes_{t=1}^T \delta_{\gamma_t(x_{[t],[N]}, u_{[t-1],[N]})}(\mathrm{d}u_{t,[N]}) \quad (3)$$

For a given $\Delta \in [0, 1]$, define the set

$$\Gamma^N_{\mu,\pi}(\Delta) \triangleq \Bigg\{\gamma = (\gamma_t)_{t\in[T]} :$$
$$\frac{1}{T}\sum_{t=1}^T \mathbb{E}^\gamma_\mu \left\|\mathsf{P}_{(X_{t,[N]},U_{t,[N]})} - \mu_t \otimes \pi_t\right\|_{\mathcal{F}} \le \Delta\Bigg\}$$

of all sequential N-codes whose induced state-action frequencies approximate the true state-action distributions under the strategic measure $\mathbf{P}^\pi_\mu$ to accuracy Δ in the $\|\cdot\|_{\mathcal{F}}$ seminorm. We are interested in the minimal communication requirements for approximating $\mathbf{P}^\pi_\mu$ in this sense, which motivates our definition of the Nth-order *operational sequential rate-distortion function for empirical coordination*:

$$\widehat{R}_{T,N}(\Delta) \triangleq \inf_{\gamma \in \Gamma^N_{\mu,\pi}(\Delta)} \frac{H\left(\gamma_t(X_{[t],[N]})_{t\in[T]}\right)}{NT},$$

where $H(\cdot)$ is the Shannon entropy. (Since γ_t's have countable ranges, the marginal law of $\boldsymbol{U}$ under $\mathbf{P}^\gamma_\mu$ is purely atomic, and therefore its Shannon entropy is well-defined.)

IV. Main results and some implications

Our main results, stated as Theorems 1 and 2 below, provide a full characterization of the operatinoal sequential rate-distortion function $\widehat{R}_{T,N}(\Delta)$. In particular, they show that, asymptotically as $N \to \infty$, $\widehat{R}_{T,N}(\Delta)$ is given by a purely information-theoretic expression that quantifies the minimum amount of information between the state process $X^T \sim \mu$ and any jointly distributed action process U^T that can be generated from X^T using a causal policy, such that the empirical frequencies of the state-action distributions can approximate the target strategic measure $\mathbf{P}^\pi_\mu$ to accuracy Δ on the given uGC class $\mathcal{F}$.

We start by defining this information-theoretic quantity. Let $\vec{\mathcal{M}}(\mathsf{U}^T|\mathsf{X}^T)$ denote the collection of all horizon-T causal

[1]Existence and uniqueness follow from the Ionescu Tulcea theorem [14, Thm. 6.17].

policies, i.e., all sequences $\tilde{\pi} = (\tilde{\pi}^{(t)})_{t=1}^T$ of Markov kernels $\tilde{\pi}^{(t)} \in \mathcal{M}(\mathsf{U}|\mathsf{X}^t \times \mathsf{U}^{t-1})$, $t \in [T]$. The state process law μ and each $\tilde{\pi} \in \vec{\mathcal{M}}(\mathsf{U}^T|\mathsf{X}^T)$ induce a unique strategic measure $\mathbf{P}_\mu^{\tilde{\pi}} \in \mathcal{P}(\mathsf{X}^T \times \mathsf{U}^T)$, just like in (1). Define the set

$$\Pi_{\mu,\pi}(\Delta) \triangleq \Bigg\{ \tilde{\pi} \in \vec{\mathcal{M}}(\mathsf{U}^T|\mathsf{X}^T) : \frac{1}{T}\sum_{t=1}^T \|\mu_t \otimes \tilde{\pi}_t - \mu_t \otimes \pi_t\|_{\mathcal{F}} \le \Delta \Bigg\},$$

where $\mu_t \in \mathcal{P}(\mathsf{X})$ is the marginal distribution of X_t under μ, and $\tilde{\pi}_t \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ is the conditional law of U_t given X_t under $\mathbf{P}_\mu^{\tilde{\pi}}$. We now introduce the *sequential rate-distortion function for empirical coordination*:

$$R_T(\Delta) \triangleq \inf_{\tilde{\pi} \in \Pi_{\mu,\pi}(\Delta)} \frac{I_{\mu,\tilde{\pi}}(X^T;U^T)}{T},$$

where $I_{\mu,\tilde{\pi}}(X^T;U^T)$ is the Shannon mutual information between the state process X^T and the action process U^T when $(X^T,U^T) \sim \mathbf{P}_\mu^{\tilde{\pi}}$.

We now state two theorems that reveal $T_T(\Delta)$ as the fundamental limit in our empirical coordination problem:

Theorem 1 (Achievability)**.** *Suppose that $R_T(\Delta) < \infty$. Then, for each $\varepsilon > 0$ there exists $N(\varepsilon) \in \mathbb{N}$, such that*

$$\widehat{R}_{T,N}(\Delta + \varepsilon) \le R_T(\Delta) + \varepsilon. \tag{4}$$

That is, under the conditions of the theorem, for each sufficiently large N we can find an N-code in $\Gamma_{\mu,\pi}^N(\Delta)$, whose per-DM, per-time step output entropy is (approximately) bounded by $R_T(\Delta)$.

Theorem 2 (Converse)**.** *For any N, T, and Δ,*

$$\widehat{R}_{T,N}(\Delta) \ge R_T(\Delta). \tag{5}$$

That is, the per-DM, per-time step output entropy of any N-code $\gamma \in \Gamma_{\mu,\pi}^N(\Delta)$ must be at least as large as $R_T(\Delta)$.

The proofs of Theorems 1 and 2 are given in Section V.

Although Theorems 1 and 2 provide a full characterization of the fundamental limits on the minimal rate of communication for sequential empirical coordination, the computation of the sequential rate distortion function $R_T(\Delta)$ is a complicated infinite-dimensional optimization problem already in the static ($T = 1$) case, which was addressed in [7]. Below, we provide two examples that nicely illustrate the difficulty of explicitly computing $R_T(\Delta)$ even for $T = 1$.

A. Kolmogorov–Smirnov criterion for one-step costs

While we have remained silent on the nature of the target strategic measure $\mathbf{P}_\mu^\pi$, it may have been selected based on considerations of expected cost. Thus, suppose that we have a function $c \in M(\mathsf{X} \times \mathsf{U})$, such that $c(x,u)$ gives the cost of taking action u in response to state x. Let $\mathcal{F}$ be the class of indicator functions of the level sets of c:

$$f(x,u) \triangleq 1\{c(x,u) \le a\}, \qquad a \in \mathbb{R}. \tag{6}$$

Then we have the following:

Proposition 1. *Let $\mathcal{F}$ denote the class of all f of the form (6). Then, for any two $P,Q \in \mathcal{P}(\mathsf{X} \times \mathsf{U})$,*

$$\|P - Q\|_{\mathcal{F}} = d_{\mathrm{KS}}(F_{P \circ c^{-1}}, F_{Q \circ c^{-1}}), \tag{7}$$

where F_μ denotes the cumulative distribution function (cdf) of a Borel probability measure μ on the reals, and

$$d_{\mathrm{KS}}(F,F') \triangleq \sup_{a \in \mathbb{R}} |F(a) - F'(a)|$$

is the Kolmogorov–Smirnov distance *between cdf's F and F'. The class $\mathcal{F}$ is a universal Glivenko–Cantelli class.*

The following is immediate from the above proposition:

Theorem 3.

$$R_T(\Delta) = \inf_{\tilde{\pi} \in \vec{\mathcal{M}}(\mathsf{U}^T|\mathsf{X}^T)} \Big\{ I_{\mu,\tilde{\pi}}(X^T;U^T) : \frac{1}{T}\sum_{t=1}^T d_{\mathrm{KS}}(F_{(\mu_t \otimes \tilde{\pi}_t)\circ c^{-1}}, F_{(\mu_t \otimes \pi_t)\circ c^{-1}}) \le \Delta \Big\}.$$

In other words, $R_T(\Delta)$ is the smallest mutual information between the state process $X^T \sim \mu$ and any action process U^T generated from X^T by a causal policy $\tilde{\pi}$, such that the time average of the Kolmogorov-Smirnov distances between the state-action costs under $\tilde{\pi}$ and the target policy π is bounded from above by Δ. Evaluating this quantity exactly is difficult even for $T = 1$.

B. Weak convergence and Wasserstein distances

Another example concerns the approximation of the target strategic measure $\mathbf{P}_\mu^\pi$ in a certain metric that metrizes the topology of weak convergence of probability measures. Suppose that $\mathsf{X} \times \mathsf{U}$ is a Polish space with a given metric d. For any $f \in M_b(\mathsf{X} \times \mathsf{U})$, define the *Lipschitz norm*

$$\|f\|_{\mathrm{Lip}} \triangleq \sup_{(x,u),(y,v) \in \mathsf{X}\times\mathsf{U}} \frac{|f(x,u) - f(y,v)|}{d((x,u),(y,v))}.$$

and the *bounded Lipschitz norm* $\|f\|_{\mathrm{BL}} \triangleq \|f\|_\infty + \|f\|_{\mathrm{Lip}}$.

Proposition 2. *Consider the function class $\mathcal{F} = \{f \in M_b^1(\mathsf{X} \times \mathsf{U}) : \|f\|_{\mathrm{BL}} \le 1\}$. Then, for any $P,Q \in \mathcal{P}(\mathsf{X} \times \mathsf{U})$,*

$$\|P - Q\|_{\mathcal{F}} = d_{\mathrm{BL}}(P,Q), \tag{8}$$

the bounded Lipschitz metric *on $\mathcal{P}(\mathsf{X} \times \mathsf{U})$ that metrizes the topology of weak convergence of probability measures. The class $\mathcal{F}$ is a universal Glivenko–Cantelli class.*

Under an additional moment condition, the bounded Lipschitz metric can be upper-bounded by the so-called Wasserstein metric. Let $\mathcal{P}_0(\mathsf{X} \times \mathsf{U}) \subseteq \mathcal{P}(\mathsf{X} \times \mathsf{U})$ be the set of all probability measures P, for which there exists some $(x_0,u_0) \in \mathsf{X} \times \mathsf{U}$, such that $\langle P, d(\cdot,(x_0,u_0))\rangle < \infty$. The *Wasserstein metric* between any two $P,Q \in \mathcal{P}_0(\mathsf{X} \times \mathsf{U})$ is

$$W_d(P,Q) \triangleq \sup_{\|f\|_{\mathrm{Lip}} \le 1} |\langle P,f\rangle - \langle Q,f\rangle|.$$

We can now give the following upper bound on the sequential rate-distortion function $R_T(\Delta)$ w.r.t. $\mathcal{F}$:

Theorem 4. *Suppose that* $\mathbf{P}^{\pi}_{\mu} \in \mathcal{P}_0(\mathsf{X} \times \mathsf{U})$. *Then*

$$R_T(\Delta) \le \inf_{\tilde{\pi} \in \vec{\mathcal{M}}(\mathsf{U}^T|\mathsf{X}^T)} \Big\{ I_{\mu,\tilde{\pi}}(X^T;U^T) : \frac{1}{T}\sum_{t=1}^T W_d(\mu_t \otimes \tilde{\pi}_t, \mu_t \otimes \pi_t) \le \Delta \Big\}.$$

Despite the clean conceptual interpretation of $R_T(\Delta)$ in terms of approximating strategic measures by empirical distributions under the bounded Lipschitz metric, it does not admit closed-form expressions even for $T = 1$.

V. Proofs

Achievability. Since $R_T(\Delta) < \infty$, for a given $\varepsilon > 0$ there exists some $\tilde{\pi} \in \vec{\mathcal{M}}(\mathsf{U}^T|\mathsf{X}^T)$, such that

$$\frac{1}{T}\sum_{t=1}^T \|\mu_t \otimes \tilde{\pi}_t - \mu_t \otimes \pi_t\|_{\mathcal{F}} \le \Delta + \frac{\varepsilon}{2} \tag{9}$$

and $I_{\mu,\tilde{\pi}}(X^T;U^T) \le R_T(\Delta) + \varepsilon/2$. For each t, let

$$\Delta_{t,N} \triangleq \Big\| \mathsf{P}_{(X_{t,[N]},U_{t,[N]})} - \mu_t \otimes \tilde{\pi}_t \Big\|_{\mathcal{F}},$$

and define the set

$$A_{t,N} \triangleq \Big\{ (X_{t,[N]}, U_{t,[N]}) : \Delta_{t,N} > \sqrt{\delta_{t,N}} \Big\},$$

where $\delta_{t,N} = \mathbb{E}^{\tilde{\pi}}_{\mu}[\Delta_{t,N}]$. By Markov's inequality, $\mathbf{P}^{\tilde{\pi}}_{\mu}[A_{t,N}] \le \sqrt{\delta_{t,N}}$. Let $M_t = e^{N(R_t+\varepsilon/2)} - 1$, where $R_t = I_{\mu,\tilde{\pi}}(X^t;U_t|U^{t-1})$. Using a random selection argument (omitted here for lack of space), it can be shown that there exist measurable mappings $\gamma_t : \mathsf{X}_{[t],[N]} \to \mathsf{U}_{t,[N]}$, $t \in [T]$, such that

$$\begin{aligned}
&\max_{t\in[T]} \mathbf{P}_{\mu}\left[(X_{t,[N]}, \gamma_t(X_{[t],[N]})) \in A_{t,N}\right] \\
&\le T \max_{t\in[T]} \Big\{ \mathbf{P}^{\tilde{\pi}}_{\mu}[A_{t,N}] \\
&\quad + \mathbf{P}^{\tilde{\pi}}_{\mu}\left[i_t(X_{[t],[N]}, U_{[t],[N]}) > N(R_t + \varepsilon/4)\right] \\
&\quad + \exp\left(-M_t e^{-N(R_t+\varepsilon/4)}\right) \Big\} \\
&\le T\exp(-e^{N\varepsilon/4}) + T \max_{t\in[T]} \Big\{ \sqrt{\delta_{t,N}} \\
&\quad + \mathbf{P}^{\tilde{\pi}}_{\mu}\left[i_t(X_{[t],[N]}, U_{[t],[N]}) > N(R_t + \varepsilon/4)\right] \Big\}.
\end{aligned} \tag{10}$$

For all sufficiently large N, the right-hand side of (10) can be made smaller than $\varepsilon/4$. Indeed, since $\mathcal{F}$ is a uGC class, $\max_{t\in[T]} \delta_{t,N} \to 0$ as $N \to \infty$. Moreover, since

$$i_t(X_{[t],[N]}, U_{[t],[N]}) = \sum_{n=1}^N \log \frac{\mathrm{d}\mathbf{P}^{\mu,\tilde{\pi}}_{U_{t,n}|X_{[t],n},U_{[t-1],n}}}{\mathrm{d}\mathbf{P}^{\mu,\tilde{\pi}}_{U_{t,n}|U_{[t-1],n}}}(X_{[t],n}, U_{[t],n})$$

is a sum of i.i.d. random variables, and

$$R_t = I_{\mu,\tilde{\pi}}(X^t;U_t|U^{t-1}) = \frac{1}{N}\mathbb{E}^{\tilde{\pi}}_{\mu}\left[i_t(X_{[t],[N]}, U_{[t],[N]})\right],$$

the quantity

$$\max_{t\in[T]} \mathbf{P}^{\tilde{\pi}}_{\mu}\left[i_t(X_{[t],[N]}, U_{[t],[N]}) > N(R_t + \varepsilon/4)\right]$$

can be made as small as we wish for all large N, by the law of large numbers. Thus, we can find a sufficiently large $N = N(\varepsilon)$ and a sequential N-code $\gamma = (\gamma_t)_{t\in[T]}$, such that

$$\max_{t\in[T]} \mathbf{P}_{\mu}\left[(X_{t,[N]}, \gamma_t(X_{[t],[N]})) \in A_{t,N}\right] \le \frac{\varepsilon}{4}.$$

Let $\mathbf{P}^{\gamma}_{\mu}$ be the joint probability law of $\boldsymbol{X} \sim \mu$ and the output of γ, cf. (3). Then, since all functions in $\mathcal{F}$ are uniformly bounded by 1, we can write

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^T \mathbb{E}^{\gamma}_{\mu}[\Delta_{t,N}] \\
&= \frac{1}{T}\sum_{t=1}^T \left(\mathbb{E}^{\gamma}_{\mu}[\Delta_{t,N}1\{A_{t,N}\}] + \mathbb{E}^{\gamma}_{\mu}[\Delta_{t,N}1\{A^c_{t,N}\}]\right) \\
&\le 2\max_{t\in[T]} \mathbf{P}^{\gamma}_{\mu}[A_{t,N}] + \max_{t\in T}\sqrt{\delta_{t,N}} \le \frac{\varepsilon}{2}.
\end{aligned} \tag{11}$$

Using (9), (11), and the triangle inequality, we get

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^T \mathbf{P}^{\gamma}_{\mu}\Big\|\mathsf{P}_{(X_{t,[N]},U_{t,[N]})} - \mu_t \otimes \pi_t\Big\|_{\mathcal{F}} \\
&\le \frac{1}{T}\sum_{t=1}^T \mathbf{P}^{\gamma}_{\mu}\Big\|\mathsf{P}_{(X_{t,[N]},U_{t,[N]})} - \mu_t \otimes \tilde{\pi}_t\Big\|_{\mathcal{F}} \\
&\quad + \frac{1}{T}\sum_{t=1}^T \|\mu_t \otimes \tilde{\pi}_t - \mu_t \otimes \pi_t\|_{\mathcal{F}} \\
&\le \Delta + \varepsilon.
\end{aligned}$$

Moreover, for $(\boldsymbol{X}, \boldsymbol{U}) \sim \mathbf{P}^{\gamma}_{\mu}$, we have

$$\begin{aligned}
\frac{H(\boldsymbol{U})}{NT} &\overset{\text{(a)}}{=} \frac{1}{NT}\sum_{t=1}^T H(U_{t,[N]}|U_{[t-1],[N]}) \\
&\overset{\text{(b)}}{\le} \frac{1}{NT}\sum_{t=1}^T \log(M_t + 1) \\
&\overset{\text{(c)}}{\le} \frac{1}{NT}\sum_{t=1}^T N\left(R_t + \frac{\varepsilon}{2}\right) \\
&= \frac{1}{T}\sum_{t=1}^T I_{\mu,\tilde{\pi}}(X^t;U_t|U^{t-1}) + \frac{\varepsilon}{2} \\
&\overset{\text{(d)}}{=} \frac{1}{T} I_{\mu,\tilde{\pi}}(X^T;U^T) + \frac{\varepsilon}{2} \le R_T(\Delta) + \frac{\varepsilon}{2}.
\end{aligned}$$

where (a) holds by the chain rule for entropy, (b) uses the fact that, for each realization of $U_{[t-1],[N]}$, $U_{t,[N]}$ can take at most $M_t + 1$ values by construction of γ, as well as the fact that the Shannon entropy of a random variable taking M values is upper-bounded by $\log M$; (c) is by the choice of the M_t's, (d) uses the fact that, since $\tilde{\pi}$ is a causal sequence of Markov kernels, U_t and $(X_{t+1}, \dots, X_T)$ are conditionally independent given (X^t, U^{t-1}) for each $t \in [T]$, and therefore

$$\begin{aligned}
\sum_{t=1}^T I_{\mu,\tilde{\pi}}(X^t;U_t|U^{t-1}) &= \sum_{t=1}^T I_{\mu,\tilde{\pi}}(X^T;U_t|U^{t-1}) \\
&= I_{\mu,\tilde{\pi}}(X^T;U^T),
\end{aligned}$$

by the chain rule for mutual information.

Thus, we have shown that, for any $\varepsilon > 0$, one can find a sufficiently large N and a sequential N-code $\gamma \in \Gamma^N_{\mu,\pi}(\Delta + \varepsilon)$, for which $\frac{H\left((\gamma_t(X_{[t],[N]}))_{t\in T}\right)}{NT} \le R_T(\Delta) + \varepsilon$. Therefore, $\widehat{R}_{T,N}(\Delta + \varepsilon) \le R_T(\Delta) + \varepsilon$, as claimed.

Converse. To prove the converse, we use the technique from [7]. Fix an arbitrary sequential N-code $\gamma \in \Gamma^N_{\mu,\pi}(\Delta)$, and let $(\boldsymbol{X}, \boldsymbol{U})$ be a realization of state-action processes drawn from the strategic measure $\mathbf{P}^\gamma_\mu$. Let J be a random variable uniformly distributed on $[N]$, independently of $(\boldsymbol{X}, \boldsymbol{U})$. Consider the random couple $(X_{[T],J}, U_{[T]}, J)$. From symmetry and independence, it follows that the marginal distribution of $X_{[T],J}$ is equal to μ. For each $t \in [T]$, let $\tilde{\pi}^{(t)} \in \mathcal{M}(\mathsf{U}|\mathsf{X}^t \times \mathsf{U}^{t-1})$ be the induced conditional law of $U_{t,J}$ given $(X_{[t],J}, U_{[t-1],J})$, and let $\tilde{\pi}_t \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ denote the induced conditional law of $U_{t,J}$ given $X_{t,J}$. Then we have the following chain of equalities and inequalities:

$$
\begin{aligned}
H\left((\gamma_t(X_{[t],[N]}))_{t\in[T]}\right) &= H(U_{[T],[N]}) \\
&\overset{\text{(a)}}{=} I(X_{[T],[N]}; U_{[T],[N]}) \\
&\overset{\text{(b)}}{\ge} \sum_{n=1}^{N} I(X_{[T],n}; U_{[T],n}) \\
&\overset{\text{(c)}}{=} N I(X_{[T],J}; U_{[T],J}|J) \\
&\overset{\text{(d)}}{=} N I(X_{[T],J}; U_{[T],J}, J) \\
&\ge N I_{\mu,\tilde{\pi}}(X^T; U^T),
\end{aligned}
$$

where (a) follows from the fact that the map $X_{[T],[N]} \mapsto U_{[T],[N]}$ is deterministic; (b) follows from the fact that $Y_1, \ldots, Y_N$ is a sequence of independent random variables and Lemma 8.8 in [13]; (c) follows from the construction of J; and (d) follows from the fact that, since $X_{[T],1}, \ldots, X_{[T],N}$ are i.i.d., J and $X_{[T],J}$ are independent (see [7, App. B]), and from chain rule for the mutual information. The remaining steps are consequences of definitions and of standard information-theoretic identities. Dividing both sides by NT, we obtain the bound

$$
\frac{H\left((\gamma_t(X_{[t],[N]}))_{t\in[T]}\right)}{NT} \ge \frac{I_{\mu,\tilde{\pi}}(X^T; U^T)}{T}.
$$

Now, for each $t \in [T]$, $X_{t,J}$ is independent of J, and has the same law as $X_{t,1}$, namely μ_t. Moreover (cf. App. B in [7]), the expected empirical distribution $\mathbb{E}^\gamma_\mu \mathsf{P}_{(X_{t,[N]}, U_{t,[N]})}$ is equal to $\mu_t \otimes \tilde{\pi}_t$. Then we have

$$
\begin{aligned}
&\sum_{t=1}^{T} \left\| \mu_t \otimes \tilde{\pi}_t - \mu_t \otimes \pi_t \right\|_{\mathcal{F}} \\
&= \sum_{t=1}^{T} \left\| \mathbb{E}^\gamma_\mu \mathsf{P}_{(X_{t,[N]}, U_{t,[N]})} - \mu_t \otimes \pi_t \right\|_{\mathcal{F}} \\
&\le \sum_{t=1}^{T} \mathbb{E}^\gamma_\mu \left\| \mathsf{P}_{(X_{t,[N]}, U_{t,[N]})} - \mu_t \otimes \pi_t \right\|_{\mathcal{F}} \le \Delta,
\end{aligned}
$$

where the first inequality is by convexity, while the second inequality is by assumption on γ. Therefore, $\tilde{\pi} = (\tilde{\pi}^{(t)})_{t\in[T]} \in \Pi_{\mu,\pi}(\Delta)$, and consequently

$$
\frac{H\left((\gamma_t(X_{[t],[N]}))_{t\in[T]}\right)}{NT} \ge \frac{I_{\mu,\tilde{\pi}}(X^T; U^T)}{T} \ge R_T(\Delta),
$$

by definition. Since this holds for every $\gamma \in \Pi^N_{\mu,\pi}(\Delta)$, it follows that $\widehat{R}_{T,N}(\Delta) \ge R_T(\Delta)$.

VI. Conclusion

We have formulated the problem of sequential empirical coordination under an output entropy constraint, where the fidelity of coordination is measured relative to a given universal Glivenko–Cantelli class of test functions. This pproblem is an extension of the empirical coordination framework of Raginsky [7] to the sequential setting introduced by Tatikonda [4]. The major departure of our set-up from sequential rate-distortion theory is that our operational criterion is not additive across the DM's; nevertheless, we have obtained a full information-theoretic characterization of the fundamental limits on the communication requirements for sequential empirical coordination. There are two major avenues for future work: (1) extending the results to allow for feedback control policies, and (2) obtaining tight and easily computable upper and lower bounds for sequential rate-distortion functions.

References

[1] T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1971.

[2] A. S. Matveev and A. V. Savkin, *Estimation and Control Over Communication Networks*, ser. Control Engineering. Birkhäuser, 2009.

[3] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization Under Information Constraints*, ser. Systems and Control: Foundations and Applications. Birkhäuser, 2013.

[4] S. Tatikonda, "Control under communication constraints," Ph.D. dissertation, MIT, Cambridge, MA, August 2000.

[5] T. Tanaka, P. Mohajerin Esfahani, and S. K. Mitter, "LQG control with minimal information: three-stage separation principle and SDP-based solution sythesis," 2015, arXiv preprint 1510.04214. [Online]. Available: http://arxiv.org/abs/1510.04214

[6] P. W. Cuff, H. H. Permuter, and T. M. Cover, "Coordination capacity," *IEEE Trans. Inform. Theory*, vol. 56, no. 9, pp. 4181–4206, September 2010.

[7] M. Raginsky, "Empirical processes, typical sequences, and coordinated actions in standard Borel spaces," *IEEE Trans. Inform. Theory*, vol. 59, no. 3, pp. 1288–1301, 2013.

[8] R. van Handel, "The universal Glivenko–Cantelli property," *Prob. Theory Rel. Fields*, vol. 155, pp. 911–934, 2013.

[9] D. L. Neuhoff and R. K. Gilbert, "Causal source codes," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 5, pp. 701–713, September 1982.

[10] K. R. Parthasarathy, *Probability Measures on Metric Spaces*. New York: Academic Press, 1967.

[11] R. M. Dudley, *Uniform Central Limit Theorems*. Cambridge Univ. Press, 1999.

[12] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*. San Francisco, CA: Holden-Day, 1964.

[13] R. M. Gray, *Entropy and Information Theory*, 2nd ed. Springer, 2011.

[14] O. Kallenberg, *Foundations of Modern Probability*, 2nd ed. Springer, 2002.