Estimating the state of a Markov chain over a noisy communication channel: A bound and an encoder

M. Anand and P. R. Kumar
Dept. of ECE and CSL, Univ. of Illinois, Urbana, IL 61801, USA
Email: {amurali2, prkumar}@illinois.edu

***Abstract*—We consider the problem of estimating the state of a Markov chain after N time units, when observed over a noisy communication channel. Specifically, a Markov chain is observed by an encoder. The encoder communicates with a decoder over the noisy communication channel. The past channel outputs are available causally to the encoder. The objective of the encoder is to maximize the mutual information between the state of the Markov chain after N time units, and the vector of channel outputs for N time units.**

We show that an outer bound on the reward under any encoding policy is N times the information-theoretic capacity of the noisy channel.

We show that the optimal encoding scheme is a function of the current state of the Markov chain, and the a-posteriori distribution of the current state given all the past channel outputs. We describe a simple encoding scheme called *posterior matching*, which has desirable properties.

I. Introduction

In networked control, the sensor and actuator may not be located in the same place. Moreover, the sensor will have to communicate its observations to the actuator or detector over a lossy or a noisy unrealiable communication network. Any packet that is transmitted over a network is subject to corruptions and drops, both of which complicate the role of the decoder. Though information theory provides a fundamental bound on the rate at which reliable communication is possible on an unreliable channel, the bound is achieved by collecting incoming data into long blocks, and later encoding them into long codewords. Clearly this approach is not tenable for a networked control system because of the delay involved in accumulating large blocks before processing them. In a control system, the state of a plant must be communicated to the controller or the estimator in a timely manner, else its use is limited. Also, the communication channel must be used every time instant. It is not possible to accumulate channel bandwidth by not transmitting for a few time units, and later using the channel multiple times on a single day. For all these reasons, the problem of state estimation over a noisy channel which is used at every time instant is a problem of much interest. It is a problem that lies at the confluence of the fields of control and information theory.

A. Control over networks

Real-time estimation and control problems have been addressed in [1] and [2], where the authors consider a plant whose states form a Markov chain or a controlled Markov chain, depending on whether it is an estimation or a control problem that is of interest. The state is observed by a sensor, which communicates it over a noisy channel to a decoder/controller. The encoder (sensor) also is assumed to have knowledge of the past channel outputs, which are causally communicated to it over a noiseless feedback channel. For the detection problem, the authors considered a traditional cost structure, where a cost $c_n(x_n, d_n)$ is paid for a decision d_n made on the n-th day, when the state is x_n. A similar cost is applicable for the control problem, where now d_n can be regarded as the control applied on time unit n. The aim is to design optimal encoding-decoding schemes to minimize the expected total cost over a span of N time units. The authors proved that a separation based architecture for the decoder/controller is optimal, where a state estimator computes the a-posteriori distribution of the current state based on all the channel outputs, and the optimal decision or control is derived based on the posterior. The authors proved that an optimal encoder is a function of the a-posteriori distribution of the previous state given all the channel observations. However, no explicit encoding scheme for general channels and arbitrary cost functions is proposed due to the apparent need to solve the dynamic programming equations for a value function defined over the space of probability density functions.

A more general approach is taken in [3], which considers control of a continuous state space Markov chain over a finite rate channel or a noisy channel. The encoder has access to the state of the decoder. Policies for minimizing an expected cost subject to an average entropy constraint on the rate are considered, and it is shown that a randomized stationary Markov policy is optimal.

References [4]–[7] analyze the stabilizability and observability of linear systems. Reference [4] considers a linear time-invariant system that is observed and controlled over a noiseless channel of finite rate. Lower bounds on the channel rates to achieve asymptotic stabilizability and asymptotic observability of the linear systems are studied. For certain information patterns, coding schemes are provided which achieve the bounds on performance. The finite rate link is replaced with a noisy channel in [5], and information-theoretic tools, specifically rate-distortion theory, are used to compute bounds on the capacity of the noisy channel to guarantee asymptotic stabilizability and observability. Reference [6] studies the classical linear quadratic Gaussian (LQG) problem with a noisy channel connecting the sensor and the controller. In order to

This material is based upon work partially supported by NSF CNS-0905397, and AFOSR under Contract FA9550-09-0121.

study the effect of delay on the squared error distortion, tools from sequential rate distortion theory are used, and it is proved that the LQG cost can be decomposed as the sum of full knowledge cost and partial knowledge cost. In [7], feedback anytime capacity above a threshold is shown to be necessary and sufficient for stabilizing a linear stochastic system.

There has been much research on analyzing the performance (mainly asymptotic stability) of stochastic linear systems over a network, and the survey article [8] is a good reference.

B. Feedback information theory

A classic result in information theory is that feedback does not improve the capacity of a noisy channel [16]. However, this does not mean that feedback is not useful in communicating information over noisy channels. Feedback helps to reduce the complexity of encoding schemes, as well as to reduce the probability of error in decoding the source's transmission.

Consider a plant whose state is a number randomly and uniformly chosen in the interval $[0,1)$. Once chosen, the state is fixed for all time. The encoder's aim is to convey the state to the decoder. Horstein [9] proposed a capacity achieving scheme for the binary symmetric channel (BSC), where at each time instant the encoder transmits a bit indicating if the message is to the left or to the right of the median of the receiver's a-posteriori distribution of the message. Later, Schalkwijk and Kailath[10] introduced a coding scheme for the additive white Gaussian noise (AWGN) channel, where the encoder transmits the mean squared error of the decoder's estimate every time, with the appropriate scaling to satisfy the average power constraint for the channel. This scheme achieves the capacity of the AWGN channel and the error probability decreases at a doubly exponential rate in the duration of transmission. The connection between these two schemes was suspected, but not proven for a while. It was also not clear how to generalize these schemes to general DMCs.

The above communication-with-feedback problem was recently solved in [11] and [12]. In [11], a new encoding scheme called *posterior matching* is proposed. Here, at each time instant, the encoder attempts to send independent information about the message to the decoder. Also, the encoder ensures that the transmitted symbol is distributed according to the capacity-achieving distribution of the channel. For the special cases of the BSC and the AWGN channel, the posterior matching scheme reduces to the scheme in [9] and [10] respectively. In [12], it is shown that the posterior matching scheme achieves the capacity of any DMC. In [13], an alternate and simpler proof of the optimality of the posterior matching scheme is given by recasting it as a stochastic control problem.

C. Our results

The above-mentioned references suggest an approach to networked control systems that is different from that of traditional control systems or traditional communication systems, yet incorporates ideas from both. In many networked control systems, some form of feedback is available to the encoder from the decoder. We begin by studying networked control systems under the assumption of perfect causal feedback of channel outputs at the encoder. However as we will later see, optimal encoding schemes do not require complete feedback, but only need the decoder to pass on a sufficient statistic of the channel outputs to the encoder, as a function of which the encoder can choose an encoding scheme. We address a problem in this field by using information-theoretic techniques to upper bound the reward (negative of the cost) in an estimation problem, and prove the optimality of a certain information pattern for the encoder, using tools from stochastic control theory. We also design an encoding scheme that is derived from the above-mentioned posterior matching scheme.

We consider the problem of estimating the state of a Markov chain $\{x_n\}$ at the end of a finite horizon of N time units. A sensor that observes the state of the Markov chain communicates with an estimator over a noisy communication channel; see Fig. 1. From a communication perspective, the sensor is the encoder and the estimator is the decoder. The sensor is also assumed to have access to the past channel outputs through a noiseless feedback channel. Note that by "feedback" we mean that the encoder has access to what the decoder sees. (Such feedback is different from the feedback for the control problem that connects the output of the plant to the input of the plant). Unlike the traditional reward function in optimum control theory that is a function of the state and the decision, we analyze encoding schemes that aim to maximize the terminal reward

$$I(x_N; y_1, y_2, \ldots, y_N),$$

which is the mutual information (see [16]) between the state x_N of the Markox chain after N time units, and the symbols received by the decoder $(y_1, y_2, \ldots, y_N)$ over a span of N time units. We develop the model further in Section II. The problem considered here is a variant of the detection problem considered in [1]. There are crucial differences though; the decoder in this case does not make a decision every day.

In Section III, we derive a simple and intuitive upper bound on the highest reward that can be achieved by any encoding scheme. This upper bound is a function of the information-theoretic channel capacity, and is derived using standard information-theoretic inequalities.

We extensively analyze the encoder in Section IV. We prove that an optimal encoder is a function of the current state, and the a-posteriori distribution of the current state of the plant given the past outputs. We also outline filtering equations that permit recursive computation of the posterior distribution at each time instant. The techniques used in proving these properties of the encoder are similar to those in [1], though they are substantially more involved due to the need to keep track of two posterior distributions at each stage. This leads to more bookkeeping while proving optimality of information patterns using dynamic programming.

Finally, we outline a simple encoding scheme, *posterior matching*, with desirable properties in Section V. We borrow the name from the encoding scheme proposed in [11], [12], since the encoding scheme used here is a variant of that in [11], with the crucial distinction lying in its use in tracking a Markov chain, instead of its use in improving the estimate of the message in [11].

The reward we consider here is the mutual information between two random variables. The mutual information is invariant to a change in the labeling of the alphabet of state space and has a more universal appeal over, say, a specific average reward function. For systems with an unknown reward function, the universality of the mutual information as a measure of the performance of the system, and the universality of posterior matching, could perhaps be exploited to design systems with provable guarantees on performance.

II. Model

A. Plant

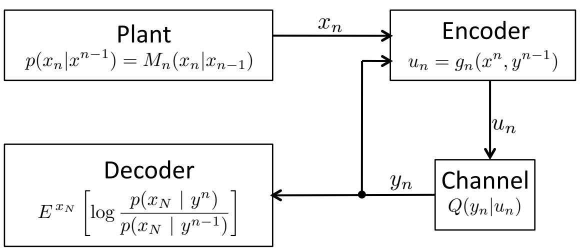

Fig. 1. Tracking the state of a Markov chain.

We consider tracking the state of a plant over a noisy network, as shown in Fig. 1. The state of the plant at time unit n is indexed by the random variable x_n, with $x_n \in \mathcal{X}$, for all n. $\{x_n\}$ forms a Markov process with

$$p(x_n \mid x^{n-1}) \;=\; M_n(x_n \mid x_{n-1}), \;\; n \geq 2. \qquad (1)$$

M_n is the transition probability matrix of the Markov chain on time unit n. The initial distribution $p(x_1)$ of the Markov chain is known to both the encoder and the decoder.

B. Encoder

The encoder communicates an encoded version of all its causally available information to a decoder over a noisy channel. Let $u_n \in \mathcal{U}$ and $y_n \in \mathcal{Y}$ respectively be the input and the output of the channel on n-th time unit. We assume that the encoder has access to all the past states x^n, and the past channel outputs y^{n-1} (perhaps communicated to the encoder over a feedback channel). That is, we are assuming here that channel feedback is available to the encoder. Later one we will see how to dispense with this assumption. Hence the information pattern [14] of the encoder at time n is $\{x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_{n-1}\}$. The channel input u_n is given by the function

$$u_n \;=\; g_n(x_n, x^{n-1}, y^{n-1}). \qquad (2)$$

The encoding function g_n is unrestricted, and in particular can vary with time.

C. Decoder

The output of the channel is given by y_n, where

$$p(y_n \mid x^n, y^{n-1}, u^n) \;=\; Q(y_n \mid u_n).$$

$Q(\cdot|\cdot)$ is the transition probability matrix of the channel. The information available to the decoder at time n is $\{y_1, y_2, \ldots, y_n\}$. At the end of time unit n, the decoder computes the reward function

$$c_n(y^n) \;=\; E^{x_N}\left[\log \frac{p(x_N \mid y^n)}{p(x_N \mid y^{n-1})}\right].$$

In the above expression, y^n is fixed according to what the decoder has received. The superscript indicates that we compute the conditional expectation over all values of $x_N \in \mathcal{X}$. We use this notation throughout the paper.

At the end of N time units, the total reward computed by the decoder is

$$c(y^N) \;=\; \sum_{n=1}^{N} c_n(y^n) \;=\; E^{x_N}\left[\log \frac{p(x_N \mid y^N)}{p(x_N)}\right].$$

D. Objective

We study a finite horizon problem where we want to track the state of the Markov chain at the end of N time units. The objective is to design an encoder $\{g_n(.)\}$ to maximize the expected reward

$$\begin{aligned} J(g) &= E^{y^N}[c(y^N)] \;=\; E^{y^N}\left[E^{x_N}\left[\log \frac{p(x_N \mid y^N)}{p(x_N)}\right]\right] \\ &= I(x_N; y_1, y_2, \ldots, y_N). \end{aligned} \qquad (3)$$

Hence, the encoder's aim is to maximize the mutual information between the state of the plant at time N and the set of all the signals received by the decoder. An optimal encoder has the expected reward (mutual information) J^* with $J^* \geq J(g)$ for any other choice of the encoding policy $\{g_n, n = 0, 1, \ldots, N-1\}$.

III. Upper bound on reward

In this section, we derive an upper bound on the optimal reward J^*. The bound is based on simple information-theoretic arguments.

Lemma 3.1: The optimal reward is bounded from above as

$$J^* \;\leq\; N\mathcal{C}, \qquad (4)$$

where $\mathcal{C}$ is the capacity of the discrete-time memoryless channel described by the transition probability matrix $Q(\cdot|\cdot)$.

Proof: Consider any encoding policy $\{g_n\}$. Then, from (3), the expected reward is given by

$$\begin{aligned} J(g) &= I(x_N; y_1, y_2, \ldots, y_N) & (5) \\ &= \sum_{n=1}^{N} I(x_N; y_n | y^{n-1}) & (6) \\ &\leq \sum_{n=1}^{N} I(x_N, x_n; y_n | y^{n-1}), & (7) \end{aligned}$$

where the above bound follows by a simple application of the chain rule and the non-negativity of mutual information (see

[16]). Further, by an application of the chain rule on each of the mutual information terms, we get

$$J(g) \leq \sum_{n=1}^{N} (I(x_n; y_n|y^{n-1}) + I(x_N; y_n|y^{n-1}, x_n)). \quad (8)$$

Now we separately bound both the mutual information terms appearing for each n. The first term can be bounded as

$$\begin{aligned} & I(x_n; y_n|y^{n-1}) \\ & \quad = \quad H(y^n|y^{n-1}) - H(y^n|y^{n-1}, x_n) \\ & \quad = \quad H(y^n|y^{n-1}) - H(y^n|y^{n-1}, x_n, u_n), \end{aligned}$$

where the above follows from $u_n = g_n(x_n, y^{n-1})$. Using the properties of of conditional entropy (see [16]), $H(y^n|y^{n-1})$ can be upper bounded by $H(y^n)$. Hence we get

$$\begin{aligned} I(x_n; y_n|y^{n-1}) \quad & \leq \quad H(y_n) - H(y_n|y^{n-1}, x_n, u_n) \quad (9) \\ & = \quad H(y_n) - H(y_n|u_n), \end{aligned}$$

where we used the fact that $(x_n, y^{n-1}) \to u_n \to y_n$ forms a Markov chain (see [16]). Therefore,

$$I(x_n; y_n|y^{n-1}) \quad \leq \quad I(u_n; y_n) \ \leq \ \mathcal{C}, \quad (10)$$

where $\mathcal{C}$ is the capacity of the channel described by the transition probability matrix $Q(\cdot|\cdot)$. Now, we expand the second mutual information term from (8) to get

$$\begin{aligned} & I(x_N; y_n|y^{n-1}, x_n) \\ & \quad = \quad H(x_N|y^{n-1}, x_n) - H(x_N|y^{n-1}, x_n, y_n) \\ & \quad = \quad H(x_N|x_n) - H(x_N|x_n) \ = \ 0. \quad (11) \end{aligned}$$

The last step uses the fact that $y^{n-1} \to x_n \to x_N$ and $y^n \to x_n \to x_N$ are Markov chains.

Substituting (10) and (11) into (8), we get the statement of the lemma. ■

IV. Analysis of the encoder

Next, we describe the architecture of the optimal encoder. As mentioned in (2), the information pattern of the encoder is (x^n, y^{n-1}). The dimension of the information pattern available to the encoder is growing with time, which may suggest optimal encoding schemes of increasing complexity with time. However, in the following theorem, we prove that the encoder's information pattern can be simplified to just $(x_n, \xi_n(y^{n-1}))$, where $\xi_n(y^{n-1})$ is the a-posteriori distribution of x_n given y^{n-1}, i.e.,

$$\xi_n(y^{n-1})(x_n) \ := \ p(x_n|y^{n-1}).$$

Theorem 4.1: The optimal encoder $g_n^*(\cdot)$ is of the form

$$g_n^*(x^n, y^{n-1}) \ = \ \nu_n(x_n, \xi_n(y^{n-1})). \quad (12)$$

We prove the theorem in a series of steps:

- First we show that the information pattern of the encoder can be simplified to just (x_n, y^{n-1}), i.e., the encoder can ignore the past states.
- Then, we prove that $\xi_n(y^{n-1})$ can be computed recursively as a function of $\xi_n(y^{n-1})$, y_n, and the coding scheme for time n.
- Finally, via dynamic programming we prove that the optimal policy at time n is a function of $\xi_n(y^{n-1})$.

A. Simplifying the encoder's information pattern by ignoring the past states

Next, we prove the optimality of an encoder that generates the channel input only as a function of the previous received signals and the current state. Hence the encoder can ignore the previous states x^{n-1}.

Lemma 4.2: The optimal encoder $g_n^*(\cdot)$ is of the form

$$g_n^*(x^n, y^{n-1}) = \gamma_n(x_n, y^{n-1}), \ \ n = 1, 2, \ldots, N. \quad (13)$$

Proof: The encoder helps the decoder in maximizing the reward. We prove that the encoder can be equivalently seen as a *controller* for an alternate Markov chain.

Define $z_n := (x_n, y^{n-1}), n = 1, 2, \ldots, N$. Then z_n describes a Markov process that is controlled by the channel input u_n:

$$\begin{aligned} & p(z_{n+1} \mid z^n, u^n) \\ & \quad = \quad p(y^n \mid x^n, y^{n-1}, u^n) p(x_{n+1} \mid x^n, y^n, u^n) \\ & \quad = \quad p(y^n \mid x_n, y^{n-1}, u_n) p(x_{n+1} \mid x_n, y^n, u_n) \\ & \quad = \quad p(x_{n+1}, y^n \mid x_n, y^{n-1}, u_n) \\ & \quad = \quad p(z_{n+1} \mid z_n, u_n). \end{aligned}$$

Now fix an encoder for the model in Section II. The expected reward $J(g)$ is given by

$$\begin{aligned} & J(g) \\ & \quad = \quad E\left[\sum_{n=1}^{N} \log \frac{p(x_N \mid y^n)}{p(x_N \mid y^{n-1})}\right] \\ & \quad \stackrel{(i)}{=} \quad E\left[\sum_{n=1}^{N} E\left[\log \frac{p(x_N \mid y^n)}{p(x_N \mid y^{n-1})} \;\middle|\; x^n, y^{n-1}, u^n\right]\right] \\ & \quad \stackrel{(ii)}{=} \quad E\left[\sum_{n=1}^{N} E\left[\log \frac{p(x_N \mid y^n)}{p(x_N \mid y^{n-1})} \;\middle|\; x_n, y^{n-1}, u_n\right]\right] \\ & \quad = \quad \mathbb{E}[\sum_{n=1}^{N} k_n(z_n, u_n)], \quad (14) \end{aligned}$$

where (i) uses iterated conditioning, and (ii) follows from observing that

$$p(x_N, y_n|x^n, y^{n-1}, u^n) \ = \ p(x_N, y_n|x_n, y^{n-1}, u_n).$$

Note that k_n is a function that depends on the encoding scheme and the reward function. But (14) suggests that, for a given encoding scheme, $J(g)$ is maximized by optimally choosing u_n to control $\{z_n\}$. The optimal control law in this case is of the form (see [15])

$$u_n \ = \ \gamma_n(z_n), \quad n = 1, 2, \ldots, N.$$

■

B. Filtering equations

Next we describe the filtering equations for recursively computing $\xi_{n+1}(y^n)$ as a function of $\xi_n(y^{n-1})$, y_n, $\gamma_n(\cdot, y^{n-1})$. Let $\mathcal{P}(\mathcal{X})$ be the space of probability distributions on $\mathcal{X}$.

Lemma 4.3: (Filtering equations) Choose an encoding scheme $\gamma_n(\cdot, y^{n-1})$ for all n. The a-posteriori distribution $\xi_n(y^{n-1})$ can be recursively computed by the rule

$$\xi_1(y^0) = p(x_1) \quad \text{(Initial distribution)} \tag{15}$$

$$\xi_n(y^{n-1}) = G_n(\xi_{n-1}(y^{n-2}), y_{n-1}, \gamma_{n-1}(\cdot, y^{n-2})), \quad \text{for all } n \geq 2. \tag{16}$$

where

$$G_n(\xi, y, \gamma)(x) = \frac{\sum_{x' \in \mathcal{X}} \xi(x')Q(y|\gamma(x))M_n(x|x')}{\sum_{\tilde{x} \in \mathcal{X}} \sum_{x' \in \mathcal{X}} \xi(x')Q(y|\gamma(\tilde{x}))M_n(\tilde{x}|x')}, \tag{17}$$

for any $\xi \in \mathcal{P}(\mathcal{X})$.

Proof: The proof follows from a simple application of Bayes' rule and we omit it due to lack of space. ■

C. Proof of Theorem 4.1

We use the results of Lemmas 4.2 and 4.3, along with dynamic programming to prove Theorem 4.1. For the setup in Sec. II, we formulate a dynamic programming to show that the optimal encoder attempts to minimize an average reward-to-go function on each day. The solution to this optimization problem will require the encoder to choose an encoding scheme that is solely a function of the posterior distribution of the state.

Choose an encoding policy $g \equiv \{g_n\}$, and define the expected reward-to-go function from time unit n as

$$J_n(g, y^n) = E\left[\sum_{k=n}^{N} c_k(y^k) \,\middle|\, y^n\right].$$

As is evident, $J_n(g, y^n)$ is the average total reward incurred when we start tallying from time n, and have knowledge of y^n.

Pick any two distributions ζ_1 and ζ_2 in $\mathcal{P}(\mathcal{X})$, and define a value function that is computed recursively as

$$V_N(\zeta_1, \zeta_2) = E\left[\log \frac{p(x_N|y^N)}{p(x_N|y^{N-1})} \middle| \xi_N = \zeta_1, p(x_N|y^N) = \zeta_2\right], \tag{18}$$

$$V_{N-1}(\zeta_1, \zeta_2) = \max_{\gamma_N:\ x_N \to y_N} E\left[\log \frac{p(x_N|y^{N-1})}{p(x_N|y^{N-2})} + V_N(\zeta_2, G(\zeta_2, y_N, \gamma_N)) \middle| \xi_{N-1} = \zeta_1, \xi_N = \zeta_2\right], \tag{19}$$

with

$$G(\xi, y, \gamma)(x) = \frac{\xi(x)Q(y|\gamma(x))}{\sum_{\tilde{x} \in \mathcal{X}} \xi(x)Q(y|\gamma(x))}, \tag{20}$$

$$V_n(\zeta_1, \zeta_2) = \max_{\gamma:\ x_{n+1} \to y_{n+1}} E\left[\log \frac{p(x_N|y^n)}{p(x_N|y^{n-1})} + V_{n+1}(\zeta_2, G_{n+2}(\zeta_2, y_{n+1}, \gamma)) \middle| \xi_n = \zeta_1, \xi_{n+1} = \zeta_2\right], \quad \text{for } n = N-2, N-3, \ldots, 1, \tag{21}$$

and

$$V_0(p(x_1)) := \max_{\gamma:\ x_1 \to y_1} E\left[V_1(p(x_1), G_2(p(x_1), y_1, \gamma)) \middle| \xi_1 = p(x_1)\right]. \tag{22}$$

The definition of the value function is somewhat involved since we need to keep track of the pair of distributions $(p(x_n|y^{n-1}), p(x_{n+1}|y^n))$ on each day.

The next goal is to show that $V_n(\xi_n)$ is an upper bound on the reward-to-go function for time unit n.

Lemma 4.4:

$$V_N(\xi_N(y^{N-1}), p(\cdot|y^N)) \geq J_N(g, y^N), \tag{23}$$

$$V_n(\xi_n(y^{n-1}), \xi_{n+1}(y^n)) \geq J_n(g, y^n), \quad n = N-1, \ldots, 1. \tag{24}$$

Proof: Equality in (23) follows by comparing the respective definitions. (24) is easy to show for $n = N-1$ using techniques similar to the proof outlined below, and we skip it to avoid repetition. By induction, we assume that the lemma holds for all $n \geq m$, for some $m \leq N$.

Then,

$$\begin{aligned}
&J_{m-1}(g, y^{m-1}) \\
&= E\left[\sum_{k=m-1}^{N} \log \frac{p(x_N \mid y^k)}{p(x_N \mid y^{k-1})} \,\middle|\, y^{m-1}\right] \\
&= E\left[\log \frac{p(x_N \mid y^{m-1})}{p(x_N \mid y^{m-2})} + E\left[\sum_{k=m}^{N} \log \frac{p(x_N \mid y^k)}{p(x_N \mid y^{k-1})} \,\middle|\, y^m\right] \,\middle|\, y^{m-1}\right] \\
&\leq E\left[\log \frac{p(x_N \mid y^{m-1})}{p(x_N \mid y^{m-2})} + V_m(\xi_m(y^{m-1}), \xi_{m+1}(y^m)) \,\middle|\, y^{m-1}\right] &(25) \\
&= E\left[E\left[\log \frac{p(x_N \mid y^{m-1})}{p(x_N \mid y^{m-2})} + V_m(\xi_m(y^{m-1}), \xi_{m+1}(y^m)) \middle| y^{m-1}, \xi_m, \xi_{m-1}\right] \middle| y^{m-1}\right] \\
&= E\left[E\left[\log \frac{p(x_N \mid y^{m-1})}{p(x_N \mid y^{m-2})} + V_m(\xi_m(y^{m-1}), \xi_{m+1}(y^m)) \middle| \xi_m, \xi_{m-1}\right] \middle| y^{m-1}\right] &(26) \\
&\leq E\left[V_{m-1}(\xi_{m-1}(y^{m-2}), \xi_m(y^{m-1})) \,\middle|\, y^{m-1}\right] &(27) \\
&= V_{m-1}(\xi_{m-1}(y^{m-2}), \xi_m(y^{m-1})). &(28)
\end{aligned}$$

In the above, (25) follows by applying induction. (26) is a crucial step where we used Lemmas 4.2 and 4.3 to remove conditioning on y^{m-1} in the inner conditional expectation. It is easy to verify (26) by rewriting the expanded versions of the conditional expectations. (27) follows from the definition of V_{m-1}, and the last step (28) is obvious.

Hence (24) is true for $n = N-1, N-2, \ldots, 1$, and the proof for (22) is similar and simpler. The lemma follows. ■

We observe that for an encoding policy g, $V_0(p(x_1)) \geq J(g)$. The optimal policy g^* achieves equality in the above and is obtained by solving the dynamic programming equations (18)–(22) recursively. At time instant n, from Lemma 4.4, the optimal policy is only a function of the posterior $\xi_n(y^{n-1})(x_n) = p(x_n|y^{n-1})$, and Theorem 4.1 follows.

V. An encoding scheme: Posterior matching

In this Section, we present a encoding scheme with desirable properties. From Theorem 4.1, we can narrow down our search to schemes that are only a function of the a-posteriori distribution of x_n given all the previous channel outputs.

Let F_x describe the cumulative distribution function (cdf) of any random variable x. Let F_x^{-1} describe the corresponding inverse function, i.e.,

$$F_x^{-1}(y) \ := \ \inf \{\tilde{x} : F_x(\tilde{x}) > y\}.$$

In the above, for ease of exposition, we have assumed that the random variable x is real.

Pick an input distribution $p(u)$ for the channel that achieves the channel capacity (see [16] for the definition), and let the corresponding cdf be F_u. We use the same input distribution for constructing the channel inputs on all time units.

We obtain ideas on how to construct an encoding scheme from the proof in Section III. The upper bound in (9) suggests that the channel output y_n at time instant n must be independent of y^{n-1}. Also, the steps following (9) suggest that the induced distribution on the channel input u_n must be the capacity achieving distribution. With these insights, we define the *posterior matching* encoding scheme.

At time instant n, the encoder transmits the function

$$u_n \ = \ F_u^{-1} F_{x_n|y^{n-1}}(x_n|y^{n-1}).$$

$F_{x_n|y^{n-1}}$ is the a-posteriori cdf corresponding to the distribution $\xi_n(y^{n-1})$. For a fixed y^{n-1}, $F_{x_n|y^{n-1}}(x_n|y^{n-1})$ is a random variable that is uniformly distributed on $[0,1]$. Hence it is independent of y^{n-1}. Therefore the input u_n is independent of y^{n-1}, due to which y_n is also independent of y^{n-1}. Then, the inverse cdf function F_u^{-1} applied to a uniformly distributed random variable transforms it into a random variable with the capacity-achieving distribution $p(u)$. Hence the channel input u_n has the chosen capacity achieving distribution. Posterior matching has some of the properties that we required of an optimal scheme. In fact, from the above arguments, the inequalities in (9) and (10) are met with equality under posterior matching and the mutual information term $I(x_n; y_n|y^{n-1})$ is maximized to equal $\mathcal{C}$. Hence,

$$\sum_{k=1}^{N} I(x_n; y_n|y^{n-1}) \ = \ N\mathcal{C}.$$

However, this does not imply that $I(x_N; y^N)$ is maximized under posterior matching and it is not clear if this encoding scheme is optimal.

VI. Conclusions

We have used ideas from feedback information theory to analyze an estimation problem that is motivated by networked control systems. We have considered the design of an encoding policy to maximize the mutual information $I(x_N; y^N)$ between the state x_N of a Markov chain after N time units and the set of signals $(y_1, \ldots, y_N)$ received by the destination over N time units at the output of a noisy communication link. We have proved an upper bound of $N\mathcal{C}$ on the reward, where $\mathcal{C}$ is the information-theoretic capacity of the noisy channel connecting the encoder and the destination. We have proved that the optimal channel input at time n is only a function of the current state x_n and the a-posteriori distribution of x_n given all the channel outputs. Also, we have described an encoding scheme called *posterior matching* with desirable properties.

References

[1] J. Walrand and P. Varaiya, "Optimal causal coding-decoding problems, *IEEE Transactions on Information Theory*, vol. 29, no. 6, pp. 814–820, Nov. 1983.

[2] ———, "Causal coding and control of markov chains, System and Control Letters", vol. 3, pp. 189192, 1983.

[3] V. Borkar, S. Mitter, and S. Tatikonda, "Markov control problems under communication constraints," *Communication in Information and Systems*, vol. 1, no. 1, pp/ 15–32, Jan. 2001.

[4] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1056–1068, July 2004.

[5] S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1196-1201, July 2004

[6] S. Tatikonda, A. Sahai, and S. Mitter, "Stochastic linear control over a communication channel," *IEEE Transactions on Information Theory*, vol. 49, no.9, pp. 1549–1560, Sept. 2004.

[7] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link – Part I: Scalar systems," *IEEE Transactions on Information Theory*, vol. 52, no.8, pp. 3369–3395, Aug. 2006.

[8] J. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, Jan. 2007.

[9] M. Horstein, "Sequential transmission using noiseless feedback," *IEEE Trans. Information Theory*, vol. 9, no. 3, pp. 136–143, July 1963.

[10] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback-I: No bandwidth constraint," *IEEE Transactions on Information Theory*, vol.12, no. 2, pp. 172–182, Apr. 1966.

[11] O. Shayevitz and M. Feder, "Communication with feedback via posterior matching," *IEEE International Symposium on Information Theory*, Nice, France, June 2007.

[12] ———, "The posterior matching feedback scheme: capacity achieving and error analysis," *IEEE International Symposium on Information Theory*, Toronto, Canada, July 2008.

[13] T. P. Coleman, "A Stochastic Control Viewpoint on Posterior Matching-style Feedback Communication Schemes," *IEEE International Symposium on Information Theory*, Seoul, Korea, June 2009.

[14] H. S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proceedings of the IEEE*, vol. 59, no. 11, pp. 1557–1566, Nov. 1971.

[15] P. R. Kumar and P. Varaiya, *Stochastic systems: estimation, identification, and adaptive control*. Prentice Hall, Inc., Englewood Cliffs, NJ, 1986.

[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.