

# Optimal Design of Sequential Real-Time Communication Systems

Aditya Mahajan, *Member, IEEE*, and Demosthenis Teneketzis, *Fellow, IEEE*

**Abstract**—Optimal design of sequential real-time communication of a Markov source over a noisy channel is investigated. In such a system, the delay between the source output and its reconstruction at the receiver should equal a fixed prespecified amount. An optimal communication strategy must minimize the total expected symbol-by-symbol distortion between the source output and its reconstruction. Design techniques or performance bounds for such real-time communication systems are unknown. In this paper a systematic methodology, based on the concepts of information structures and information states, to search for an optimal real-time communication strategy is presented. This methodology trades off complexity in communication length (linear in contrast to doubly exponential) with complexity in alphabet sizes (doubly exponential in contrast to exponential). As the communication length is usually order of magnitudes bigger than the alphabet sizes, the proposed methodology simplifies the search for an optimal communication strategy. In spite of this simplification, the resultant optimality equations cannot be solved efficiently using existing algorithmic techniques. The main idea is to formulate a zero-delay communication problem as a dynamic team with nonclassical information structure. Then, an appropriate choice of information states converts the dynamic team problem into a centralized stochastic control problem in function space. Thereafter, Markov decision theory is used to derive nested optimality equations for choosing an optimal design. For infinite horizon problems, these optimality equations give rise to a fixed point functional equation. Communication systems with fixed finite delay constraint, a higher-order Markov source, and channels with memory are treated in the same manner after an appropriate expansion of the state space. Thus, this paper presents a comprehensive methodology to study different variations of real-time communication.

**Index Terms**—Dynamic teams, information states, joint source-channel coding, nonclassical information structures, real-time communication, zero-delay communication.

Manuscript received January 23, 2006; revised January 29, 2009. Current version published October 21, 2009. This work was supported in part by ONR under Grant N00014-03-1-0232, NSF under Grant CCR-0325571, and NASA under Grant NNX06AD47G. The material in this paper was presented in part at the 43rd Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, September, 2005.

A. Mahajan was with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA. He is now with the Department of Electrical Engineering, Yale University, New Haven, CT 06520 USA (e-mail: aditya.mahajan@yale.edu).

D. Teneketzis is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: teneket@eecs.umich.edu).

Communicated by M. Effros, Associate Editor for Source Coding.  
Digital Object Identifier 10.1109/TIT.2009.2030462

## I. INTRODUCTION

### A. Motivation

**I**N many controlled informationally decentralized systems information must be transmitted within bounded delay. Examples of such systems include networks with quality of service (QoS) requirements (e.g., bounded end-to-end delay), distributed routing in wired and wireless networks, decentralized detection in sensor networks, traffic flow control in transportation networks, resource allocation and consensus in partially synchronous systems, and decentralized resource allocation problems in economic systems. To understand how to design such systems it is necessary to understand how to communicate information with a hard deadline on communication delay, i.e., understand real-time communication of information.

In this paper we consider the simplest instance of a real-time communication system: a point-to-point real-time communication system shown in Fig. 1. The simplest model consists a dynamic first-order Markov source that is to be sequentially transmitted over a discrete memoryless channel to a receiver. Encoding and decoding must be done with zero-delay to minimize either a total expected distortion over a finite horizon or a total expected discounted distortion over an infinite horizon or a total expected average distortion per unit time over an infinite horizon. We later consider the models when communication must take place with a fixed finite delay, the source dynamics are higher order Markov, and the channel has finite state memory.

### B. Literature Overview

The real-time constraint on information transmission makes the real-time communication problem drastically different from the classical information theoretic formulation [1] which has no delay constraint. Information theory is an asymptotic theory; the fundamental concepts of information theory like source entropy and channel capacity are asymptotic concepts; the performance bounds that it provides are only tight for asymptotically large values of delay. Real-time communication is not asymptotic. Hence, the basic concepts and results from information theory are not appropriate for real-time communication. Nevertheless, many variations of the real-time communication problem have been studied in the literature.

Real-time source coding using finite memory encoders and decoders was studied in [2] and [3]. Real-time source coding of individual sequences was considered in [4]–[6]; these techniques were extended to zero-delay communication of individual sequences over noisy channels in [7]. Causal source coding, a weaker constraint than real-time communication, was investigated in [8]–[11]. Properties of optimal systems for real-time communication over noisy channels were obtained in [12]–[15]. Bounds on performance of communication systems

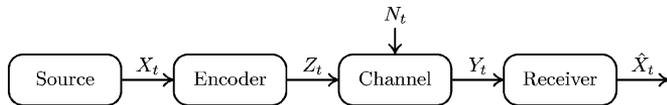


Fig. 1. A point-to-point real-time communication system.

with a real-time or finite-delay constraint on the information transmission were obtained in [16]–[23]. Error exponents of zero-delay source coding were obtained in [24].

For Markovian sources structural properties of optimal encoders and decoders have also been identified. Properties of real-time decoders for noisy observations of a Markov source were considered in [25] and [26]. Properties of real-time encoders for transmitting Markov sources through noiseless channels were investigated in [27] and [28]. The structure of optimal real-time encoding and decoding strategies for systems with noisy channels and noiseless feedback from the decoder to encoder was investigated in [29] and [30]. Applications of results developed in [30] appeared in [31] and [32]. Structural properties of optimal real-time encoding and decoding strategies for systems with Markov source, noisy channels with no feedback and finite memory at the receiver were presented in [33]. Such structural properties are derived by assuming that either the encoder or the decoder is fixed, considering the communication problem at the other end as a stochastic optimization problem, and identifying an information state sufficient for performance analysis.

For real-time source coding and real-time joint source channel coding with noiseless feedback optimality equations to determine optimal encoders and decoders were obtained in [28] and [29], respectively. However, optimality equations that determine optimal encoders and decoders for real-time joint source–channel coding with no feedback are unknown. In this paper, we use the structural results of [33] to obtain such optimality equations.

The real-time communication over noisy channel without feedback (which is considered in this paper) is fundamentally different from the real-time source coding and the real-time communication over noisy channel with noiseless feedback (which were considered in [27]–[29]). The difference lies in the information structures of the models. In the models of [27]–[29] the encoder “knows” all the information available at the decoder. In the model considered in this paper neither the encoder nor the decoder “knows” what is “known” to the other. The models considered in [27]–[29] can be solved by considering the problem from the decoder’s point of view. The model considered in this paper cannot be solved from the point of view of either the encoder or the decoder.

In this paper, we formulate the joint design of real-time encoders and decoders as a decentralized multiagent sequential stochastic optimization problem; the encoder and the decoder can be thought of as the two agents of the system. As mentioned above, the problem has a nonclassical information structure because of the noise in the communication channel. As such, it cannot be solved *directly* by Markov decision theory [34], which is only applicable to stochastic optimization problems with *classical* information structure, or equivalently, problems with centralized information.

In real-time communication both agents (the encoder and the decoder) have a common objective (minimize a total expected distortion). As such, it is a *team problem*. However, systematic methods to solve general team problems are unknown. Nevertheless, in this paper we show that a systematic solution for the real-time communication problem can be obtained by an appropriate choice of information state. This systematic solution is *sequential decomposition*; it breaks the one-shot (brute-force) optimization problem into a sequence of nested optimization problems. This decomposition exponentially simplifies the complexity of the search for optimal real-time encoding and decoding strategies. Moreover, these optimality equations can be extended to infinite horizon problems. This is in contrast to the only other methodology to sequentially decompose team problems—Witsenhausen’s standard form [35], which is only applicable to finite horizon problems.

### C. Contributions

The main contribution of this paper is a solution framework for the optimal design of a real-time communication system. So far, there are no existing communication schemes or tight performance bounds for communicating stochastic sources over noisy channels within a (small) finite delay. The framework presented in this paper has significantly lower complexity than a brute force search; it is also amenable to approximations, though currently no good approximation algorithms exist.

In a finite-horizon communication system, the complexity of searching for an optimal communication scheme by brute force is doubly exponential in the communication length and exponential in the alphabet sizes. We present a search framework whose complexity is linear in the communication length and doubly exponential in the alphabet sizes. We extend this framework to infinite horizon communication systems. For such systems, an optimal communication strategy is obtained from the fixed point of a functional equation. In contrast, a brute force search cannot be used to find an optimal communication scheme for an infinite horizon problem.

We are unaware of any existing algorithm to solve the optimality equations presented in this paper. As a consequence, although the results presented in this paper drastically simplify the search for an optimal communication scheme, yet they do not completely solve the real-time communication problem. Nevertheless, we believe that these optimality equations are useful because they are amenable to approximations. The optimality equations of this paper are similar in structure to the dynamic programming equations for POMDPs (partially observable Markov decision process) with continuous state and action spaces. Approximation algorithms for such POMDPs is an active area of research (see [36]–[39]). Advances in such approximation algorithms along with the framework presented in this paper will provide a complete methodology to obtain real-time communication schemes that perform well for small delays.

Another contribution of this paper is presenting real-time communication as an example in which both finite and infinite horizon variations of dynamic teams with nonclassical information structure can be sequentially decomposed. A general method to sequentially decompose finite and infinite horizon team problems with nonclassical information structures does

not exist; even examples where such decomposition can be obtained are few and far between. We believe that the insights and guidelines on choosing an information state presented in this paper will also be useful for general dynamic teams with nonclassical information structure.

#### D. Organization

The rest of this paper is organized as follows. The finite horizon problem with zero-delay constraint is described in Section II. Information state and its properties are explained in Section III. Then, variables that satisfy these properties are identified, and used to sequentially decompose the optimal design problem. The solution approach for the finite horizon problem is extended to infinite horizon problems in Section IV. Extensions to fixed-finite delay, higher order Markov sources, and channels with memory are considered in Sections V–VII. These extensions are treated in the same manner as the original problem after an appropriate expansion of the state space. Computational aspects are discussed in Section VIII. We contrast the philosophy of our approach with that of information theory and coding theory in Section IX. Concluding remarks and future directions are presented in Section X.

#### E. Notation

Throughout this paper, we use the following notion. Uppercase letters ( $X, Y, Z$ ) represent random variables, lowercase letters ( $x, y, z$ ) represent their realizations, and calligraphic letters ( $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ ) represent their alphabets. Script letters ( $\mathcal{C}, \mathcal{G}, \mathcal{L}$ ) represent family of functions and Gothic letters ( $\mathfrak{F}, \mathfrak{E}, \mathfrak{R}$ ) represent  $\sigma$ -algebras. For random variables and functions,  $x^t$  is a short hand for the sequence  $x_1, \dots, x_t$ , and  $x_a^b$  is a short hand for  $x_a, \dots, x_b$ .  $\mathbb{E}\{\cdot\}$  denotes the expectation of a random variable,  $\Pr(\cdot)$  denotes the probability of an event,  $\mathbb{1}[\cdot]$  denotes the indicator function of a statement, and  $\mathcal{P}(\mathcal{X})$  denotes the space of all PMF (probability mass functions) on  $\mathcal{X}$ . We use  $\mathbb{E}\{\cdot|\phi\}$  and  $\Pr(\cdot|\phi)$  to denote that the expectation of a random variable or the probability of an event depends on a function  $\phi$ . This slightly unusual notation is chosen since we want to keep track of all functional dependencies and the conventional notation of  $\mathbb{E}^\phi\{\cdot\}$  and  $\Pr^\phi\{\cdot\}$  is too cumbersome to use.

## II. THE FINITE HORIZON PROBLEM

### A. Problem Formulation

We first study the finite horizon version of the problem. Consider a discrete time communication system shown in Fig. 1. A first-order Markov source produces a random sequence  $\{X_t, t = 1, \dots, T\}$ . For simplicity of exposition we assume that  $X_t$  takes values in a finite alphabet  $\mathcal{X}$ . Let  $P_{X_1}$  denote the PMF (probability mass function) of the first output  $X_1$ , and  $P_{X_{t+1}|X_t}$  denote the transition probability at time  $t$ .

At each stage  $t$ , the encoder transmits a symbol  $Z_t$  taking values in a finite alphabet  $\mathcal{Z}$ . This encoded symbol is causally generated in real-time using all the source outputs until that time according to an encoding rule  $c_t$ , i.e.

$$Z_t = c_t(X_1, \dots, X_t), \quad t = 1, \dots, T \quad (1)$$

and transmitted through a  $|\mathcal{Z}|$ -input  $|\mathcal{Y}|$ -output DMC (discrete memoryless channel) producing a channel output  $Y_t$  which belongs to a finite alphabet  $\mathcal{Y}$ . The channel is given by

$$Y_t = h_t(Z_t, N_t), \quad t = 1, \dots, T, \quad (2)$$

where  $h_t(\cdot)$  denotes the channel function at time  $t$ , and  $N_t$ , which belongs to a finite alphabet  $\mathcal{N}$ , denotes the channel noise at time  $t$ .  $\{N_t, t = 1, \dots, T\}$  is a sequence of independent random variables that are also independent of the source output  $\{X_t, t = 1, \dots, T\}$ . Let  $P_{N_t}$  denote the PMF of  $N_t$ .

We assume that the receiver has a memory of  $\log_2 |\mathcal{M}|$  bits. So, after some time, the receiver cannot store all the past observations and must selectively *shed* information. This is modeled by assuming that the contents of the memory belong to a finite alphabet  $\mathcal{M}$ . The memory is arbitrarily initialized with  $M_0 = 1$  and then updated at each stage according to the memory update rule  $l_t$ , i.e.

$$M_t = l_t(Y_t, M_{t-1}), \quad t = 1, \dots, T-1. \quad (3)$$

The decoder generates an estimate  $\hat{X}_t$  of the source output in real-time. This estimate takes values in a finite set  $\hat{\mathcal{X}}$  and is generated from the present channel output  $Y_t$  and the memory contents  $M_{t-1}$  according to the decoding rule  $g_t$ , i.e.,

$$\hat{X}_t = g_t(Y_t, M_{t-1}), \quad t = 1, \dots, T. \quad (4)$$

The performance of the system is determined by a sequence of distortion functions,  $\rho_t : \mathcal{X} \times \hat{\mathcal{X}} \rightarrow [0, \rho_{\max}]$ , where  $\rho_{\max} < \infty$ . The function  $\rho_t(X_t, \hat{X}_t)$  measures the distortion at stage  $t$ .

The collection  $C := (c_1, c_2, \dots, c_T)$  of encoding rules for the entire horizon is called an *encoding strategy*. Similarly, the collection  $G := (g_1, g_2, \dots, g_T)$  of decoding rules is called a *decoding strategy* and the collection  $L := (l_1, l_2, \dots, l_T)$  of memory update rules is called a *memory update strategy*. Furthermore, the choice  $(C, G, L)$  of communication rules for the entire horizon is called a *communication strategy* or a *design*. The performance of a communication strategy is quantified by the expected total distortion under that strategy and is given by

$$\mathcal{J}_T(C, G, L) := \mathbb{E} \left\{ \sum_{t=1}^T \rho_t(X_t, \hat{X}_t) \middle| C, G, L \right\}. \quad (5)$$

We are interested in the following optimization problem.

*Problem 2.1:* Assume that the encoder and the receiver know the statistics of the source (i.e., PMF of  $X_1$  and the transition probabilities  $P_{X_{t+1}|X_t}$ ), the channel function  $h_t$ , the statistics  $P_{N_t}$  of the noise, the distortion function  $\rho_t(\cdot, \cdot)$ , and the time horizon  $T$ . Choose a communication strategy  $(C^*, G^*, L^*)$  that is optimal with respect to the performance criterion of (5), i.e.,

$$\mathcal{J}_T(C^*, G^*, L^*) = \mathcal{J}_T^* := \min_{\substack{C \in \mathcal{C}^T \\ G \in \mathcal{G}^T \\ L \in \mathcal{L}^T}} \mathcal{J}_T(C, G, L) \quad (6)$$

where  $\mathcal{C}^T := \mathcal{C}_1 \times \dots \times \mathcal{C}_T$ ,  $\mathcal{C}_t$  is the family of functions from  $\mathcal{X}^t$  to  $\mathcal{Z}$ ,  $\mathcal{G}^T := \mathcal{G} \times \dots \times \mathcal{G}$  ( $T$ -times),  $\mathcal{G}$  is the family of functions from  $\mathcal{Y} \times \mathcal{M}$  to  $\hat{\mathcal{X}}$ ,  $\mathcal{L}^T := \mathcal{L} \times \dots \times \mathcal{L}$  ( $T$ -times), and  $\mathcal{L}$  is the family of functions from  $\mathcal{Y} \times \mathcal{M}$  to  $\mathcal{M}$ .

In Problem 2.1 we want to identify a globally optimal communication strategy to communicate the outputs of a first-order Markov source over a DMC when both the encoding and the decoding have to be done in real-time. Due to this real-time constraint on communication, separate source and channel coding is not optimal. So, we are looking for joint source–channel coding strategies. A globally optimal communication strategy always exists because there are only a finite number of communication strategies and we can always choose one with the best performance. The number of possibly time-varying communication strategies are doubly exponential in the communication length and exponential in the alphabet sizes. Thus, a brute force search for an optimal solution is intractable. As a result, a systematic approach to search for an optimal communication strategy is required. In this paper we present one such systematic approach called *sequential decomposition*, which determines an optimal communication strategy sequentially by proceeding backward in time. The resultant simplified nested optimization problems have linear complexity in the communication length and doubly exponential complexity in the alphabet sizes. In the next section we present an example for a real-time communication system.

### B. An Example

Consider a real-time communication system that runs for three time steps ( $T = 3$ ) with  $\mathcal{X} = \mathcal{Z} = \mathcal{N} = \mathcal{Y} = \hat{\mathcal{X}} = \{0, 1\}$  and  $\mathcal{M} = \{0, 1, \dots, 7\}$ . Suppose the source statistics are

$$P_{X_1} = [0.4 \quad 0.6], \quad P_{X_2|X_1} = P_{X_3|X_2} = \begin{bmatrix} 1.0 & 0.0 \\ 0.1 & 0.9 \end{bmatrix}$$

and the channel is a  $Z$ -channel with crossover probability 0.1, which can be written as

$$h_t(Z_t, N_t) = Z_t \cdot N_t, \quad P_{N_1} = P_{N_2} = P_{N_3} = [0.1 \quad 0.9].$$

The distortion metric is probability of correct reconstruction, i.e.

$$\rho_t(x_t, \hat{x}_t) = \mathbf{1}[x_t = \hat{x}_t].$$

We are going to consider the complexity and performance of two classes of communication strategies for this system: memoryless and real-time. Memoryless communication strategies are a subclass of real-time communication strategies which encode and decode based on only the current symbol, i.e., encoding and decoding is of the form

$$Z_1 = c_1(X_1), \quad Z_2 = c_2(X_2), \quad Z_3 = c_3(X_3), \quad (7a)$$

$$\hat{X}_1 = g_1(Y_1), \quad \hat{X}_2 = g_2(Y_2), \quad \hat{X}_3 = g_3(Y_3). \quad (7b)$$

Since the memory is large enough for the decoder to store all its past observations, real-time communication strategies can be written as

$$Z_1 = c_1(X_1), Z_2 = c_2(X_1, X_2), Z_3 = c_3(X_1, X_2, X_3), \quad (8a)$$

$$\hat{X}_1 = g_1(Y_1), \hat{X}_2 = g_2(Y_1, Y_2), \hat{X}_3 = g_3(Y_1, Y_2, Y_3). \quad (8b)$$

Observe that there are  $(2^2 \times 2^2 \times 2^2)^2 = 4096$  memoryless communication strategies of the form (7), while there are  $(2^2 \times 2^4 \times 2^8)^2 \approx 2.6 \times 10^8$  real-time communication strategies of

the form (8). An optimal memoryless communication strategy is<sup>1</sup>

$$c_1(X_1) = 1 - X_1, \quad c_2(X_2) = 1 - X_2, \quad c_3(X_3) = X_3$$

and

$$g_1(Y_1) = 1 - Y_1, \quad g_2(Y_2) = 1 - Y_2, \quad g_3(Y_3) = Y_3.$$

The total expected distortion of this memoryless communication strategy is 0.1346. An optimal real-time communication strategy is<sup>1</sup>

$$c_1(X_1) = 1 - X_1, \quad c_2(X_1, X_2) = 1 - X_1 \cdot X_2$$

$$c_3(X_1, X_2, X_3) = 1 - X_1 \cdot X_2 \cdot X_3$$

and

$$g_1(Y_1) = 1 - Y_1, \quad g_2(Y_1, Y_2) = (1 - Y_1) \cdot (1 - Y_2)$$

$$g_3(Y_1, Y_2, Y_3) = (1 - Y_1) \cdot (1 - Y_2) \cdot (1 - Y_3).$$

The total expected distortion of this real-time communication strategy is 0.0564. Thus, for this example, real-time communication strategies provide 2.39 times smaller distortion than memoryless communication strategies.

It was shown in [14], [15] that for memoryless sources, i.e., when  $\{X_t, t = 1, \dots, T\}$  is i.i.d. (independent and identically distributed), memoryless strategies are optimal for real-time communication. The above example shows that this need not be the case when the source is Markovian.

The main difficulty with finding optimal real-time communication strategies is that the number of communication strategies increase doubly exponentially with the communication length. The above example, which is one of the simplest real-time communication systems, has around  $10^8$  real-time communication strategies (with the dominant term being the  $2^{2^3}$  real-time encoding and decoding strategies at stage 3). This doubly exponential dependence of the number of real-time communication strategies on the communication length makes a brute force search impractical for systems that operate for large horizons. In the rest of this paper, we present a systematic method to search for an optimal real-time communication strategy, which reduces the search complexity to be linear in the communication length (at the cost of searching over a Borel space instead of a discrete space). We now present some concepts and notation needed for the rest of the paper.

### C. Primitive Random Variables

In this paper we work with conditional probabilities, probability measures of probability measures, and  $\sigma$ -fields. To be precise in our analysis we need to define the probability space clearly. For that matter, we first define the primitive random variables of the system.

Any Markov chain  $\{X_t, t = 1, \dots, T\}$  can be defined using a sequence of independent uniform  $[0, 1]$  random variables  $\{\theta_t, t = 1, \dots, T\}$  such that

$$X_{t+1} = f_t(X_t, \theta_t) \quad (9)$$

<sup>1</sup>The optimal strategy was obtained numerically, it is simply presented in analytic form here.

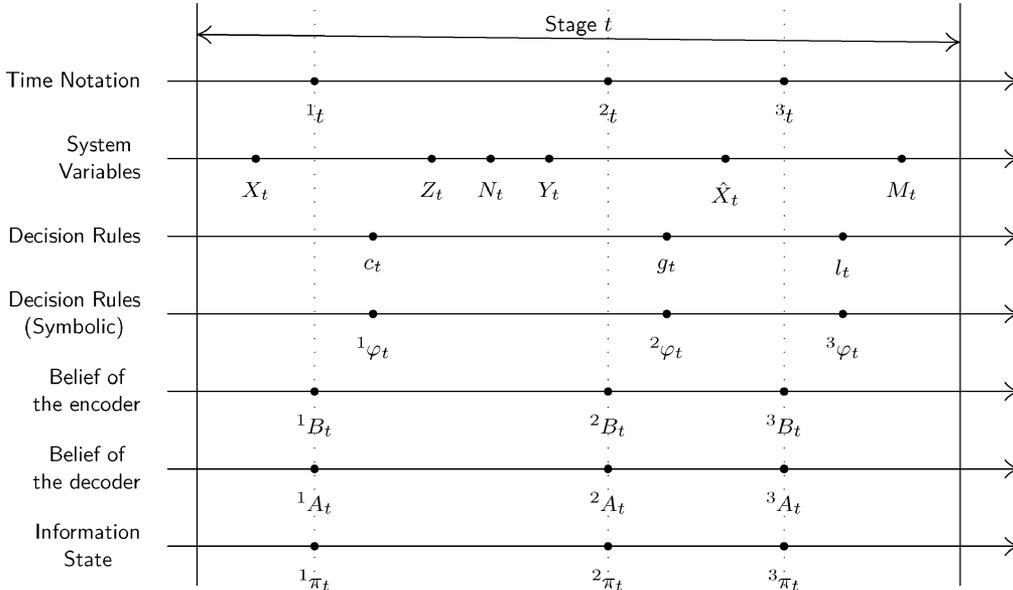


Fig. 2. The sequential ordering of the variables of the real time communication system.  $1_t$ ,  $2_t$ , and  $3_t$  are refinements of stage  $t$ .

where for  $x, x' \in \mathcal{X}$ ,  $\theta \in [0, 1]$

$$f_t(x, \theta) = x' \text{ iff } F_{X_{t+1}|X_t}(x' - 1|x) < \theta \leq F_{X_{t+1}|X_t}(x'|x)$$

and  $F_{X_{t+1}|X_t}(\cdot|x)$  is the CDF (cumulative distribution function) of  $P_{X_{t+1}|X_t}(\cdot|x)$ . This method is called the inverse transform method or Smirnov transform (see [40, Sec. 4.11]). Assume that  $\{X_t, t = 1, \dots, T\}$  is generated from independent variables  $(X_1, \theta_1, \dots, \theta_T)$  using (9). Since, the noise in the communication channel is independent of the source output, the random variables  $(X_1, \theta_1, \dots, \theta_T, N_1, \dots, N_T)$  are independent. These random variables are called the primitive random variables. We assume that all primitive random variables are defined on a common probability space  $(\Omega, \mathfrak{F}, P)$ . If the communication strategy is fixed, all system variables are defined in terms of the primitive random variables, and are  $(\Omega, \mathfrak{F}, P)$  measurable. In the sequel, all (random) variables are assumed to be defined on  $(\Omega, \mathfrak{F}, P)$ .

#### D. Problem Classification

Problem 2.1 is a sequential stochastic optimization problem in the sense of [35]. The sequential nature of the problem can be understood by refining the notion of time. We call each step of the system a *stage*. For each stage consider three time instances<sup>2</sup>:  $1_t$ ,  $2_t$  and  $3_t$ . Assume that the system has three “agents”: the encoder (agent 1), the decoder (agent 2), and the memory update (agent 3). There is no loss of generality in assuming that these agents act sequentially at  $1_t$ ,  $2_t$ , and  $3_t$ , respectively. The choice of decision rules and the realization of primitive random variables do not affect the order in which the agents act. Hence, Problem 2.1 is a sequential problem. The sequential ordering of the system variables is shown in Fig. 2 (some of these variables will be defined later).

In Problem 2.1, all agents have the same objective given by (5). Multiagent problems in which all agents have the same objective are called teams [41], and are further classified as static

<sup>2</sup>The actual values of these time instances are irrelevant; we need three values in increasing order.

or dynamic teams on the basis of their information structure. In static teams, an agent’s information is a function of primitive random variables only. In contrast, in dynamic teams an agent’s information depends on the functional form of the decision rules of other agents. In Problem 2.1 the receiver’s information depends on the functional form of the encoding rule. Thus, Problem 2.1 is a dynamic team. Dynamic teams are, in general, functional optimization problems having a complex interdependence among the decision rules [42]. This interdependence leads to nonconvex (in policy space) optimization problems that are hard to solve.

#### E. Information Fields

For the ease of notation let  ${}^i\phi_t$  and  ${}^i\phi^{t-1}$  denote the current and all the past decision rules at time  ${}^i t$ ,  $i = 1, 2, 3$ , i.e.,

$$\begin{aligned} {}^1\phi^{t-1} &:= (c^{t-1}, g^{t-1}, l^{t-1}), & {}^1\phi_t &:= c_t, \\ {}^2\phi^{t-1} &:= (c^t, g^{t-1}, l^{t-1}), & {}^2\phi_t &:= g_t, \\ {}^3\phi^{t-1} &:= (c^t, g^t, l^{t-1}), & {}^3\phi_t &:= l_t. \end{aligned}$$

Recall that  $(\Omega, \mathfrak{F}, P)$  is the probability space on which all primitive random variables are defined. Suppose  ${}^i O_t$  is the observation of agent  $i$  at time  ${}^i t$ . For any choice  ${}^i\phi^{t-1}$  of the last decision rules,  ${}^i O_t$  is measurable with respect to  $\mathfrak{F}$ . All the information (about the randomness in  $\mathfrak{F}$ ) that agent  $i$  can collect from his observations  ${}^i O_t$  is called the *information field* of agent  $i$  at time  ${}^i t$ . It is equal to the smallest subfield of  $\mathfrak{F}$  with respect to  ${}^i O_t$  is measurable and is denoted by  $\sigma({}^i O_t; {}^i\phi^{t-1})$ . The information fields at the encoder’s and the receiver’s site are given below.

**Definition 2.1 (Encoder’s Information):** Let  ${}^i E_t$  denote the observation and  ${}^i \mathfrak{E}_t$  denote the information field at the encoder’s site at time  ${}^i t$ ,  $i = 1, 2, 3$ . Then

$${}^1 E_t := (X^t, Z^{t-1}), \quad {}^2 E_t := (X^t, Z^t), \quad {}^3 E_t := {}^2 E_t \quad (10a)$$

and

$${}^i \mathfrak{E}_t := \sigma({}^i E_t; {}^i\phi^{t-1}), \quad i = 1, 2, 3. \quad (10b)$$

Let  ${}^i \mathcal{E}_t$  denote the space of realizations of  ${}^i E_t$ ,  $i = 1, 2, 3$ .

*Definition 2.2 (Decoder's Information):* Let  ${}^iR_t$  denote the observation and  ${}^i\mathfrak{R}_t$  denote the information field at the receiver's site at time  ${}^it$ ,  $i = 1, 2, 3$ . Then

$${}^1R_t := M_{t-1}, \quad {}^2R_t := (Y_t, M_{t-1}), \quad {}^3R_t := {}^2R_t \quad (11a)$$

and

$${}^i\mathfrak{R}_t := \sigma({}^iR_t; {}^i\phi^{t-1}), \quad i = 1, 2, 3. \quad (11b)$$

Let  ${}^i\mathcal{R}$  denote the space of realizations of  ${}^iR_t$ ,  $i = 1, 2, 3$ .

The above defined information fields highlight the following features of the problem.

F1) *Nonclassical information structure.*

In general, the noise in the channel makes the information fields at the encoder's and the receiver's sites non-compatible, i.e.,  ${}^i\mathcal{E}_t \not\subseteq {}^i\mathfrak{R}_t$  and  ${}^i\mathcal{E}_t \not\supseteq {}^i\mathfrak{R}_t$ . Thus, at no time during the evolution of the system does the encoder "know" what is known to the receiver and *vice versa*. Hence, the information in the system is decentralized and Problem 2.1 has a nonclassical information structure.

F2) *Shedding of information at the receiver due to finite memory.*

The encoder has perfect memory, i.e., it remembers all the past observations. As a result the information fields at the encoder are nested, i.e.,  ${}^1\mathcal{E}_t \subseteq {}^2\mathcal{E}_t \subseteq {}^3\mathcal{E}_t \subseteq {}^1\mathcal{E}_{t+1}$  and so on. On the other hand, the receiver has finite memory. As a result, the information fields at the receiver are non-nested: although  ${}^1\mathfrak{R}_t \subseteq {}^2\mathfrak{R}_t \subseteq {}^3\mathfrak{R}_t$ , at time  ${}^3t$  when the receiver updates its memory  ${}^3\mathfrak{R}_t \not\subseteq {}^1\mathfrak{R}_{t+1}$ , i.e., at time  ${}^3t$  the receiver sheds information.

F. *Agents' Belief and Their Evolution*

As explained in F1) above, the encoder does not "know" what is "known" to the receiver and *vice versa*. So, we need to characterize what the encoder "thinks" that the receiver has "seen" and what the receiver "thinks" that the encoder has "seen." This is captured by the encoder's belief about the observations at the receiver and the receiver's belief about the observations at the encoder. These beliefs are given below.

*Definition 2.3 (Encoder's Beliefs):* Let  ${}^iB_t$  denote the encoder's belief about the receiver's observation at time  ${}^it$ ,  $i = 1, 2, 3$ . Then for  ${}^ir \in {}^i\mathcal{R}$

$${}^iB_t({}^ir) := \Pr({}^iR_t = {}^ir \mid {}^i\mathcal{E}_t). \quad (12)$$

Let  ${}^i\mathcal{B} := \mathcal{P}({}^i\mathcal{R})$  denote the space of realizations of  ${}^iB_t$

*Definition 2.4 (Receiver's Beliefs):* Let  ${}^iA_t$  denote the receiver's belief about the encoder's observation at time  ${}^it$ ,  $i = 1, 2, 3$ . Then for  ${}^ie_t \in {}^i\mathcal{E}_t$

$${}^iA_t({}^ie_t) := \Pr({}^iE_t = {}^ie_t \mid {}^i\mathfrak{R}_t). \quad (13)$$

Let  ${}^i\mathcal{A}_t := \mathcal{P}({}^i\mathcal{E}_t)$  denote the space of realizations of  ${}^iA_t$ . Furthermore, let  $\hat{A}_t$  denote the receiver's belief about the source output at time instant  ${}^2t$ , i.e., for  $x \in \mathcal{X}$

$$\hat{A}_t(x) := \Pr(X_t = x \mid {}^2\mathfrak{R}_t). \quad (14)$$

The sequential ordering of the beliefs is shown in Fig. 2. For any particular realization  ${}^1e_t$  of  ${}^1E_t$ , and any arbitrary (but fixed)

choice of  ${}^1\phi^{t-1}$ , the realization  ${}^1b_t$  of  ${}^1B_t$  is a PMF on  $\mathcal{M}$ . If  ${}^1E_t$  is a random vector, then  ${}^1B_t$  is a random vector belonging to  $\mathcal{P}(\mathcal{M})$ , the space of PMFs on  $\mathcal{M}$ . Similar interpretations hold for  ${}^2B_t$ ,  ${}^2B_t$ ,  ${}^1A_t$ ,  ${}^2A_t$ , and  ${}^3A_t$ .

The belief of the encoder evolve as follows.

*Lemma 2.1 (Evolution of the Encoder's Beliefs):* For each stage  $t$ , there exist deterministic functions  ${}^1F_t$  and  ${}^3F$  such that

$${}^2B_t = {}^1F_t({}^1B_t, Z_t), \quad {}^3B_t = {}^2B_t, \quad {}^1B_{t+1} = {}^3F({}^3B_t, l_t).$$

This is proved in Appendix A.

G. *Structural Properties*

The structural properties of optimal real-time encoders and decoders were derived in [33]. We summarize these properties below and will later use them to develop a methodology for determining globally optimal communication strategies.

*Theorem 2.1 (Structure of Optimal Real-Time Encoders):* Consider Problem 2.1 for any arbitrary (but fixed) decoding and memory update strategies,  $G = (g_1, g_2, \dots, g_T)$  and  $L = (l_1, l_2, \dots, l_T)$ , respectively. Then there is no loss of optimality in restricting attention to encoding rules of the form

$$Z_t = c_t(X_t, {}^1B_t), \quad t = 2, \dots, T. \quad (15)$$

*Theorem 2.2 (Structure of Optimal Real-Time Decoders):* Consider Problem 2.1 for any arbitrary (but fixed) encoding and memory update strategies,  $C = (c_1, c_2, \dots, c_T)$  and  $L = (l_1, l_2, \dots, l_T)$ , respectively. Then there is no loss of optimality in restricting attention to decoding rules of the form

$$\hat{X}_t = \hat{g}_t(\hat{A}_t) := \arg \min_{\hat{x} \in \hat{\mathcal{X}}} \sum_{x \in \mathcal{X}} \rho_t(x, \hat{x}) \hat{A}_t(x). \quad (16)$$

H. *Implication of the Structural Results*

Let  $\hat{\mathcal{C}}$  denote the space of functions from  $\mathcal{X} \times {}^1\mathcal{B}$  to  $\mathcal{Z}$ . The result of Theorem 2.1 states that instead of choosing an encoding rule from the space  $\mathcal{C}_t$  at time  $t$ , an encoding rule from the space  $\hat{\mathcal{C}}$  can be chosen. Therefore, we have the following corollary.

*Corollary 2.1:* The optimal performance  $\mathcal{J}_T^*$  given by (6) can be determined by

$$\mathcal{J}_T^* := \inf_{\substack{C \in \hat{\mathcal{C}}^T \\ G \in \mathcal{G}^T \\ L \in \mathcal{L}^T}} \mathcal{J}_T(C, G, L) \quad (17)$$

where  $\hat{\mathcal{C}}^T := \hat{\mathcal{C}} \times \dots \times \hat{\mathcal{C}}$  ( $T$ -times), and  $\mathcal{G}^T$  and  $\mathcal{L}^T$  are defined as before.

Hence, in Problem 2.1 rather than choosing a communication strategy  $(C^*, G^*, L^*)$  belonging to  $(\mathcal{C}^T \times \mathcal{G}^T \times \mathcal{L}^T)$  to minimize (6), a communication strategy  $(C^*, G^*, L^*)$  belonging to  $(\hat{\mathcal{C}}^T \times \mathcal{G}^T \times \mathcal{L}^T)$  to minimize (17) can be chosen. The domain of an encoding rule belonging to  $\mathcal{C}_t$  increases with  $t$ . In contrast, the domain of an encoding rule belonging to  $\hat{\mathcal{C}}$  does not depend on  $t$ . Hence, using the structural results of Theorem 2.1, Problem 2.1 can be reformulated such that the encoding rules at each time are

chosen from a time-invariant space. The reformulated problem is as follows.

*Problem 2.2:* Under the assumptions of Problem 2.1, choose a communication strategy  $(C^*, G^*, L^*)$  belonging to  $(\hat{\mathcal{C}}^T \times \mathcal{G}^T \times \mathcal{L}^T)$  that is optimal with respect to the performance criterion of (17).

In the sequel, we will concentrate on Problem 2.2.

### III. GLOBAL OPTIMIZATION

In this section we provide a methodology for sequential decomposition of Problem 2.2. The class of problems consisting of Problem 2.2 and its infinite horizon extensions belong to a category for which no sequential decomposition methodology is known in general. In order to obtain a sequential decomposition, we need to find “information states sufficient for performance evaluation.” We explain the properties that such information states should satisfy, and then guess information states with these properties and show how they lead to a sequential decomposition.

#### A. Information Structures and Information State

In a multiagent system, the collection of sets of data available to each agent as arguments of its decision rule is called the *information structure* (or *information pattern*) of the system. Multiagent systems can be classified according to their information structures. In [43] three classes of information structures are defined: classical, quasi-classical, and nonclassical information structures. A system is said to have a classical information structure if all agents observe the same data and have perfect recall (or equivalently, if the information fields of all agents at a given time stage are equal and the information fields across time are nested). The system is said to have a *quasi-classical* information structure if each agent knows the information available to any other agent whose actions affect its observations. A system that has neither classical nor quasi-classical information structure is said to have a (strictly) nonclassical information structure.

Markov decision theory explains how to obtain a sequential decomposition of problems with a classical information structure. Witsenhausen’s standard form [35] explains how to obtain a sequential decomposition of a subclass of problems with a nonclassical information structure. Methodologies to sequentially decompose a general problem with a nonclassical information structure are unknown. Problem 2.1 and its infinite horizon extensions have a nonclassical information structure so they cannot be solved by Markov decision theory; the infinite horizon extensions of Problem 2.1 belong to a subclass that cannot be solved using the standard form. So, we need to develop a new methodology for sequential decomposition of these problems.

Identifying information states sufficient for performance evaluation is a critical step in sequentially decomposing problems with nonclassical information structures. An information state is a sufficient statistic that satisfies certain properties. Unfortunately, all definitions of information states in the literature

are in terms of their properties for systems with a classical information structure; there is no explanation of the properties of information states for systems with a nonclassical information structure. The idea behind an information state is best described in [44]: “The [information] state should be a summary (“compression”) of some data (the “past”) known to someone (an observer or a controller) and sufficient for some purposes (input–output map, optimization, dynamic programming).”

In this section we define the properties that the information states sufficient for performance analysis should satisfy and explain what these properties mean in the context of real-time communication. These properties are as follows.

#### P1) *Sufficient summary of past information*

The information state should be a representation of all the past information that is sufficient for future performance evaluation. This has the following interpretation. The real-time communication problem is a controlled stochastic input–output system. The stochastic inputs are  $\{X_t, t = 1, \dots, T\}$  and  $\{N_t, t = 1, \dots, T\}$ , and the outputs are  $\{\hat{X}_t, t = 1, \dots, T\}$ . The designer has to choose a communication strategy  $(C, G, L)$ . Suppose the system is at time  $t$ : nature has produced  $(x^t, n^t)$ , the designer has chosen  $i\phi^{t-1}$ , and the system has produced  $\hat{x}^{t-1}$  and incurred a distortion  $\sum_{s=1}^{t-1} \rho_s(x_s, \hat{x}_s)$ . The designer now wants to choose  $i\phi_t^T$  to minimize the expected future distortion  $\mathbb{E} \left\{ \sum_{s=t}^T \rho_s(X_s, \hat{X}_s) \middle| i\phi^{t-1}, i\phi_t^T \right\}$ .

Different choices of the past communication rules are equivalent for the purpose of evaluating future performance if any choice of future decision rules lead to the same expected future performance. In other words, two choices of past decision rules  $i\phi^{t-1,(1)}$  and  $i\phi^{t-1,(2)}$  are equivalent, denoted by  $i\phi^{t-1,(1)} \sim i\phi^{t-1,(2)}$ , if for any choice of future decision rules  $i\phi_t^T$

$$\begin{aligned} \mathbb{E} \left\{ \sum_{s=t}^T \rho_s(X_s, \hat{X}_s) \middle| i\phi^{t-1,(1)}, i\phi_t^T \right\} \\ = \mathbb{E} \left\{ \sum_{s=t}^T \rho_s(X_s, \hat{X}_s) \middle| i\phi^{t-1,(2)}, i\phi_t^T \right\}. \end{aligned}$$

Assume that the designer has already chosen  $i\phi^{t-1}$  and wants to choose  $i\phi_t^T$  to minimize the expected future cost. If  $i\phi^{t-1,(1)} \sim i\phi^{t-1,(2)}$  then the optimal future communication rules are same for both of them. So, to evaluate future performance and choose future communication rules, it is sufficient for the designer to keep track of the equivalence class of the past communication rules.

Let  $i\Phi^{t-1}$  denote the space of realization of all past decision rules, and let  $i\Pi_t$  be any arbitrary space. Suppose  $i\pi_t : i\Phi^{t-1} \rightarrow i\Pi_t$  is a function such that for any  $i\phi^{t-1,(1)}, i\phi^{t-1,(2)} \in i\Phi^{t-1}$  if  $i\pi_t(i\phi^{t-1,(1)}) = i\pi_t(i\phi^{t-1,(2)})$ , then  $i\phi^{t-1,(1)} \sim i\phi^{t-1,(2)}$ . Any such  $i\pi_t$  is a sufficient statistic for future performance evaluation.

#### P2) *Common knowledge and sequential update.*

All agents in the system should be able to solve the sequential decomposition of the problem. So, the information state cannot depend on data that is observed locally by one of the agents. In fact, the information state should be common knowledge in the sense of Aumann [45], and the agents should be able to keep track of how the information state evolves with time.

In centralized stochastic optimization (i.e., problems with classical information structure), the conditional expectation of the state conditioned on the agent's data is an information state appropriate for performance evaluation. However, in decentralized stochastic optimization (i.e., problems with nonclassical information structures) such conditional expectations cannot be information states as they are not common knowledge: the data observed at each agent is not common knowledge, hence conditional expectations based on this data is not common knowledge. The sufficient statistics  ${}^i\pi_t$  of (P1) are derived from past decision rules, which are common knowledge. So, they can be evaluated both at the encoder and the receiver.

Furthermore, for the purpose of sequential decomposition,  ${}^2\pi_t(2\phi^{t-1})$  has to be a function of  ${}^1\pi_t(1\phi^{t-1})$  and  ${}^1\phi_t$  (recall that  ${}^1\phi^{t-1} = ({}^1\phi^{t-1}, {}^1\phi_t)$ ),  ${}^3\pi_t(3\phi^{t-1})$  has to be a function of  ${}^2\pi_t(2\phi^{t-1})$  and  ${}^2\phi_t$ , and  ${}^1\pi_{t+1}(1\phi^t)$  has to be a function of  ${}^3\pi_t(3\phi^{t-1})$  and  ${}^3\phi_t$ .

Any sequence  $\{{}^i\pi_t, i = 1, 2, 3, t = 1, \dots, T\}$  that has properties P1) and P2) is a valid choice of information state, and can be used to obtain a sequential decomposition for the finite horizon problem. We want to develop a methodology that can be extended to infinite horizon problem. For that matter, we require the following additional property.

3) *Time invariant domain.*

Identify functions  ${}^i\pi_t : {}^i\Phi^{t-1} \rightarrow {}^i\Pi$  such that  $\{{}^i\pi_t, i = 1, 2, 3, t = 1, \dots, T\}$  satisfy (P1) and (P2) and the sets  ${}^1\Pi$ ,  ${}^2\Pi$ , and  ${}^3\Pi$  do not depend on the time horizon  $T$ .

An information state should provide representation of past knowledge that is efficient, both in calculating optimal decision rules and in their implementation. The smaller the set of all realizations of the information state, the more efficient is it to compute optimal communication rules. So, the following property is desirable.

4) *Minimality.*

If more than one appropriate information state exist, we want to use the "smallest" information state. However, we have not been able to establish a good way of comparing information states, especially, information states satisfying P3). So, in the rest of the paper, we do not consider minimality.

For a given communication rule  $(C, G, L)$ ,  ${}^i\pi_t(i\phi^{t-1})$  is the information state at time  $t$  and denoted by  ${}^i\pi_t$ . In summary, these information states satisfy the following properties.

- S1) *The information state is a summary of past information.* Thus,  ${}^1\pi_t$  is a function of  ${}^1\phi^{t-1}$ ,  ${}^2\pi_t$  is a function of  ${}^2\phi^{t-1}$ , and  ${}^3\pi_t$  is a function of  ${}^3\phi^{t-1}$ .
- S2) *Both the encoder and the receiver should be able to keep track of the information states.*

This means that  ${}^2\pi_t$  can be determined from  ${}^1\pi_t$  and  ${}^1\phi_t$  (i.e.,  ${}^1\pi_t$  and  $c_t$ ),  ${}^3\pi_t$  can be determined from  ${}^2\pi_t$  and  ${}^2\phi_t$  (i.e.,  ${}^2\pi_t$  and  $g_t$ ), and  ${}^1\pi_{t+1}$  can be determined from  ${}^3\pi_t$  and  ${}^3\phi_t$  (i.e.,  ${}^3\pi_t$  and  $l_t$ ).

- S3) *The information state should be sufficient for performance evaluation, that is, it should absorb the affect of past decisions on future performance.*

This means that

$$\begin{aligned} & \mathbb{E} \left\{ \sum_{s=t}^T \rho_s(X_s, \hat{X}_s) \middle| C, G, L \right\} \\ &= \mathbb{E} \left\{ \sum_{s=t}^T \rho_s(X_s, \hat{X}_s) \middle| {}^3\pi_{t-1}, c_t^T, g_t^T, l_{t-1}^T \right\} \\ &= \mathbb{E} \left\{ \sum_{s=t}^T \rho_s(X_s, \hat{X}_s) \middle| {}^1\pi_t, c_t^T, g_t^T, l_t^T \right\} \\ &= \mathbb{E} \left\{ \sum_{s=t}^T \rho_s(X_s, \hat{X}_s) \middle| {}^2\pi_t, c_{t+1}^T, g_t^T, l_t^T \right\}. \end{aligned} \quad (18)$$

In general, it is difficult to check (18). Nevertheless, the state update relation of S2) suggests the following sufficient condition of (18):

$$\begin{aligned} & \mathbb{E} \left\{ \rho_t(X_t, \hat{X}_t) \middle| C, G, L \right\} \\ &= \mathbb{E} \left\{ \rho_t(X_t, \hat{X}_t) \middle| {}^3\pi_{t-1}, l_{t-1}, c_t, g_t \right\} \\ &= \mathbb{E} \left\{ \rho_t(X_t, \hat{X}_t) \middle| {}^1\pi_t, c_t, g_t \right\} \\ &= \mathbb{E} \left\{ \rho_t(X_t, \hat{X}_t) \middle| {}^2\pi_t, g_t \right\}. \end{aligned} \quad (19)$$

- S4) *The information states should belong to time-invariant spaces.*

This means that there exist spaces  ${}^1\Pi$ ,  ${}^2\Pi$ , and  ${}^3\Pi$  such that for all  $t$ ,  ${}^i\pi_t \in {}^i\Pi$ ,  $i = 1, 2, 3$ .

Properties S1) and S2) are equivalent to property P1), properties S1) and S3) are equivalent to P2), and property S4) is equivalent to P3).

In order to obtain a sequential decomposition, we need to identify information states  ${}^1\pi_t$ ,  ${}^2\pi_t$ , and  ${}^3\pi_t$  that satisfy properties S1)–S4). As mentioned earlier, there is no general method of identifying appropriate information states for problems with a nonclassical information structure. Next, we first guess information states that satisfy the above properties, and then show how to obtain a sequential decomposition using these information states.

*Definition 3.1:* Define  ${}^1\pi_t$ ,  ${}^2\pi_t$ , and  ${}^3\pi_t$  as follows:

$$\begin{aligned} {}^1\pi_t &:= \Pr(X_t, {}^1B_t | {}^1\phi^{t-1}), \\ {}^2\pi_t &:= \Pr(X_t, {}^2B_t | {}^2\phi^{t-1}), \\ {}^3\pi_t &:= \Pr(X_t, {}^3B_t | {}^3\phi^{t-1}). \end{aligned}$$

Let  ${}^i\Pi$ ,  $i = 1, 2, 3$ , denote the space of probability measures on  $(\mathcal{X} \times {}^i\mathcal{B})$ . Then  ${}^i\pi_t$  takes values in  ${}^i\Pi$ .

The above definitions are to be interpreted as follows. Let  $(\Omega, \mathfrak{F}, P)$  denote the probability space on which all primitive random variables are defined. For any choice  ${}^i\phi^{t-1}$  of past decision rules for agent  $i$ ,  $i = 1, 2, 3$ , the beliefs  ${}^iB_t$  are  $\mathfrak{F}$ -measur-

able. Thus, for any choice of  ${}^i\phi^{t-1}$ ,  $(X_t, {}^iB_t)$  is  $\mathfrak{F}$ -measurable.  ${}^i\pi_t$  is the corresponding induced measure on  $(\mathcal{X} \times {}^i\mathcal{B})$ .

The above defined probability measures are related as follows.

*Lemma 3.1:* For encoding rules of the form (15),  ${}^1\pi_t$ ,  ${}^2\pi_t$ , and  ${}^3\pi_t$  are information states for the encoder, the decoder, and the memory update respectively, i.e.

1) there exist linear transformations  ${}^1Q_t$  and  ${}^3Q_t$  such that

$${}^2\pi_t = {}^1Q_t(c_t) {}^1\pi_t, \quad {}^3\pi_t = {}^2\pi_t, \quad {}^1\pi_{t+1} = {}^3Q_t(l_t) {}^3\pi_t. \quad (20)$$

2) The expected instantaneous cost can be expressed as

$$\mathbb{E} \left\{ \rho_t(X_t, \hat{X}_t) \middle| {}^2\phi^{t-1}, g_t \right\} = \hat{\rho}_t({}^2\pi_t, g_t). \quad (21)$$

This is proved in Appendix B. Observe that by definition  ${}^i\pi_t$  satisfies S1). Part 1 of Lemma 3.1 shows that they satisfy S2); part 2 shows that they satisfy S3). S4) is satisfied by definition. Next we show how these information states lead to a sequential decomposition.

### B. An Equivalent Optimization Problem

Consider a centralized deterministic optimization problem with state space alternating between  ${}^1\Pi$ ,  ${}^2\Pi$ , and  ${}^3\Pi$  and action space alternating between  $\hat{\mathcal{C}}$ ,  $\mathcal{G}$ , and  $\mathcal{L}$ . The system dynamics are given by (20) and at each stage  $t$  the decision rules  $c_t$ ,  $g_t$ , and  $l_t$  are determined according to *meta-functions* or *meta-rules*  ${}^1\Delta_t$ ,  ${}^2\Delta_t$ , and  ${}^3\Delta_t$ , where  ${}^1\Delta_t$  is a function from  ${}^1\Pi$  to  $\hat{\mathcal{C}}$ ,  ${}^2\Delta_t$  is a function from  ${}^2\Pi$  to  $\mathcal{G}$ , and  ${}^3\Delta_t$  is a function from  ${}^3\Pi$  to  $\mathcal{L}$ . Thus, the system (20) can be written as

$$c_t = {}^1\Delta_t({}^1\pi_t), \quad {}^2\pi_t = {}^1Q(c_t) {}^1\pi_t, \quad (22a)$$

$$g_t = {}^2\Delta_t({}^2\pi_t), \quad {}^3\pi_t = {}^2\pi_t, \quad (22b)$$

$$l_t = {}^3\Delta_t({}^3\pi_t), \quad {}^1\pi_{t+1} = {}^3Q(l_t) {}^3\pi_t. \quad (22c)$$

The initial state  ${}^1\pi_1 = P_{X_1}$  is given. At stage  $t$ , an instantaneous cost  $\hat{\rho}_t({}^2\pi_t, g_t)$  is incurred. The choice  $({}^1\Delta_1, {}^2\Delta_1, {}^3\Delta_1, \dots, {}^1\Delta_T, {}^2\Delta_T, {}^3\Delta_T)$  is called a *meta-strategy* or a *meta-design* and denoted by  $\Delta^T$ . The performance of a meta-strategy is given by the total cost incurred by that meta-strategy, i.e.,

$$\mathcal{J}_T(\Delta^T | {}^1\pi_1) = \sum_{t=1}^T \hat{\rho}_t({}^2\pi_t, g_t). \quad (23)$$

Now consider the following optimization problem.

*Problem 3.1:* Consider the dynamic system (22) with known transformations  ${}^1Q$  and  ${}^3Q$ . The initial state  ${}^1\pi_1$  is given. Determine a meta-strategy  $\Delta^T$  to minimize the total cost given by (23).

Given any meta-strategy  $\Delta^T$ , the time evolution of  ${}^i\pi_t$  is deterministic;  ${}^i\pi_t$  and the corresponding  ${}^i\phi_t$  can be determined from (22). Thus, for a given initial states  ${}^1\pi_t$ , there is a communication strategy corresponding to any choice of meta-strategy.

Furthermore, the performance criterion of (5) can be rewritten as

$$\begin{aligned} \mathcal{J}_T(C, G, L) &= \mathbb{E} \left\{ \sum_{t=1}^T \rho_t(X_t, \hat{X}_t) \middle| C, G, L \right\} \\ &\stackrel{(a)}{=} \sum_{t=1}^T \mathbb{E} \left\{ \rho_t(X_t, \hat{X}_t) \middle| {}^2\phi^{t-1}, g_t \right\} \\ &\stackrel{(b)}{=} \sum_{t=1}^T \hat{\rho}_t({}^2\pi_t, g_t) \\ &=: \mathcal{J}_T(\Delta^T | {}^1\pi_1) \end{aligned}$$

where (a) follows from the sequential ordering of system variables and (b) follows from Lemma 3.1. Thus, if  $\Delta^{*,T}$  is an optimal meta-strategy for Problem 3.1, and  $(C^*, G^*, L^*)$  is the communication strategy corresponding to  $\Delta^{*,T}$ , then  $(C^*, G^*, L^*)$  is an optimal communication strategy for Problem 2.2 and thereby also for Problem 2.1. Hence, Problem 3.1 is equivalent to Problems 2.1 and 2.2. Now we provide an algorithm to determine an optimal meta-strategy for Problem 3.1.

### C. The Global Optimization Algorithm

Problem 3.1 can be formulated as a classical centralized stochastic control problem by considering the information state  ${}^i\pi_t$  as the ‘‘controlled state’’ at time  $t$ , the communication rule  ${}^i\phi_t$  ( $c_t$ ,  $g_t$ , or  $l_t$  depending on  $i$ ) as the ‘‘control action’’ (or decision) at time  $t$ , and the meta-function  ${}^i\Delta_t$  as the ‘‘control law’’ at time  $t$ . Hence, an optimal meta-strategy for Problem 3.1 is given by the optimal ‘‘control strategy’’ of the centralized stochastic control problem and can be determined as follows:

*Theorem 3.1 (Global Optimization Algorithm):* An optimal meta-strategy  $\Delta^{*,T}$  for Problem 3.1 (and consequently an optimal communication strategy for Problem 2.1) can be determined by the solution of the following nested optimality equations. For all  ${}^i\pi \in {}^i\Pi$ ,  $i = 1, 2, 3$ , define

$${}^1V_{T+1}({}^1\pi) = 0, \quad (24a)$$

and for  $t = 1, \dots, T$

$${}^1V_t({}^1\pi) = \inf_{c_t \in \hat{\mathcal{C}}} {}^2V_t({}^1Q_t(c_t) {}^1\pi), \quad (24b)$$

$${}^2V_t({}^2\pi) = \min_{g_t \in \mathcal{G}} \left[ \hat{\rho}_t({}^2\pi, g_t) + {}^3V_t({}^2\pi) \right], \quad (24c)$$

$${}^3V_t({}^3\pi) = \min_{l_t \in \mathcal{L}} {}^1V_{t+1}({}^3Q_t(l_t) {}^3\pi). \quad (24d)$$

The functions  ${}^iV_t$  are called *value functions*; they represent the minimum expected future cost that the system in state  ${}^i\pi$  will incur from time  $t$  onwards. These value functions can be determined iteratively by moving backwards in time. The optimal performance of Problem 3.1 (and Problem 2.1) is given by

$$\mathcal{J}_T^* = {}^1V_1({}^1\pi_1). \quad (25)$$

For any  ${}^i\pi$  and  ${}^i\pi$ , the  $\arg \min$  (or  $\arg \inf$ ) in the RHS of  ${}^iV_t({}^i\pi)$  equals to the optimal value of the meta-function  ${}^i\Delta_t({}^i\pi_t)$ . Thus,

solving for the value functions for all values of the information state also determines an optimal meta-strategy  $\Delta^{*,T}$  for Problem 3.1. Relations (22) can be used to determine optimal communication strategy for Problem 2.1.

*Proof:* As explained in the previous paragraph, Problem 3.1 can be formulated as a centralized stochastic control problem. Therefore, the optimality of the sequential decomposition of (24) follows from standard results in controlled Markov processes, see [34, Ch. 2]. The communication policy corresponding to  $\Delta^{*,T}$  is optimal for Problem 2.1 because of the equivalence between Problems 2.1 and 3.1.

Observe that the three step  $T$ -stage sequential decomposition of (24) can be combined into a one-step  $T$ -stage sequential decomposition

$${}^1V_t({}^1\pi) = \inf_{\substack{c \in \hat{\mathcal{C}}^T \\ g \in \mathcal{G}^T \\ l \in \mathcal{L}^T}} \left[ \hat{\rho}_t({}^1Q_t(c_t) {}^1\pi, g_t) + {}^1V_{t+1} \left( \left( {}^3Q_t(l_t) \circ {}^1Q_t(c_t) \right) {}^1\pi \right) \right] \quad (26)$$

which is a deterministic dynamic program in function space. The decomposition of Theorem 3.1 is based on the refinement of time presented in Section II-D; consequently, it has a smaller search space than the decomposition of (26).

#### D. The Time Homogeneous Case

In many scenarios the system is time-homogeneous, i.e., the source statistics  $P_{X_{t+1}|X_t}$ , the channel function  $h_t(\cdot)$ , the noise statistics  $P_{N_t}$  and the distortion function  $\rho_t(\cdot)$ , do not depend on time  $t$ . If the system of Problem 2.1 is time-homogeneous, some of the results derived in the previous section can be simplified. The function  ${}^1F_t$  in Lemma 2.1 does not depend on  $t$ ; the transformations  ${}^1Q_t$ ,  ${}^3Q_t$  and the function  $\hat{\rho}_t$  of Lemma 3.1 also do not depend on  $t$ ; thus, we drop the subscripts  $t$  and simply denote them by  ${}^1F$ ,  ${}^1Q$ ,  ${}^3Q$  and  $\hat{\rho}$ , respectively. Hence, Problem 3.1 reduces to a time-homogeneous problem—the state space, the action space, the system update equations, and the instantaneous distortion do not depend on  $t$ . Furthermore, Theorem 3.1 simplifies as follows.

*Corollary 3.1:* If the system of Problem 2.1 is time-homogeneous, the nested optimality equations (24) can be written as

$${}^1V_{T+1}({}^1\pi) = 0 \quad (27a)$$

and for  $t = 1, \dots, T$

$${}^1V_t({}^1\pi) = \inf_{c_t \in \hat{\mathcal{C}}} {}^2V_t({}^1Q(c_t) {}^1\pi) \quad (27b)$$

$${}^2V_t({}^2\pi) = \min_{g_t \in \mathcal{G}} \left[ \hat{\rho}({}^2\pi, g_t) + {}^3V_t({}^2\pi) \right] \quad (27c)$$

$${}^3V_t({}^3\pi) = \min_{l_t \in \mathcal{L}} {}^1V_{t+1}({}^3Q(l_t) {}^3\pi). \quad (27d)$$

Notice that in the above equations  ${}^1Q$ ,  ${}^3Q$ , and  $\hat{\rho}$  do not depend on  $t$ .

#### IV. THE INFINITE HORIZON TIME-HOMOGENEOUS PROBLEM

In this section we extend the time-homogeneous model of Section III-D to an infinite horizon ( $T \rightarrow \infty$ ) using two criteria:

the expected discounted distortion and the average distortion per unit time. Let  $(C, G, L)$ ,  $C := (c_1, c_2, \dots)$ ,  $G := (g_1, g_2, \dots)$ ,  $L := (l_1, l_2, \dots)$  denote an infinite horizon communication strategy or an infinite horizon design. The two performance criteria that we consider are as follows.

- 1) The expected discounted cost criteria.

Under this criteria the performance of a communication strategy is given by

$$\mathcal{J}^\beta(C, G, L) := \mathbb{E} \left\{ \sum_{t=1}^{\infty} \beta^{t-1} \rho(X_t, \hat{X}_t) \middle| C, G, L \right\} \quad (28)$$

where  $0 < \beta < 1$  is called the discount factor.

- 2) The average cost per unit time criteria.

Under this criteria the performance of a communication strategy is given by

$$\bar{\mathcal{J}}(C, G, L) := \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left\{ \sum_{t=1}^T \rho(X_t, \hat{X}_t) \middle| C, G, L \right\}. \quad (29)$$

Time-invariant communication strategies are desirable while implementing a communication scheme for infinite horizon. This motivates the following definitions.

*Definition 4.1 (Stationary Communication Strategy):* A communication strategy  $(C, G, L)$ ,  $C = (c_1, c_2, \dots)$ ,  $G = (g_1, g_2, \dots)$ ,  $L = (l_1, l_2, \dots)$  is called stationary (or time-invariant) if  $c_1 = c_2 = \dots = c$ ,  $g_1 = g_2 = \dots = g$ , and  $l_1 = l_2 = \dots = l$ . Such a stationary communication strategy is equivalently denoted by  $(c^\infty, g^\infty, l^\infty)$ .

*Definition 4.2 (Stationary Meta-Strategy):* A meta-strategy  $\tilde{\Delta}^\infty = (\tilde{\Delta}_1, \tilde{\Delta}_2, \dots)$ , where  $\tilde{\Delta}_t = ({}^1\Delta_t, {}^2\Delta_t, {}^3\Delta_t)$ , is called stationary (or time-invariant) if  $\tilde{\Delta}_1 = \tilde{\Delta}_2 = \dots = \tilde{\Delta}$ .

In time-homogeneous infinite-horizon stochastic optimization problems with classical information structures, under some technical conditions there is no loss of optimality in restricting attention to stationary strategies (see [34]). This result drastically simplifies the search for an optimal solution. It is not known whether, in general, restricting attention to stationary strategies is optimal for problems with nonclassical information structures. (Recall that Problem 2.1 has a nonclassical information structure.) In this section we show that for the time-homogeneous infinite-horizon extensions of Problem 2.1, stationary communication strategies may not be optimal. However, there is no loss of optimality in restricting attention to stationary meta-strategies: for the expected discounted distortion criterion there exist stationary meta-strategies that are optimal; for the average cost per unit time criterion, under a technical condition, there exist stationary meta-strategies that are arbitrarily close to optimal.

However, the optimal communication strategy corresponding to the stationary meta-strategy is, in general, time-varying.

##### A. The Expected Discounted Distortion Problem

Consider a time-homogeneous infinite-horizon problem with the expected discounted distortion criterion of (28). With a

slight modification to the proof of [33] one can show that the structural results of Section II-G are valid for this case, hence we can restrict attention to encoders belonging to  $\hat{\mathcal{C}}$ . Consider  ${}^1\pi_t$ ,  ${}^2\pi_t$ , and  ${}^3\pi_t$  as in Definition 3.1: they satisfy the properties of Lemma 3.1; furthermore, since the system is time-invariant, the transformations  ${}^1Q$  and  ${}^3Q$  and the expected instantaneous distortion  $\hat{\rho}$  do not depend on  $t$ . Let  $\gamma_t := (c_t, g_t, l_t)$  denote the communication rules at time  $t$  and  $\Gamma$  denote the space  $\hat{\mathcal{C}} \times \mathcal{G} \times \mathcal{L}$ . We can combine (22) as

$${}^1\pi_{t+1} \tilde{Q}(\gamma_t) {}^1\pi_t, \quad \gamma_t = \tilde{\Delta}_t({}^1\pi_t) \quad (30)$$

where

$$\tilde{Q}(\gamma_t) = {}^3Q(l_t) \circ {}^1Q(c_t)$$

and

$$\tilde{\Delta}_t({}^1\pi_t) = ({}^1\Delta_t({}^1\pi_t), {}^2\Delta_t({}^1Q({}^1\Delta_t({}^1\pi_t)) {}^1\pi_t), {}^3\Delta_t({}^1Q({}^1\Delta_t({}^1\pi_t)) {}^1\pi_t)).$$

The instantaneous distortion at time  $t$  can be written as

$$\tilde{\rho}({}^1\pi_t, {}^1\gamma_t) = \hat{\rho}({}^1Q(c_t), {}^1\pi_t, g_t).$$

Hence, the time-homogeneous infinite horizon problem with the expected discounted cost criterion of (28) is equivalent to the following deterministic optimization problem.

*Problem 4.1:* Consider a deterministic system with state space  ${}^1\Pi$  and action space  $\Gamma$ . The system dynamics are given by

$${}^1\pi_{t+1} = \tilde{Q}(\gamma_t) {}^1\pi_t, \quad \gamma_t = \tilde{\Delta}_t({}^1\pi_t) \quad (31)$$

where  $\tilde{Q}$  is a known transformation and  $\tilde{\Delta} : \Pi \rightarrow \Gamma$  is a meta-function. At each time an instantaneous cost  $\tilde{\rho}({}^1\pi_t, \gamma_t)$  is incurred. The initial state  ${}^1\pi_1$  is known. The objective is to choose meta-strategy  $\tilde{\Delta}^\infty = (\tilde{\Delta}_1, \tilde{\Delta}_2, \dots)$  so as to minimize the discounted infinite horizon total distortion given by

$$\mathcal{J}^\beta(\tilde{\Delta}^\infty) = \sum_{t=1}^{\infty} \beta^{t-1} \tilde{\rho}({}^1\pi_t, \gamma_t). \quad (32)$$

Problem 4.1 is a standard deterministic time-invariant infinite horizon problem with total discounted distortion (cost) criterion. As we have assumed  $\rho(\cdot)$  to be uniformly bounded,  $\hat{\rho}$  and  $\tilde{\rho}$  are also uniformly bounded. Therefore, an optimal meta-strategy is guaranteed to exist and can be obtained as follows.

*Theorem 4.1:* For Problem 4.1 and consequently for the infinite horizon expected discounted cost problem with the performance criterion given by (28) one can restrict attention to stationary meta-strategies without any loss of optimality. Specifically there exists a stationary meta-strategy  $\tilde{\Delta}^{*,\infty} := (\tilde{\Delta}^*, \tilde{\Delta}^*, \dots)$ , and a corresponding infinite horizon communication strategy  $(C^*, G^*, L^*)$ ,  $C^* := (c_1, c_2, \dots)$ ,  $G := (g_1, g_2, \dots)$ ,  $L := (l_1, l_2, \dots)$  such that

$$\mathcal{J}^\beta(\tilde{\Delta}^{*,\infty}) = V({}^1\pi_1), \quad (33)$$

where  $V$  is the unique uniformly bounded fixed point of

$$V({}^1\pi) = \min_{\gamma \in \Gamma} \{ \tilde{\rho}({}^1\pi, \gamma) + \beta V(\tilde{Q}(\gamma)({}^1\pi)) \}, \quad (34)$$

and  $\tilde{\Delta}^*$  satisfies

$$V({}^1\pi) = \tilde{\rho}({}^1\pi, \tilde{\Delta}^*({}^1\pi)) + \beta V(\tilde{Q}(\tilde{\Delta}^*({}^1\pi))({}^1\pi)). \quad (35)$$

Optimal communication rules  $(c_t^*, g_t^*, l_t^*)$  at time  $t$  are given by

$$(c_t^*, g_t^*, l_t^*) := \gamma_t^* = \tilde{\Delta}^*({}^1\pi_t). \quad (36)$$

*Proof:* Since Problem 4.1 is a centralized stochastic control problem with uniformly bounded cost, the result follows from standard results for infinite horizon discounted cost problems in centralized stochastic control, see [46, Ch. 6]. The fixed point equation (34) can be decomposed into its ‘‘natural’’ sequential form as

$${}^1V({}^1\pi) = \inf_{c \in \hat{\mathcal{C}}} {}^2V({}^1Q(c) {}^1\pi), \quad (37a)$$

$${}^2V({}^2\pi) = \min_{g \in \mathcal{G}} \tilde{\rho}({}^2\pi, g) + \beta {}^3V({}^2\pi), \quad (37b)$$

$${}^3V({}^3\pi) = \min_{l \in \mathcal{L}} {}^1V({}^3Q(l) {}^3\pi). \quad (37c)$$

These equations are the infinite horizon analogue of (24).

### B. The Average Distortion Per Unit Time Problem

Consider the time-homogeneous infinite horizon problem with the average distortion per unit time criterion of (29). Using the arguments similar to the first paragraph of Section IV-B, this problem is equivalent to the following deterministic problem:

*Problem 4.2:* Consider a deterministic system with state space  ${}^1\Pi$  and action space  $\Gamma$ . The system dynamics are given by

$${}^1\pi_{t+1} = \tilde{Q}(\gamma) {}^1\pi_t, \quad \gamma_t = \tilde{\Delta}_t({}^1\pi_t) \quad (38)$$

where  $\tilde{Q}$  is a known transformation and  $\tilde{\Delta} : \Pi \rightarrow \Gamma$  is a meta-function. At each time an instantaneous cost  $\tilde{\rho}({}^1\pi_t, \gamma_t)$  is incurred. The initial state  ${}^1\pi_1$  is known. The objective is to choose meta-strategy  $\tilde{\Delta}^\infty = (\tilde{\Delta}_1, \tilde{\Delta}_2, \dots)$  so as to minimize the average distortion per unit time over an infinite horizon given by

$$\bar{\mathcal{J}}(\tilde{\Delta}^\infty) = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \tilde{\rho}({}^1\pi_t, \gamma_t). \quad (39)$$

Problem 4.2 *cannot* be solved by taking the limit  $\beta \rightarrow 1$  in the result of Theorem 4.1. Such a result is valid only if the problem has finite state and action space (see [47, Theorem 31.5.2]) which is not the case here. See [48] for a survey of various results connecting the expected discounted cost problem with the average cost per unit time problem.

For Problem 4.2 an optimal meta-strategy may not exist. However, under suitable conditions, meta-strategies that are arbitrarily close to optimal exist. Specifically, we have the following result.

*Theorem 4.2:* For Problem 4.2 and correspondingly for the infinite horizon average cost per unit time problem with the performance criterion given by (29), assume the following.

- A1) For any  $\varepsilon > 0$  there exist bounded measurable functions  $v(\cdot)$  and  $r(\cdot)$  and meta-function  $\tilde{\Delta}^* : \Pi \rightarrow \Gamma$  such that for all  ${}^1\pi$

$$v({}^1\pi) = \min_{\gamma \in \Gamma} v(\tilde{Q}(\gamma) {}^1\pi) = v(\tilde{Q}(\tilde{\Delta}^*({}^1\pi)) {}^1\pi) \quad (40)$$

and

$$\begin{aligned} \tilde{\rho}({}^1\pi, \tilde{\Delta}^*({}^1\pi)) + r(\tilde{Q}(\tilde{\Delta}^*({}^1\pi)) {}^1\pi) &\leq v({}^1\pi) + r({}^1\pi) \\ &\leq \min_{\gamma \in \Gamma} \left\{ \tilde{\rho}({}^1\pi, \gamma) + r(\tilde{Q}(\gamma) {}^1\pi) \right\} + \varepsilon. \end{aligned} \quad (41)$$

Then for any horizon  $T$  and any meta-strategy  $\tilde{\Delta}^{*T} := (\tilde{\Delta}_1^*, \tilde{\Delta}_2^*, \dots, \tilde{\Delta}_T^*)$ , the stationary meta-strategy  $\tilde{\Delta}^{*,T} := (\tilde{\Delta}^*, \dots, \tilde{\Delta}^*)$  ( $T$ -times) satisfies

$$\mathcal{J}_T(\tilde{\Delta}^{*,T}) \leq r({}^1\pi_1) + T v({}^1\pi_1) \leq \mathcal{J}_T(\tilde{\Delta}^{*T}) + T\varepsilon. \quad (42)$$

Furthermore, the stationary meta-strategy  $\tilde{\Delta}^{*,\infty} := (\tilde{\Delta}^*, \tilde{\Delta}^*, \dots)$  is  $\varepsilon$ -optimal (i.e.,  $\varepsilon$  close to optimal). That is, for any infinite horizon meta-strategy  $\tilde{\Delta}^\infty := (\tilde{\Delta}_1^*, \tilde{\Delta}_2^*, \dots)$  we have

$$\overline{\mathcal{J}}(\tilde{\Delta}^{*,\infty}) \leq v({}^1\pi_1) \leq \underline{\mathcal{J}}(\tilde{\Delta}^\infty) + \varepsilon \quad (43)$$

where

$$\overline{\mathcal{J}}(\tilde{\Delta}^{*,\infty}) := \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \tilde{\rho}({}^1\pi_t, \tilde{\Delta}^*(\pi_t)) \quad (44)$$

with  ${}^1\pi_{t+1} = \tilde{Q}(\tilde{\Delta}^*({}^1\pi_t)) {}^1\pi_t$  and

$$\underline{\mathcal{J}}(\tilde{\Delta}^\infty) := \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \tilde{\rho}({}^1\pi_t, \tilde{\Delta}_t^*(\pi_t)) \quad (45)$$

with  ${}^1\pi_{t+1} = \tilde{Q}(\tilde{\Delta}_t^*({}^1\pi_t)) {}^1\pi_t$ .  $\varepsilon$ -optimal communication rules  $(c_t^*, g_t^*, l_t^*)$  at time  $t$  are given by

$$(c_t^*, g_t^*, l_t^*) =: \gamma_t^* = \tilde{\Delta}^*({}^1\pi_t). \quad (46)$$

*Proof:* Since Problem 4.1 is a centralized stochastic control problem with uniformly bounded cost, the result follows from standard results for infinite horizon average cost per unit time problems in centralized stochastic control, see [46, Ch. 7].

Conditions that guarantee that assumption A1) of Theorem 4.2 is satisfied are fairly technical and do not provide much insight into the properties of the source, the channel, and the distortion functions that will guarantee the existence of such policies. The interested reader may look at [46, Ch. 7, Sec.10] and [49, Ch. 3, Sec. 3.3].

### C. Discussion of the Results

The discussion of Section III-B shows that one can view the real-time communication problem as an equivalent deterministic optimization problem by considering the information state

as the ‘‘controlled state,’’ the communication rule as the ‘‘control action’’ and the meta-function as the ‘‘control law’’ at each time. In classical infinite-horizon deterministic optimization problems, there is no loss of optimality in restricting attention to stationary control laws; by analogy, in the infinite-horizon real-time communication problem, there is no loss of optimality in restricting attention to stationary meta-strategies. In classical infinite-horizon deterministic optimization problems, stationary actions are not optimal in general; by analogy, in infinite-horizon real-time communication problem, stationary communication strategies are not optimal in general. In the absence of a systematic framework, the task of finding and implementing an optimal infinite-horizon communication strategy is infeasible. The methodology of this section provides one systematic framework: obtain and implement time-varying optimal infinite-horizon communication strategies by obtaining and implementing stationary infinite-horizon meta-strategies. The off-line search simplifies to finding the fixed point of a functional equation. Once an optimal stationary meta-strategy is obtained, both the encoder and the decoder can store it, and use it to obtain the current optimal communication rules by keeping track of the current information state. This greatly simplifies the on-line implementation of a time-varying optimal communication strategy.

## V. FIXED-FINITE DELAY

For many applications where communication delay is important, the acceptable delay is finite and fixed but nonzero. In this section we consider the case when the distortion metric tolerates a fixed-finite delay  $\delta$ , i.e., at time  $t$ ,  $t > \delta$ , the decoder tries to estimate the source output at time  $t - \delta$  and a distortion  $\rho_t(X_{t-\delta}, \hat{X}_t)$  is incurred. This case can be modeled by modifying the model of Section II-A as follows.

- M1) The variables  $\hat{X}_1, \dots, \hat{X}_\delta$  are simply not generated; the receiver spends the first  $\delta$  periods just accumulating the observations  $Y_1, \dots, Y_\delta$  and updating its memory accordingly.
- M2) The performance of a communication scheme  $(C, G, L)$ ,  $C := (c_1, c_2, \dots, c_T)$ ,  $G := (g_{\delta+1}, \dots, g_T)$ ,  $L := (l_1, l_2, \dots, l_T)$  is given by

$$\mathcal{J}_T(C, G, L) := \mathbb{E} \left\{ \sum_{t=\delta+1}^T \rho_t(X_{t-\delta}, \hat{X}_t) \middle| C, G, L \right\}. \quad (47)$$

Assume that the system runs for more than  $\delta$  steps, that is,  $T > \delta$ . We are interested in the following optimization problem.

*Problem 5.1:* Consider the model of Problem 2.1 with modifications M1) and M2) defined above. Choose a communication strategy  $(C^*, G^*, L^*)$  that is optimal with respect to the performance criterion of (47).

### A. Transformation to a Zero-Delay Problem

In this section, we show how to convert the fixed-finite delay problem into a zero-delay problem. This can be done using the sliding window repackaging of the source as presented in [27]

and [33]. For that matter, define a process  $\{\bar{X}_t, t = 1, \dots, T\}$  as follows:

$$\bar{X}_t := \begin{cases} (X_t, \dots, X_t), & t \leq \delta \\ (X_{t-\delta}, \dots, X_t), & t > \delta. \end{cases} \quad (48)$$

Let  $\bar{\mathcal{X}}_t$  denote the space of realizations of  $\bar{X}_t$ , i.e.,  $\bar{\mathcal{X}}_t = \mathcal{X}^t$  for  $t \leq \delta$  and  $\bar{\mathcal{X}}_t = \mathcal{X}^{\delta+1}$  for  $t > \delta$ . Observe that  $X^t = \bar{X}^t$ , so for any  $c_t \in \mathcal{C}_t$  we can find a  $\bar{c}_t: \bar{\mathcal{X}}^t \rightarrow \mathcal{Z}$  such that

$$Z_t = c_t(X_1, \dots, X_t) = \bar{c}_t(\bar{X}_1, \dots, \bar{X}_t).$$

Let  $\bar{\mathcal{C}}_t$  denote the collection of all  $\bar{c}_t$  corresponding to all  $c_t \in \mathcal{C}_t$ . Define  $\bar{g}_t$  as follows

$$\bar{g}_t(Y_t, M_{t-1}) := \begin{cases} 0, & t \leq \delta \\ g_t(Y_t, M_{t-1}), & t > \delta. \end{cases}$$

Let  $\bar{\mathcal{G}}_t = \emptyset$  for  $t \leq \delta$  and  $\bar{\mathcal{G}}_t = \mathcal{G}$ . Furthermore, define a modified distortion function  $\bar{\rho}_t$  as follows:

$$\bar{\rho}_t(\bar{X}_t, \hat{X}_t) := \begin{cases} 0, & t \leq \delta; \\ \rho_t(X_{t-\delta}, \hat{X}_t), & t > \delta. \end{cases} \quad (49)$$

Using these transformations, the total distortion under a communication strategy  $(C, G, L)$  can be written as

$$\begin{aligned} \mathcal{J}_T(C, G, L) &= \mathbb{E} \left\{ \sum_{t=\delta+1}^T \rho_t(X_{t-\delta}, \hat{X}_t) \middle| C, G, L \right\} \\ &= \mathbb{E} \left\{ \sum_{t=1}^T \bar{\rho}_t(\bar{X}_t, \hat{X}_t) \middle| \bar{\mathcal{C}}, \bar{\mathcal{G}}, L \right\} \\ &=: \bar{\mathcal{J}}_T(\bar{\mathcal{C}}, \bar{\mathcal{G}}, L). \end{aligned} \quad (50)$$

Hence, Problem 5.1 is equivalent to the following problem.

*Problem 5.2:* Consider Problem 5.1 with the sliding window repackaging of the source given by (48). Choose a communication strategy  $(\bar{\mathcal{C}}^*, \bar{\mathcal{G}}^*, L^*)$  that is optimal with respect to the performance criterion of (50), i.e.

$$\bar{\mathcal{J}}_T(\bar{\mathcal{C}}^*, \bar{\mathcal{G}}^*, L^*) = \bar{\mathcal{J}}_T^* := \min_{\substack{\bar{\mathcal{C}} \in \bar{\mathcal{C}}^T \\ \bar{\mathcal{G}} \in \bar{\mathcal{G}}^T \\ L \in \mathcal{L}^T}} \mathcal{J}_T(\bar{\mathcal{C}}, \bar{\mathcal{G}}, L) \quad (51)$$

where  $\bar{\mathcal{C}}^T := \bar{\mathcal{C}}_1 \times \dots \times \bar{\mathcal{C}}_T$ ,  $\bar{\mathcal{G}}^T := \bar{\mathcal{G}}_1 \times \dots \times \bar{\mathcal{G}}_T$ , and  $\bar{\mathcal{C}}_t, \bar{\mathcal{G}}_T$ , and  $\mathcal{L}^T$  are as defined earlier.

### B. Agents' Beliefs and Structural Results

Problem 5.2 is a zero-delay real-time communication problem. So, the analysis and results of Sections II–IV can be applied to this problem. We can define the encoder's and receiver's beliefs as in Definitions 2.3 and 2.4. We need to modify  $\hat{A}_t(x)$  in Definition 2.4 as follows:

$$\hat{A}_t(x) := \Pr(X_{t-\delta} = x | \mathfrak{R}_t), \quad t = \delta + 1, \dots, T. \quad (52)$$

The structural properties of optimal encoders and decoders can be restated as follows.

*Theorem 5.1:* Consider Problem 5.2 for any arbitrary (but fixed) decoding and memory update strategies,

$\bar{\mathcal{C}} = (g_1, g_2, \dots, g_T)$  and  $L = (l_1, l_2, \dots, l_T)$ , respectively. Then there is no loss of optimality in restricting attention to encoding rules of the form

$$Z_t = \bar{c}_t(\bar{X}_t, {}^1B_t), \quad t = 2, \dots, T.$$

which is equivalent to

$$Z_t = c_t(X_1, \dots, X_t, B_t), \quad t = 2, \dots, \delta \quad (53a)$$

$$Z_t = c_t(X_{t-\delta}, \dots, X_t, B_t), \quad t = \delta + 1, \dots, T. \quad (53b)$$

*Theorem 5.2:* Consider Problem 5.2 for any arbitrary (but fixed) encoding and memory update strategies,  $\bar{\mathcal{C}} = (\bar{c}_1, \bar{c}_2, \dots, \bar{c}_T)$  and  $L = (l_1, l_2, \dots, l_T)$ , respectively. Then there is no loss of optimality in restricting attention to decoding rules of the form

$$\hat{X}_t = \hat{g}_t(\hat{A}_t) := \arg \min_{\hat{x} \in \hat{\mathcal{X}}} \sum_{x \in \mathcal{X}} \rho_t(x, \hat{x}) \hat{A}_t(x). \quad (54)$$

where  $\hat{A}_t$  is defined in (52).

### C. Global Optimization

Globally optimal communication strategies for Problem 5.2 can be obtained along the lines of Section III. For that matter, define information states  ${}^1\bar{\pi}$ ,  ${}^2\bar{\pi}$ , and  ${}^3\bar{\pi}$  as follows:

$${}^i\bar{\pi}_t := \Pr(\bar{X}_t, {}^iB_t) = \begin{cases} \Pr(X_1, \dots, X_t, {}^iB_t), & t \leq \delta \\ \Pr(X_{t-\delta}, \dots, X_t, {}^iB_t), & t > \delta. \end{cases} \quad (55)$$

These information states are related in the same manner as Lemma 3.1. So, we can formulate an equivalent optimization problem as in Section III-B and solve it along the lines of the nested optimality equations in Section III-C. Observe that in Problem 5.2,  $\bar{\rho}_t(\bar{X}_t, \hat{X}_t) = 0$  for  $t \leq \delta$ , so the global optimization algorithm simplifies as follows.

*Theorem 5.3:* An optimal meta-strategy  $\Delta^{*,T}$  for Problem 5.2 can be determined by the solution of the following nested optimality equations. Define value function  ${}^iV_t, i = 1, 2, 3, t = 1, \dots, T$  as follows: for all  ${}^i\bar{\pi}_t \in {}^i\bar{\Pi} := \mathcal{P}(\mathcal{X}^{\delta+1} \times {}^i\mathcal{B})$  and

$${}^iV_{T+1}({}^1\bar{\pi}) = 0 \quad (56a)$$

and for  $t = \delta + 1, \dots, T$

$${}^1V_t({}^1\bar{\pi}) = \inf_{\bar{c}_t \in \bar{\mathcal{C}}} {}^2V_t({}^1Q_t(\bar{c}_t) {}^1\bar{\pi}), \quad (56b)$$

$${}^2V_t({}^2\bar{\pi}) = \min_{g_t \in \mathcal{G}} \left[ \hat{\rho}_t({}^2\bar{\pi}, g_t) + {}^3V_t({}^2\bar{\pi}) \right] \quad (56c)$$

$${}^3V_t({}^3\bar{\pi}) = \min_{l_t \in \mathcal{L}} {}^1V_{t+1}({}^3Q_t(l_t) {}^3\bar{\pi}). \quad (56d)$$

and for  $t = 1, \dots, \delta$  and  ${}^i\bar{\pi}_t \in {}^i\bar{\Pi}_t := \mathcal{P}(\mathcal{X}^t \times {}^i\mathcal{B})$

$${}^1V_t({}^1\bar{\pi}_t) = \inf_{\bar{c}_t \in \bar{\mathcal{C}}} {}^2V_t({}^1Q_t(\bar{c}_t) {}^1\bar{\pi}_t), \quad (57a)$$

$${}^2V_t({}^2\bar{\pi}_t) = {}^3V_t({}^2\bar{\pi}_t), \quad (57b)$$

$${}^3V_t({}^3\bar{\pi}_t) = \min_{l_t \in \mathcal{L}} {}^1V_{t+1}({}^3Q_t(l_t) {}^3\bar{\pi}_t). \quad (57c)$$

The value function can be determined iteratively by moving backwards in time. The optimal performance  $\mathcal{J}_T^*$  and optimal meta-strategy  $\Delta^{*,T}$  can be determined from the value functions in the same way as in Theorem 3.1.

#### D. The Infinite Horizon Time-Homogeneous Problem

In this section we assume that the system is time-homogeneous and consider the infinite-horizon problem for the two performance criteria described in Section IV. Observe that a time-homogeneous model for Problem 5.1 *does not* imply that the transformed zero delay problem is time-homogeneous. This is because for the first  $\delta$  steps  $\bar{X}_t$  takes values in a space that is increasing with time. Thus, for the first  $\delta$  time steps, the system is not time-homogeneous; from  $\delta + 1$  onwards, it is time-homogeneous. Therefore, the infinite horizon problems can be broken into two phases:

- 1) the initialization phase, and
- 2) the sliding window phase.

The initialization phase is for the first  $\delta$  time steps, and the sliding window phase is from  $\delta + 1$  onwards.

Now, an optimal communication strategy for the infinite-horizon problem with the expected discounted distortion criterion can be obtained by first obtaining the value function and an optimal meta-strategy for the sliding window phase and then obtaining an optimal meta-strategy for the sliding window phase by treating it as a finite horizon problem. This is explained in detail below.

*Theorem 5.4:* An optimal meta-strategy for the time-homogeneous infinite-horizon expected discounted distortion problem can be determined as follows.

##### 1) The sliding window phase.

The sliding window phase can be transformed into a zero-delay time-homogeneous infinite-horizon expected discounted distortion problem using the transformation of Section V-B. Therefore, we can use Theorem 4.1 to find the value function  $V$  and a stationary meta-strategy  $\Delta^*$  that is optimal for  $\delta + 1$  onwards.  $V$  is given by the unique fixed point of (34) and  $\Delta^*$  satisfies (35).

##### 2) The initialization phase.

The initialization phase corresponds to a finite horizon problem where there is no instantaneous distortion, and only a final expected distortion corresponding to the value function  $V$  which was determined in the sliding window phase, i.e., for  ${}^i\bar{\pi}_t \in {}^i\bar{\Pi}_t := \mathcal{P}(\mathcal{X}^t \times {}^i\mathcal{B})$ , we have

$$V_{\delta+1}({}^i\bar{\pi}_{\delta+1}) = \beta^\delta V({}^i\bar{\pi}_{\delta+1}) \quad (58a)$$

and for  $t = 1, \dots, \delta$

$${}^1V_t({}^1\bar{\pi}_t) = \inf_{\bar{c}_t \in \bar{\mathcal{C}}} {}^2V_t({}^1Q_t(\bar{c}_t) {}^1\bar{\pi}_t), \quad (58b)$$

$${}^2V_t({}^2\bar{\pi}_t) = {}^3V_t({}^2\bar{\pi}_t), \quad (58c)$$

$${}^3V_t({}^3\bar{\pi}_t) = \min_{l_t \in \mathcal{L}} {}^1V_{t+1}({}^3Q_t(l_t) {}^3\bar{\pi}_t). \quad (58d)$$

For the average distortion per unit time problem, it does not matter what communication strategy is used in the initialization phase. Since the sliding window phase is time-homogeneous,

we can use Theorem 4.2 to find an optimal infinite-horizon meta-strategy for  $\delta + 1$  onwards, and use any policy for the initialization phase.

## VI. HIGHER-ORDER MARKOV SOURCES

In many applications the source statistics are higher-order Markov rather than first-order Markov. Such applications can be modeled by making the following modification to the model of Section II-A:

M1') The source output  $\{X_t, t = 1, \dots, T\}$  is  $k$ th order Markov, i.e., for  $t > k$  and  $x_1, \dots, x_{t+1} \in \mathcal{X}$ , we have

$$\begin{aligned} \Pr(X_{t+1} = x_{t+1} | X^t = x^t) \\ = \Pr(X_{t+1} = x_{t+1} | X_{t-k+1}^t = x_{t-k+1}^t). \end{aligned}$$

This model was considered in [33, Sec. III-B] and it was shown that

- 1) there is no loss of optimality in restricting attention to encoders of the form

$$Z_t = c_t(X_{t-k+1}, \dots, X_t, {}^1B_t), \quad t = k + 1, \dots, T;$$

- 2) the structure of the optimal receiver is the same as that of the model of Section II-A.

In this section, we show how to obtain globally optimal communication strategies for this model. The key idea is to transform the problem into a first-order Markov source, in the same manner as the finite-delay problem is transformed into a zero-delay problem. For that matter, define the process  $\{\bar{X}_t, t = 1, \dots, T\}$  as follows:

$$\bar{X}_t := \begin{cases} (X_1, \dots, X_t), & t \leq k \\ (X_{t-k+1}, \dots, X_t), & t > k. \end{cases} \quad (59)$$

Observe that  $\{\bar{X}_t, t = 1, \dots, T\}$  is a first-order Markov process. The structural results of [33] state that we can restrict attention to encoders of the form

$$Z_t = c_t(\bar{X}_t, {}^1B_t). \quad (60)$$

Let  $\bar{\mathcal{C}}_t$  denote the class of all encoders at time  $t$  of the form (60).

The channel and the receiver are the same as in the model of Section II-A. Define a modified distortion function  $\bar{\rho}_t$  as follows:

$$\bar{\rho}_t(\bar{X}_t, \hat{X}_t) := \rho_t(X_t, \hat{X}_t). \quad (61)$$

With these modifications we can formulate an optimization problem similar to Problem 2.2. The finite horizon problem can be solved in the same manner as Section III. For the infinite horizon variation, we need to break the problem into two phases.

- 1) The initialization phase which lasts for the first  $k$  steps.
- 2) The sliding window phase which starts from  $k + 1$  onwards. The infinite horizon problem for both the expected discounted distortion and average distortion per unit time criteria can be solved over these two phases along the lines of the solution for the fixed-finite delay problem presented in Section V-D.

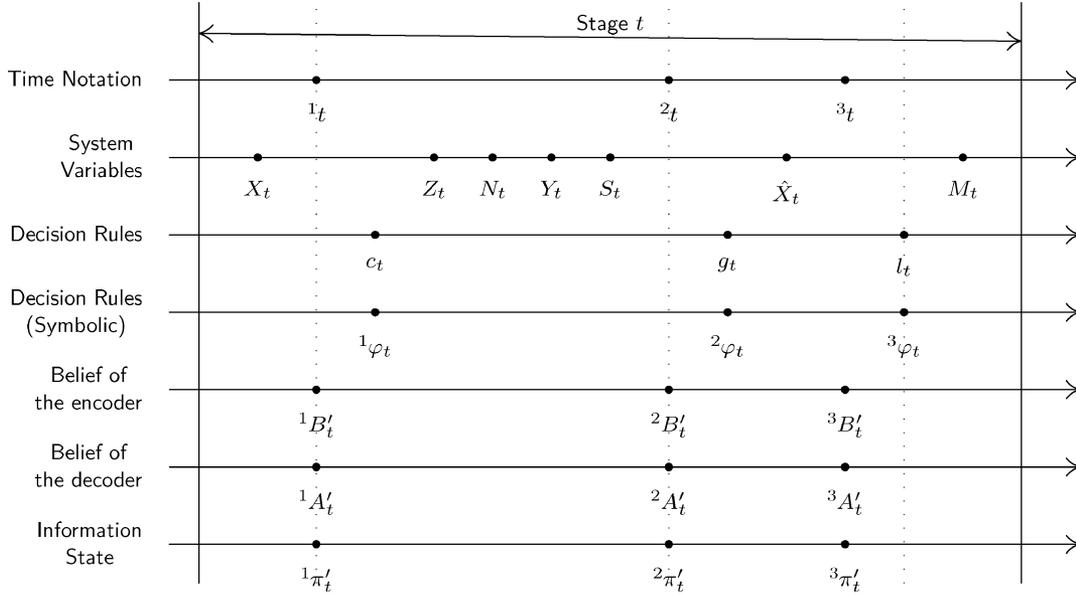


Fig. 3. The sequential ordering of the variables of the real time communication system for channels with memory.  $1_t$ ,  $2_t$ , and  $3_t$  are refinements of stage  $t$ .

## VII. CHANNELS WITH MEMORY

So far in this paper, we have assumed a memoryless channel. In realistic scenarios, the channel may have memory because of either changing physical environment (e.g., wireless fading channel), or inter symbol interference, or a combination of both. Such channels can be modeled as discrete channel with *state* (see [50] and [51]). In this section we extend our methodology for jointly optimal encoding, decoding and memory update for channels with memory.

### A. Problem Formulation

Consider the same problem as in Section II with one change. Instead of assuming the channel to be memoryless, assume that it has a finite state memory. The state  $S_t$  of the belongs to a finite set  $\mathcal{S}$ . The channel is described as

$$Y_t = h_t(Z_t, N_t, S_{t-1}) \quad (62)$$

where  $h_t(\cdot)$  denotes the channel function at time  $t$ ,  $N_t$  which belongs to  $\mathcal{N}$  denotes the channel noise at time  $t$ , and  $S_{t-1}$  which belongs to  $\mathcal{S}$  denotes the channel state at time  $t - 1$ . Assume that  $\{N_t, t = 1, \dots, T\}$  is a sequence of independent random variables and the PMF of  $N_t$  is  $P_{N_t}$ . Also assume that  $\{N_t, t = 1, \dots, T\}$  is independent of the source output  $\{X_t, t = 1, \dots, T\}$  and the initial state  $S_0$  of the channel. The channel state is updated according to

$$S_t = \hat{h}_t(Z_t, N_t, S_{t-1}) \quad (63)$$

where  $\hat{h}_t(\cdot)$  is the channel update function. Assume that the initial state  $S_0$  of the channel has distribution  $P_{S_0}$ .

The source, the encoder, the receiver, and the distortion models are the same as in Section II-A. The sequential ordering of the variables is shown in Fig. 3. We are interested in the following optimization problem.

**Problem 7.1:** Consider the real-time communication system of Section II-A with a channel with memory given by (62) and

(63). Choose a communication strategy  $(C^*, G^*, L^*)$  that is optimal with respect to the performance criterion of (5), i.e.

$$\mathcal{J}_T(C^*, G^*, L^*) = \mathcal{J}_T^* := \min_{\substack{C \in \mathcal{C}^T \\ G \in \mathcal{G}^T \\ L \in \mathcal{L}^T}} \mathcal{J}_T(C, G, L), \quad (64)$$

where  $\mathcal{C}^T$ ,  $\mathcal{G}^T$ , and  $\mathcal{L}^T$  are defined in Problem 2.1.

### B. Agents' Beliefs and Their Evolution

For this model, define information fields at the encoder and the receiver as in Definitions 2.1 and 2.2. For ease of notation, define  ${}^i S_t$  as the state of the channel at time  ${}^i t$ , i.e.

$${}^1 S_t = S_{t-1}, \quad {}^2 S_t = S_t, \quad {}^3 S_t = S_t. \quad (65)$$

The beliefs of the encoder and the receiver are modified to take the uncertainty of the channel state into account as follows.

**Definition 7.1 (Encoder's Beliefs):** Let  ${}^i B'_t$  denote the encoder's belief about the receiver's observation and the state of the channel at time  ${}^i t$ ,  $i = 1, 2, 3$ ,  $t = 1, \dots, T$ . Then for  ${}^i r \in {}^i \mathcal{R}$ , and  ${}^i s \in \mathcal{S}$

$${}^i B'_t({}^i r, {}^i s) := \Pr({}^i R_t = {}^i r, {}^i S_t = {}^i s | {}^i \mathcal{E}_t). \quad (66)$$

Let  ${}^i \mathcal{B}' = \mathcal{P}({}^i \mathcal{R} \times \mathcal{S})$  denote the space of realizations of  ${}^i \mathcal{B}_t$ .

**Definition 7.2 (Receiver's Beliefs):** Let  ${}^i A'_t$  denote the receiver's belief about the encoder's observations and the channel state at time  ${}^i t$ ,  $i = 1, 2, 3$ ,  $t = 1, \dots, T$ . Then for  ${}^i e_t \in {}^i \mathcal{E}_t$  and  ${}^i s \in \mathcal{S}$

$${}^i A'_t({}^i e_t, {}^i s) := \Pr({}^i E_t = {}^i e_t, {}^i S_t = {}^i s | {}^i \mathcal{R}_t). \quad (67)$$

Let  ${}^i \mathcal{A}'_t := \mathcal{P}({}^i \mathcal{E}_t \times \mathcal{S})$  denote the space of realizations of  ${}^i \mathcal{A}'_t$ . Furthermore, let  $\hat{A}_t$  denote the receiver's belief about the source output at time instant  ${}^2 t$ , given by (14).

The beliefs of the encoder evolve in a manner similar to Lemma 2.1.

*Lemma 7.1 (Evolution of the Encoder's Beliefs):* For each stage  $t$ , there exist deterministic functions  ${}^1F'_t$  and  ${}^3F'$  such that

$${}^2B'_t = {}^1F'_t({}^1B'_t, Z_t), \quad {}^3B'_t = {}^2B'_t, \quad {}^1B'_{t+1} = {}^3F'({}^3B'_t, l_t).$$

This is proved in Appendix C.

### C. Structural Results

In this section we derive qualitative properties of optimal encoders (respectively, decoders) that are true for any arbitrary but fixed decoding and memory update strategies (respectively, encoding and memory update strategies).

*Theorem 7.1:* Consider Problem 7.1 with arbitrary but fixed decoding and memory update strategies,  $G := (g_1, g_2, \dots, g_T)$  and  $L := (l_1, l_2, \dots, l_T)$ , respectively. Then there is no loss of optimality in restricting attention to encoding rules of the form

$$Z_t = c_t(X_t, {}^1B'_t), \quad t = 2, \dots, T. \quad (68)$$

We follow the methodology of the alternative proof of the structural results in [33].

*Proof:* Look at the problem from the encoder's point of view. The process  $\{X_t, t = 1, \dots, T\}$  is a Markov process independent of the noise in the forward channel. This fact together with the result of Lemma 7.1 imply that the process  $\{(X_t, {}^1B'_t), t = 1, \dots, T\}$  is a controlled Markov process with control action  $Z_t$ , i.e., for any  $x^{t+1} \in \mathcal{X}^{t+1}$ ,  ${}^1b'^{t+1} \in \mathcal{B}'^{t+1}$ ,  $z^t \in \mathcal{Z}^t$ , and any choice of  ${}^1\phi^t$

$$\begin{aligned} & \Pr(X_{t+1} = x_{t+1}, {}^1B'_{t+1} = {}^1b'_{t+1} \mid \\ & \quad X^t = x^t, {}^1B'^t = {}^1b'^t, Z^t = z^t; {}^1\phi^t) \\ &= \Pr({}^1B'_{t+1} = {}^1b'_{t+1} \mid \\ & \quad X^{t+1} = x^{t+1}, {}^1B'^t = {}^1b'^t, Z^t = z^t; {}^1\phi^t) \\ & \quad \times \Pr(X_{t+1} = x_{t+1} \mid X^t = x^t, {}^1B'^t = {}^1b'^t, Z^t = z^t; {}^1\phi^t) \\ & \stackrel{(a)}{=} \mathbb{1} [{}^1b'_{t+1} = {}^3F'({}^1F'_t({}^1b'_t, z_t), l_t)] P_{X_{t+1}|X_t}(x_{t+1} \mid x_t) \\ &= \Pr(X_{t+1} = x_{t+1}, {}^1B'_{t+1} = {}^1b'_{t+1} \mid \\ & \quad X_t = x_t, {}^1B'_t = {}^1b'_t, Z_t = z_t; l_t) \end{aligned}$$

where (a) follows from the Markov property of the source and Lemma 7.1. Thus, for a fixed memory update strategy  $L$ ,  $\{(X_t, {}^1B'_t), t = 1, \dots, T\}$  is a controlled Markov process with control action  $Z_t$ . Furthermore, the conditional expected instantaneous cost can be written as

$$\begin{aligned} & \mathbb{E} \left\{ \rho_t(X_t, \hat{X}_t) \mid {}^2\mathfrak{E}_t \right\} \\ &= \mathbb{E} \left\{ \rho_t(X_t, \hat{X}_t) \mid X^t, Z^t; c^t, g^t, l^{t-1} \right\} \\ &= \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \rho_t(X_t, g_t(y_t, m_{t-1})) \\ & \quad \times \Pr(Y_t = y_t, M_{t-1} = m_{t-1} \mid X^t, Z^t; c^t, g^t, l^{t-1}) \\ &= \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \rho_t(X_t, g_t(y_t, m_{t-1})) {}^3B'_t(y_t, m_{t-1}) \end{aligned}$$

$$\begin{aligned} & \stackrel{(b)}{=} \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \rho_t(X_t, g_t(y_t, m_{t-1})) {}^1F'_t({}^1B'_t, Z_t)(y_t, m_{t-1}) \\ &=: \tilde{\rho}_t(X_t, {}^1B'_t, Z_t, g_t) \end{aligned}$$

where (b) follows from Lemma 7.1. Thus, the total expected cost can be written as

$$\begin{aligned} & \mathbb{E} \left\{ \sum_{t=1}^T \rho_t(X_t, \hat{X}_t) \mid C, G, L \right\} \\ &= \mathbb{E} \left\{ \sum_{t=1}^T \mathbb{E} \left\{ \rho_t(X_t, \hat{X}_t) \mid {}^2\mathfrak{E}_t \right\} \mid C, G, L \right\} \\ &= \mathbb{E} \left\{ \sum_{t=1}^T \tilde{\rho}_t(X_t, {}^1B'_t, Z_t, g_t) \mid C, G, L \right\}. \quad (69) \end{aligned}$$

Hence, from the encoder's point of view,  $\{(X_t, {}^1B'_t), t = 1, \dots, T\}$  is a perfectly observed controlled Markov process with control action  $Z_t$  and an instantaneous cost  $\tilde{\rho}_t(X_t, {}^1B'_t, Z_t, g_t)$  (recall that  $G$  is fixed). From Markov decision theory [34, Chapter 6], we know that there is no loss of optimality in restricting attention to encoding rules of the form (68).

*Theorem 7.2:* Consider Problem 7.1 for any arbitrary but fixed encoding and memory update strategies,  $C := (c_1, c_2, \dots, c_T)$  and  $L := (l_1, l_2, \dots, l_T)$ , respectively. Then there is no loss of optimality in restricting attention to decoding rules of the form

$$\hat{X}_t = \hat{g}_t(\hat{A}_t) := \arg \min_{\hat{x} \in \hat{\mathcal{X}}} \sum_{x \in \mathcal{X}} \rho_t(x, \hat{x}) \hat{A}_t(x). \quad (70)$$

*Proof:* Look at the problem from the decoder's point of view. Since decoding is a filtering problem, minimizing the total distortion  $\mathcal{J}_T(C, G, L)$  is equivalent to minimizing the conditional expected instantaneous distortion  $\mathbb{E} \left\{ \rho_t(X_t, \hat{X}_t) \mid {}^2\mathfrak{R}_t \right\}$  for each time  $t$ . This conditional expected instantaneous distortion can be written as

$$\begin{aligned} \mathbb{E} \left\{ \rho_t(X_t, \hat{X}_t) \mid {}^2\mathfrak{R}_t \right\} &= \sum_{x_t \in \mathcal{X}} \rho_t(x_t, \hat{X}_t) \Pr(x_t \mid {}^2\mathfrak{R}_t) \\ &= \sum_{x_t \in \mathcal{X}} \rho_t(x_t, \hat{X}_t) \hat{A}_t(x_t) \end{aligned}$$

and is minimized by the decoding rule given in (70).

We can use the structural results of Theorem 7.1 to choose encoding rules from a space of functions that is not changing with time. Let  $\hat{\mathcal{C}}'$  denote the space of functions from  $\mathcal{X} \times {}^1\mathcal{B}'$  to  $\mathcal{Z}$ , and  $\hat{\mathcal{C}}'^T$  denote  $\hat{\mathcal{C}}' \times \dots \times \hat{\mathcal{C}}'$  ( $T$ -times). Then, the result of Theorem 7.1 implies the following.

*Corollary 7.1:* For Problem 7.1, the optimal performance  $\mathcal{J}_T^*$  can be determined by

$$\mathcal{J}_T^* := \inf_{\substack{C \in \hat{\mathcal{C}}'^T \\ G \in \mathcal{G}^T \\ L \in \mathcal{L}^T}} \mathcal{J}_T(C, G, L). \quad (71)$$

Consequently, Problem 7.1 can be reformulated in a manner similar to the reformulation presented in Section II-H.

*Problem 7.2:* Under the assumptions of Problem 7.1 choose a communication strategy  $(C^*, G^*, L^*)$  belonging to  $(\hat{\mathcal{C}}^{T} \times \mathcal{G}^T \times \mathcal{L}^T)$  that is optimal with respect to the performance criterion of (71).

#### D. Global Optimization

We use the structural results of Section VII-C to identify information states to obtain a sequential decomposition of Problem 7.2. This decomposition determines a globally optimal communication strategy for Problem 7.2 (and consequently for Problem 7.1). We follow the philosophy and approach of Sections III and IV. For that matter, we define the following.

*Definition 7.3:* Define  ${}^1\pi'$ ,  ${}^2\pi'$ , and  ${}^3\pi'$  as follows:

$${}^i\pi'_t = \Pr(X_t, {}^iB'_t | {}^i\phi^{t-1}), \quad i = 1, 2, 3. \quad (72)$$

Let  ${}^i\Pi'$  denote the space of probability measures on  $(\mathcal{X} \times {}^i\mathcal{B}')$ . Then  ${}^i\pi'_t$  takes values in  ${}^i\Pi'$ .

These probability measures are related to one another in a manner similar to Lemma 3.1. Consequently, Problem 7.2 can be reformulated in the same manner as Problem 2.2; furthermore, a sequential decomposition similar to the nested optimality equations of Theorem 3.1 can be obtained. This sequential decomposition can be extended to the infinite horizon along the lines of the results of Section IV. All these results can be derived along the lines of the analysis in Sections III and IV, so we omit the details.

### VIII. COMPUTATIONAL ISSUES

As mentioned earlier, all the variations of real-time communication problems considered in this paper are decentralized stochastic optimization problems. Consequently, they cannot be solved by classical Markov decision theory. However, the results presented in this paper show that an appropriate choice of information states can transform the decentralized stochastic optimization problem into a centralized deterministic optimization problem; albeit one where the objective is to choose an optimal function for each realization of the information state. Contrast this to the problems with classical information structure (which include Markov decision problems) where the objective is to choose an optimal action for each realization of the information state. This difference is the key reason for the difficulty in numerically solving these problems. Numerically solving decentralized stochastic control problems is known to be NEXP complete [52], i.e., it can be *proved* that, in general, these problems cannot be solved in polynomial time. The result of this paper shows that we can write sufficient conditions for finding optimal meta-strategies (the nested optimality conditions of Theorem 3.1 and the fixed point equations of Theorems 4.1 and 4.2) which are similar *in structure* to the sufficient conditions for finding optimal strategies in POMDPs (partially observable Markov decision processes). The information state in our decomposition is a probability vector on a finite dimensional real vector; thus it is the same as an information state for POMDPs where the unobserved state is a finite dimensional real vector. The action space in our decomposition is an uncountable

function space; thus it is similar to the action space of POMDPs with uncountable (Borelian) action spaces. Hence, one of the various approximation techniques for POMDPs [53]–[57] could be used to obtain an approximate numerical solution of the sequential decomposition presented in this paper. It could also be possible to break the curse of dimensionality by using randomization, or exploiting special structure, or taking advantage of the “knowledge capital” as explained in [58], [59].

When we move to the more general models of fixed-finite delay, higher-order Markov sources, or channels with memory, the above remarks remain valid: in the corresponding sequential decomposition the information state is still a probability measure on finite dimensional real vector and the action space is still an uncountable function space. However, all of these more general modeling assumptions increase the complexity of the solution because of the following two reasons.

- 1) The structure of optimal encoders has higher complexity, e.g., in problems with fixed-finite delay we can only restrict to encoders of the form

$$Z_t = c_t(X_{t-\delta}, \dots, X_t, {}^1B_t)$$

as compared to problems with zero-delay where we can restrict to

$$Z_t = c_t(X_t, {}^1B_t).$$

- 2) The information state is a probability measure on a higher dimensional space, e.g., in problems with fixed-finite delay  $\delta$ , the information state belongs to  $\mathcal{P}(\mathcal{X}^{\delta+1} \times {}^i\mathcal{B})$  as compared to problems with zero-delay where the information state belongs to  $\mathcal{P}(\mathcal{X} \times {}^i\mathcal{B})$ .

Both these factors make it computationally more difficult to solve the sufficient conditions for optimality derived in this paper.

It is important to remember that the high computational complexity is of the *off-line* computation of the optimal solution; the *on-line* implementation of the optimal solution is relatively straightforward.

### IX. COMPARISON WITH THE PHILOSOPHY OF INFORMATION THEORY AND CODING THEORY

This paper takes a drastically different approach to the design of a communication system than the traditional approach taken in information theory and coding theory. In this section we explain the reason for taking this different approach; we also explain the step that needs to be added to our approach in order to provide a complete solution methodology to determining good communication strategies for real-time communication systems.

When designing a communication system, we are interested in communication strategies that perform nearly optimally and are easy to implement. When there is no restriction on communication delay, information theory and coding theory break down the design of a communication system into two steps:

- 1) first, information theory is used to determine the fundamental limits of performance of a communication system;
- 2) then, coding theory investigates codes that are easy to implement and perform close to the fundamental performance limits determined by information theory.

However, when small communication delay is required, the above approach fails because the information theoretic bounds are, in general, not tight for small delays. Therefore, fundamental limits of performance for small delays are unknown.<sup>3</sup> Consequently, benchmarks for performance evaluation of communication strategies do not exist, and it is difficult to determine if particular family of codes performs close to optimal.

Given the current state of knowledge, one can take two approaches to the design of real-time communication systems: either determine tight bounds on optimal performance (and then find codes that come close to those bounds), or use some other technique to find good codes. In this paper we follow the second approach. We formulate the real-time communication problem as a decentralized stochastic optimization problem and develop a methodology to systematically search for an optimal communication strategy. This methodology exponentially simplifies the search for an optimal solution. In spite of this simplification, numerically solving the resultant optimality equations is a formidable task. We are not aware, at this point, of good approximation techniques for solving the optimality equations of Sections III and IV. (As pointed out in Section VIII, the approximation techniques for POMDP are an obvious candidate, but we do not know if it can be proved that they provide good approximations of the optimality equations of Sections III and IV). If such approximation techniques are discovered, only then would the results of this paper along with those techniques provide a complete methodology to determining communication strategies that perform well for small delays.

## X. CONCLUSION

Real-time communication is a notoriously hard problem: the difficulties are both conceptual and computational. In this paper we present a conceptual framework to study real-time communication. This framework is based on the notions of information structure and information state, and provides a systematic way of searching for an optimal real-time communication strategy. The framework is fairly general: we show that it is applicable to finite and infinite horizon zero-delay communication systems and can be extended to fixed-finite delay communication systems, to systems where the source statistics are higher-order Markov, and to systems where the channels have memory. Thus, this framework provides a unified method of investigating different variations of real-time communication.

The conceptual results presented in this paper exponentially simplify the computational complexity of the problem. In spite of this simplification, numerically determining globally optimal communication strategies is a formidable task. Furthermore, finding computationally efficient algorithms to approximately solve the optimality equations of this paper is a difficult unsolved problem.

The approach taken in this paper is similar in spirit to Witsenhausen's standard form [35]. In our solution, we exploit the structural results of [33] to identify information states that belong to a space that does not increase with time. This is in

<sup>3</sup>Recently, nonasymptotic upper and lower bounds for channel coding have been developed [60], [61]. However, corresponding bounds for joint source-channel coding remain unknown.

contrast to the information states in [35] which belong to a space that increases with time. This feature allows us to extend our methodology to infinite horizon problem; in contrast, the standard form is applicable to only finite horizon problems. It is worth noting that for infinite horizon problems, stationary (time-invariant) communication strategies are not necessarily optimal.

The methodology presented in this paper can be used, in principle, for arbitrary values of acceptable communication delay. However, the increase in computational complexity with the increase in delay implies that the methodology presented in this paper can only be used for applications where the acceptable delay is small. Information theory, on the other hand, provides tight performance bounds for applications where the acceptable communication delay is large. Finding a methodology for communication problems where the acceptable delay is medium (i.e., the delay is large enough to make the framework presented in this paper computationally intractable, but small enough so that the asymptotic laws of probability are not applicable) remains a challenging open problem.

## APPENDIX A

### RELATION BETWEEN THE BELIEFS

*Proof of Lemma 2.1:* We prove the three results separately.

- 1) Consider any  ${}^2e_t = ({}^1e_t, z_t) \in {}^2\mathcal{E}_t$ ,  ${}^2r_t = (y_t, m_{t-1}) \in {}^2\mathcal{R}_t$ ,  ${}^2\phi^{t-1} = ({}^1\phi^{t-1}, c_t)$ . Then

$$\begin{aligned} {}^2b_t({}^2r_t) &= \Pr({}^2R_t = {}^2r_t \mid {}^2E_t = {}^2e_t; {}^2\phi^{t-1}) \\ &= \Pr(Y_t = y_t, M_{t-1} = m_{t-1} \mid \\ &\quad {}^1E_t = {}^1e_t, Z_t = z_t; {}^1\phi^{t-1}, c_t) \\ &= \Pr(Y_t = y_t \mid M_{t-1} = m_{t-1}, {}^1E_t = {}^1e_t, \\ &\quad Z_t = z_t; {}^1\phi^{t-1}, c_t) \\ &\quad \times \Pr(M_{t-1} = m_{t-1} \mid {}^1E_t = {}^1e_t, Z_t = z_t; \\ &\quad {}^1\phi^{t-1}, c_t) \\ &\stackrel{(a)}{=} P_{N_t}(n_t \in \mathcal{N} : y_t = h_t(z_t, n_t)) \\ &\quad \times \Pr(M_{t-1} = m_{t-1} \mid {}^1E_t = {}^1e_t; {}^1\phi^{t-1}) \\ &= P_{N_t}(n_t \in \mathcal{N} : y_t = h_t(z_t, n_t)) \times {}^1b_t(m_{t-1}) \\ &=: {}^1F_t({}^1b_t, z_t)(y_t, m_{t-1}) \\ &= {}^1F_t({}^1b_t, z_t)({}^2r_t) \end{aligned} \quad (73)$$

where (a) follows from the sequential order in which the system variables are generated. Observe that the dependence of  ${}^2F_t(\cdot)$  on  $t$  is through the dependence of  $h_t(\cdot)$  and the noise statistics  $P_{N_t}$  on  $t$ .

- 2) Consider any  ${}^3e_t \in {}^3\mathcal{E}_t$ ,  ${}^3r_t \in {}^3\mathcal{R}_t$ , and  ${}^3\phi^{t-1} = ({}^2\phi^{t-1}, g_t)$ . Recall that  ${}^3E_t = {}^2E_t$  and  ${}^3R_t = {}^2R_t$ . Then

$$\begin{aligned} {}^3b_t({}^3r_t) &= \Pr({}^3R_t = {}^3r_t \mid {}^3E_t = {}^3e_t; {}^3\phi^{t-1}) \\ &= \Pr({}^2R_t = {}^3r_t \mid {}^2E_t = {}^3e_t; {}^2\phi^{t-1}, g_t) \\ &\stackrel{(b)}{=} \Pr({}^2R_t = {}^3r_t \mid {}^2E_t = {}^3e_t; {}^2\phi^{t-1}, g_t) \\ &= \Pr({}^2R_t = {}^3r_t \mid {}^2E_t = {}^3e_t; {}^2\phi^{t-1}) \\ &= {}^2b_t({}^3r_t) \end{aligned} \quad (74)$$

where (b) follows from the sequential order in which the system variables are generated.

3) Consider any  ${}^1e_{t+1} = ({}^3e_t, x_{t+1}) \in {}^1\mathcal{E}_{t+1}$ ,  ${}^1r_{t+1} = m_t \in {}^1\mathcal{R}$ , and  ${}^1\phi^t = ({}^3\phi^{t-1}, l_t)$ . Then

$$\begin{aligned}
& {}^1b_{t+1}({}^1r_{t+1}) \\
&= \Pr({}^1R_{t+1} = {}^1r_{t+1} \mid {}^1E_{t+1} = {}^1e_{t+1}; {}^1\phi^t) \\
&= \Pr(M_t = m_t \mid {}^3E_t = {}^3e_t, X_{t+1} = x_{t+1}; {}^3\phi^{t-1}, l_t) \\
&= \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \Pr(Y_t = y_t, M_t = m_t, M_{t-1} = m_{t-1} \mid \\
&\quad {}^3E_t = {}^3e_t, X_{t+1} = x_{t+1}; {}^3\phi^{t-1}, l_t) \\
&= \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \Pr(M_t = m_t \mid Y_t = y_t, M_t = m_{t-1}, \\
&\quad {}^3E_t = {}^3e_t, X_{t+1} = x_{t+1}; {}^3\phi^{t-1}, l_t) \\
&\quad \times \Pr(Y_t = y_t, M_{t-1} = m_{t-1} \mid \\
&\quad {}^3E_t = {}^3e_t, X_{t+1} = x_{t+1}; {}^3\phi^{t-1}, l_t) \\
&\stackrel{(c)}{=} \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \mathbf{1}[m_t = l_t(y_t, m_{t-1})] \\
&\quad \times \Pr(Y_t = y_t, M_{t-1} = m_{t-1} \mid {}^3E_t = {}^3e_t; {}^3\phi^{t-1}) \\
&= \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \mathbf{1}[m_t = l_t(y_t, m_{t-1})] {}^3b_t(y_t, m_{t-1}) \\
&=: {}^3F({}^3b_t, l_t)(m_t) \\
&=: {}^3F({}^3b_t, l_t)({}^1r_{t+1}) \tag{75}
\end{aligned}$$

where (c) follows from the sequential order in which the system variables are generated.

## APPENDIX B

### RELATION BETWEEN INFORMATION STATES

*Proof of Lemma 3.1:* We prove the three results separately.

1) Consider any  $x_t \in \mathcal{X}$ ,  ${}^2b_t \in {}^2\mathcal{B}$ , and  ${}^2\phi^{t-1} = ({}^1\phi^{t-1}, c_t)$ . A component of  ${}^2\pi_t$  is given by

$$\begin{aligned}
& {}^2\pi_t(x_t, {}^2b_t) = \Pr(X_t = x_t, {}^2B_t = {}^2b_t \mid {}^2\phi^{t-1}) \\
&= \int \sum_{\substack{z_t \in \mathcal{Z} \\ {}^1b_t \in {}^1\mathcal{B}}} \Pr(X_t = x_t, Z_t = z_t, {}^2B_t = {}^2b_t, \\
&\quad {}^1B_t = {}^1b_t \mid {}^1\phi^{t-1}, c_t) d{}^1b_t \\
&= \int \sum_{\substack{z_t \in \mathcal{Z} \\ {}^1b_t \in {}^1\mathcal{B}}} \Pr({}^2B_t = {}^2b_t \mid X_t = x_t, Z_t = z_t, \\
&\quad {}^1B_t = {}^1b_t; {}^1\phi^{t-1}, c_t) d{}^1b_t \\
&\quad \times \Pr(Z_t = z_t \mid X_t = x_t, {}^1B_t = {}^1b_t; {}^1\phi^{t-1}, c_t) \\
&\quad \times \Pr(X_t = x_t, {}^1B_t = {}^1b_t \mid {}^1\phi^{t-1}, c_t) d{}^1b_t \\
&\stackrel{(a)}{=} \int \sum_{\substack{z_t \in \mathcal{Z} \\ {}^1b_t \in {}^1\mathcal{B}}} \mathbf{1}[{}^2b_t = {}^1F_t({}^1b_t, z_t)] \\
&\quad \times \mathbf{1}[z_t = c_t(x_t, {}^1b_t)] \\
&\quad \times \Pr(X_t = x_t, {}^1B_t = {}^1b_t \mid {}^1\phi^{t-1}) d{}^1b_t \\
&= \int \sum_{\substack{z_t \in \mathcal{Z} \\ {}^1b_t \in {}^1\mathcal{B}}} \mathbf{1}[{}^2b_t = {}^1F_t({}^1b_t, z_t)] \\
&\quad \times \mathbf{1}[z_t = c_t(x_t, {}^1b_t)] {}^1\pi_t(x_t, {}^1b_t) d{}^1b_t \\
&=: ({}^1Q(c_t) {}^1\pi_t)(x_t, {}^2b_t) \tag{76}
\end{aligned}$$

where (a) follows from the sequential order in which the system variables are generated.

2) Consider  $x_t \in \mathcal{X}$ ,  ${}^3b_t \in {}^3\mathcal{B}$ , and  ${}^3\phi^{t-1} = ({}^2\phi^{t-1}, g_t)$ . Then a component of  ${}^3\pi_t$  is given by

$$\begin{aligned}
& {}^3\pi_t(x_t, {}^3b_t) = \Pr(X_t = x_t, {}^3B_t = {}^3b_t \mid {}^3\phi^{t-1}) \\
&= \Pr(X_t = x_t, {}^2B_t = {}^3b_t \mid {}^2\phi^{t-1}, g_t) \\
&\stackrel{(b)}{=} \Pr(X_t = x_t, {}^2B_t = {}^3b_t \mid {}^2\phi^{t-1}) \\
&= {}^2\pi_t(x_t, {}^3b_t) \tag{77}
\end{aligned}$$

where (b) follows from the sequential order in which the system variables are generated.

3) Consider  $x_{t+1} \in \mathcal{X}$ ,  ${}^1b_{t+1} \in {}^1\mathcal{B}$ , and  ${}^1\phi^t = ({}^3\phi^{t-1}, l_t)$ . Then a component of  ${}^1\pi_{t+1}$  is given by

$$\begin{aligned}
& {}^1\pi_{t+1}(x_{t+1}, {}^1b_{t+1}) \\
&= \Pr(X_{t+1} = x_{t+1}, {}^1B_{t+1} = {}^1b_{t+1} \mid {}^1\phi^t) \\
&= \int \sum_{\substack{x_t \in \mathcal{X} \\ {}^3b_t \in {}^3\mathcal{B}}} \Pr(X_{t+1} = x_{t+1}, X_t = x_t, {}^3B_t = {}^3b_t, \\
&\quad {}^1B_{t+1} = {}^1b_{t+1}, \mid {}^3\phi^{t-1}, l_t) d{}^3b_t \\
&\stackrel{(c)}{=} \int \sum_{\substack{x_t \in \mathcal{X} \\ {}^3b_t \in {}^3\mathcal{B}}} P_{X_{t+1}|X_t}(x_{t+1} \mid x_t) \\
&\quad \times \mathbf{1}[{}^1b_{t+1} = {}^3F({}^3b_t, l_t)] \\
&\quad \times \Pr(X_t = x_t, {}^3B_t = {}^3b_t \mid {}^3\phi^{t-1}) d{}^3b_t \\
&= \int \sum_{\substack{x_t \in \mathcal{X} \\ {}^3b_t \in {}^3\mathcal{B}}} P_{X_{t+1}|X_t}(x_{t+1} \mid x_t) \\
&\quad \times \mathbf{1}[{}^1b_{t+1} = {}^3F({}^3b_t, l_t)] {}^3\pi_t(x_t, {}^3b_t) d{}^3b_t \\
&=: ({}^3Q_t(l_t) {}^3\pi_t)(x_{t+1}, {}^1b_{t+1}) \tag{78}
\end{aligned}$$

where (c) follows from the sequential order in which the system variables are generated.

4) Consider  ${}^2r_t = (y_t, m_{t-1}) \in {}^2\mathcal{R}$ . Then

$$\begin{aligned}
& \mathbb{E} \left\{ \rho_t(X_t, \hat{X}_t) \mid {}^2\phi^{t-1}, g_t \right\} = \sum_{\substack{x_t \in \mathcal{X} \\ {}^2r_t \in {}^2\mathcal{R}_t}} \rho_t(x_t, g_t({}^2r_t)) \\
&\quad \times \Pr(X_t = x_t, {}^2R_t = {}^2r_t \mid {}^2\phi^{t-1}, g_t). \tag{79}
\end{aligned}$$

Now consider

$$\begin{aligned}
& \Pr(X_t = x_t, {}^2R_t = {}^2r_t \mid {}^2\phi^{t-1}, g_t) \\
&= \int \Pr(X_t = x_t, {}^2R_t = {}^2r_t, {}^2B_t = {}^2b_t \mid \\
&\quad {}^2\phi^{t-1}, g_t) d{}^2b_t \\
&= \int \Pr({}^2R_t = {}^2r_t \mid X_t = x_t, {}^2B_t = {}^2b_t; {}^2\phi^{t-1}, g_t) \\
&\quad \times \Pr(X_t = x_t, {}^2B_t = {}^2b_t \mid {}^2\phi^{t-1}, g_t) d{}^2b_t \\
&\stackrel{(d)}{=} \int {}^2b_t({}^2r_t) \Pr(X_t = x_t, {}^2B_t = {}^2b_t \mid {}^2\phi^{t-1}) d{}^2b_t \\
&= \int {}^2b_t({}^2r_t) {}^2\pi_t(x_t, {}^2b_t) d{}^2b_t. \tag{80}
\end{aligned}$$

Substituting the result of (80) in (79) we get

$$\begin{aligned} & \mathbb{E} \left\{ \rho_t(X_t, \hat{X}_t) \middle| 2\phi^{t-1}, g_t \right\} \\ &= \sum_{\substack{x_t \in \mathcal{X} \\ 2r_t \in 2\mathcal{R}_t}} \rho_t(x_t, g_t(2r_t)) \int_{b_t \in \mathcal{B}_t} 2b_t(2r_t) 2\pi_t(x_t, 2b_t) d^2b_t \\ &=: \hat{\rho}_t(2\pi_t, g_t). \end{aligned} \quad (81)$$

Observe that (76) and (78) imply that the transformations  $1Q_t$  and  $3Q_t$  are linear in the sense that if  $1\pi_t^{(1)}, 1\pi_t^{(2)} \in 1\Pi_t$ ,  $c_t \in \hat{\mathcal{C}}$  and  $\lambda \in [0, 1]$ , then

$$\begin{aligned} & 1Q_t(c_t)(\lambda 1\pi_t^{(1)} + (1-\lambda) 1\pi_t^{(2)}) \\ &= \lambda 1Q_t(c_t) 1\pi_t^{(1)} + (1-\lambda) 1Q_t(c_t) 1\pi_t^{(2)} \end{aligned} \quad (82)$$

and similar relation holds for  $3Q_t$ .

### APPENDIX C

#### RELATION BETWEEN THE BELIEFS FOR CHANNELS WITH MEMORY

*Proof of Lemma 7.1:* We prove the three results separately.

1) Consider any  $2e_t = (1e_t, z_t) \in 2\mathcal{E}_t$ ,  $2r_t = (y_t, m_{t-1}) \in 2\mathcal{R}$ ,  $s_t \in \mathcal{S}$ , and  $2\phi^{t-1} = (1\phi^{t-1}, c_t)$ . Then

$$\begin{aligned} & 2b'_t(2r_t, s_t) \\ &= \Pr(2R_t = 2r_t, 2S_t = s_t \mid 2E_t = 2e_t, 2\phi^{t-1}) \\ &= \Pr(Y_t = y_t, M_{t-1} = m_{t-1}, S_t = s_t \mid \\ & \quad 1E_t = 1e_t, Z_t = z_t; 1\phi^{t-1}, c_t) \\ &= \sum_{s_{t-1} \in \mathcal{S}} \Pr(Y_t = y_t, M_{t-1} = m_{t-1}, S_{t-1} = s_{t-1} \mid \\ & \quad S_t = s_t \mid 1E_t = 1e_t, Z_t = z_t; 1\phi^{t-1}, c_t) \\ &= \sum_{s_{t-1} \in \mathcal{S}} \Pr(Y_t = y_t, S_t = s_t \mid 1E_t = 1e_t, Z_t = z_t, \\ & \quad S_{t-1} = s_{t-1}, M_{t-1} = m_{t-1}; 1\phi^{t-1}, c_t) \\ & \quad \times \Pr(M_{t-1} = m_{t-1}, S_{t-1} = s_{t-1} \mid \\ & \quad 1E_t = 1e_t, Z_t = z_t; 1\phi^{t-1}, c_t) \\ &\stackrel{(a)}{=} \sum_{s_{t-1} \in \mathcal{S}} P_{N_t}(n_t \in \mathcal{N} : y_t = h_t(z_t, n_t, s_{t-1}), \\ & \quad s_t = \hat{h}_t(z_t, n_t, s_{t-1})) \\ & \quad \times \Pr(M_{t-1} = m_{t-1}, S_{t-1} = s_{t-1} \mid \\ & \quad 1E_t = 1e_t, Z_t = z_t; 1\phi^{t-1}) \\ &= \sum_{s_{t-1} \in \mathcal{S}} P_{N_t}(n_t \in \mathcal{N} : y_t = h_t(z_t, n_t, s_{t-1}), \\ & \quad s_t = \hat{h}_t(z_t, n_t, s_{t-1})) \\ & \quad \times 1b'_t(m_{t-1}, s_{t-1}) \\ &=: 1F'_t(1b'_t, z_t)(y_t, m_{t-1}) \\ &= 1F'_t(1b'_t, z_t)(2r_t). \end{aligned} \quad (83)$$

where (a) follows from the sequential order in which the system variables are generated.

2) Consider any  $3e_t \in 3\mathcal{E}_t$ ,  $3r_t \in 3\mathcal{R}$ ,  $s_t \in \mathcal{S}$ , and  $3\phi^{t-1} = (2\phi^{t-1}, g_t)$ . Recall that  $3E_t = 2E_t$ ,  $3R_t = 2R_t$ , and  $3S_t = 2S_t$ . Then

$$\begin{aligned} & 3b'_t(3r_t, s_t) \\ &= \Pr(3R_t = 3r_t, 3S_t = s_t \mid 3E_t = 3e_t; 3\phi^{t-1}) \\ &= \Pr(2R_t = 3r_t, 2S_t = s_t \mid 3E_t = 3e_t; 2\phi^{t-1}, g_t) \\ &\stackrel{(b)}{=} \Pr(2R_t = 3r_t, 2S_t = s_t \mid 3E_t = 3e_t; 2\phi^{t-1}) \\ &= 2b'_t(3r_t, s_t). \end{aligned} \quad (84)$$

where (b) follows from the sequential order in which the system variables are generated.

3) Consider any  $1e_{t+1} = (3e_t, x_{t+1}) \in 3\mathcal{E}_{t+1}$ ,  $1r_{t+1} = m_t \in 1\mathcal{R}$ ,  $s_t \in \mathcal{S}$ , and  $1\phi^t = (3\phi^{t-1}, l_t)$ . Recall that  $1S_{t+1} = 3S_t = S_t$ . Then

$$\begin{aligned} & 1b'_{t+1}(1r_{t+1}, s_t) \\ &= \Pr(1R_{t+1}, 1S_{t+1} = s_t \mid 1E_{t+1} = 1e_{t+1}; 1\phi^t) \\ &= \Pr(M_t = m_t, 3S_t = s_t \mid 3E_t = 3e_t, \\ & \quad X_{t+1} = x_{t+1}; 3\phi^{t-1}, l_t) \\ &= \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \Pr(Y_t = y_t, M_t = m_t, M_{t-1} = m_{t-1}, \\ & \quad 3S_t = s_t \mid 3E_t = 3e_t, X_{t+1} = x_{t+1}; \\ & \quad 3\phi^{t-1}, l_t) \\ &= \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \Pr(M_t = m_t \mid Y_t = y_t, M_{t-1} = m_{t-1}, \\ & \quad 2S_t = s_t, 3E_t = 3e_t, X_{t+1} = x_{t+1}; \\ & \quad 3\phi^{t-1}, l_t) \\ & \quad \times \Pr(Y_t = y_t, M_{t-1} = m_{t-1}, 2S_t = s_t \mid \\ & \quad 3E_t = 3e_t, X_{t+1} = x_{t+1}; 3\phi^{t-1}, l_t) \\ &\stackrel{(c)}{=} \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \mathbf{1}[m_t = l_t(y_t, m_{t-1})] \\ & \quad \times \Pr(Y_t = y_t, M_{t-1} = m_{t-1}, 2S_t = s_t \mid \\ & \quad 3E_t = 3e_t; 3\phi^{t-1}) \\ &= \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \mathbf{1}[m_t = l_t(y_t, m_{t-1})] 3b'_t(y_t, m_{t-1}, s_t) \\ &=: 3F'(3b'_t, l_t)(m_t) \\ &= 3F'(3b'_t, l_t)(1r_{t+1}) \end{aligned} \quad (85)$$

where (c) follows from the sequential order in which the system variables are generated.

### ACKNOWLEDGMENT

The authors are grateful to A. Anastasopoulos, S. Pradhan, and J. Lagarias for insightful discussions and to Serdar Yüksel for useful feedback. They are also grateful to the reviewers for their detailed and careful comments that helped us to significantly improve the presentation of the paper.

### REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 22, pp. 379–423, Jul. 1948.
- [2] N. T. Gaarder and D. Slepian, "On optimal finite-state digital transmission systems," in *Proc. Int. Symp. Inf. Theory*, Grignano, Italy, Jun. 1979.
- [3] N. T. Gaarder and D. Slepian, "On optimal finite-state digital transmission systems," *IEEE Trans. Inf. Theory*, vol. 28, pp. 167–186, 1982.
- [4] T. Linder and G. Lugosi, "A zero-delay sequential scheme for lossy coding of individual sequences," *IEEE Trans. Inf. Theory*, vol. 47, pp. 2533–2538, 2001.
- [5] T. Weissman and N. Merhav, "On limited-delay lossy coding and filtering of individual sequences," *IEEE Trans. Inf. Theory*, vol. 48, pp. 721–733, 2002.

- [6] A. György, T. Linder, and G. Lugosi, "Efficient adaptive algorithms and minimax bounds for zero-delay lossy source coding," *IEEE Trans. Signal Processing*, vol. 52, pp. 2337–2347, 2004.
- [7] S. Matloub and T. Weissman, "Universal zero-delay joint source–channel coding," *IEEE Trans. Inf. Theory*, vol. 52, pp. 5240–5250, Dec. 2006.
- [8] S. P. Lloyd, "Rate versus fidelity for binary source," *Bell Syst. Tech. J.*, vol. 56, pp. 427–437, 1977.
- [9] P. Piret, "Causal sliding block encoders with feedback," *IEEE Trans. Inf. Theory*, vol. 25, pp. 237–240, Mar. 1979.
- [10] D. L. Neuhoff and R. K. Gilbert, "Causal source codes," *IEEE Trans. Inf. Theory*, vol. 28, pp. 701–713, Sep. 1982.
- [11] T. Linder and R. Zamir, "Causal coding of stationary sources and individual sequences with high resolution," *IEEE Trans. Inf. Theory*, vol. 52, pp. 662–680, Feb. 2006.
- [12] T. Fine, "Properties of an optimum digital system and applications," *IEEE Trans. Inf. Theory*, vol. 10, pp. 287–296, Oct. 1964.
- [13] B. McMillan, "Communication systems which minimize coding noise," *Bell Syst. Tech. J.*, vol. 48, no. 9, pp. 3091–3113, Sep. 1969.
- [14] T. Ericson, Delayless Information Transmission Linköping University, Department of Elec. Eng., Linköping, Sweden, Internal Publication LiTH-ISY-I-0260, 1979.
- [15] T. Ericson, "A result on delayless information transmission," in *International Symposium on Information Theory*, Grignano, Italy, Jun. 1979.
- [16] H. S. Witsenhausen, "Informational aspects of stochastic control," in *Proc. Oxford Confe. Stoch. Optim.*, 1978.
- [17] D. Teneketzis, "Communication in Decentralized Control," Ph.D. dissertation, Department of EECS, Massachusetts Inst. Technol., Cambridge, MA, 1979.
- [18] G. Munson, "Causal Information Transmission with Feedback," Ph.D. dissertation, Department of Electrical Engineering, Cornell University, Ithaca, NY, 1981.
- [19] A. K. Gorbunov and P. S. Pinsker, "Non-anticipatory and prognostic epsilon entropies and message generation rates," *Probl. Inf. Trans.*, vol. 9, pp. 184–191, 1973.
- [20] A. K. Gorbunov and P. S. Pinsker, "Prognostic epsilon entropy of a Gaussian message and a Gaussian source," *Probl. Inf. Trans.*, vol. 10, pp. 93–109, 1974.
- [21] P. S. Pinsker and A. K. Gorbunov, "Delayed epsilon entropy with small mean-square reproduction error," *Probl. Inf. Transm.*, vol. 23, no. 2, pp. 91–95, Apr.–Jun. 1987.
- [22] S. Engell, "New results on the real-time transmission problem," *IEEE Trans. Inf. Theory*, vol. 33, pp. 210–218, Mar. 1987.
- [23] Y.-C. Ho, M. Kastner, and E. Wong, "Teams, signaling, and information theory," *IEEE Trans. Autom. Control*, vol. 23, pp. 305–312, 1978.
- [24] N. Merhav and I. Kontoyiannis, "Source coding exponents for zero-delay coding with finite memory," *IEEE Trans. Inf. Theory*, vol. 49, pp. 609–625, 2003.
- [25] A. W. Drake, "Observation of A Markov Process Through a Noisy Channel," Sc.D. dissertation, Department of Elec. Eng., Massachusetts Inst. Technol., Cambridge, MA, 1962.
- [26] J. Devore, "A note on the observation of a Markov source through a noisy channel," *IEEE Trans. Inf. Theory*, vol. 20, pp. 762–764, 1974.
- [27] H. S. Witsenhausen, "On the structure of real-time source coders," *Bell Syst. Tech. J.*, vol. 58, no. 6, pp. 1437–1451, Jul.–Aug. 1979.
- [28] V. S. Borkar, S. K. Mitter, and S. Tatikonda, "Optimal sequential vector quantization of Markov sources," *SIAM J. Opt. Contr.*, vol. 40, no. 1, pp. 135–148, Jan. 2001.
- [29] J. C. Walrand and P. Varaiya, "Optimal causal coding–decoding problems," *IEEE Trans. Inf. Theory*, vol. 29, pp. 814–820, Nov. 1983.
- [30] R. S. Lipster and A. N. Shiryaev, *Statistics of Random Processes*. New York: Springer-Verlag, 1977, vol. II, Applications.
- [31] T. Başar and R. Bansal, "Optimal design of measurement channels and control policies for linear-quadratic stochastic systems," *Europ. J. Oper. Res.*, vol. 93, pp. 226–236, 1994.
- [32] R. Bansal and T. Başar, "Simultaneous design of measurement channels and control strategies for stochastic systems with feedback," *Automatica*, vol. 25, pp. 679–694, 1989.
- [33] D. Teneketzis, "On the structure of optimal real-time encoders and decoders in noisy communication," *IEEE Trans. Inf. Theory*, pp. 4017–4035, Sep. 2006.
- [34] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation Identification and Adaptive Control*. Englewood Cliffs, NJ: Prentice Hall, 1986.
- [35] H. S. Witsenhausen, "A standard form for sequential stochastic control," *Mathematical Systems Theory*, vol. 7, no. 1, pp. 5–11, 1973.
- [36] S. Thrun, "Monte Carlo POMDPs," in *Advances in Neural Information Processing Systems*, S. Solla, T. Leen, and K.-R. Müller, Eds. Cambridge, MA: MIT Press, 2000, vol. 12, pp. 1064–1070.
- [37] J. M. Porta, N. Vlassis, M. T. J. Spaan, and P. Poupart, "Point-based value iteration for continuous POMDPs," *J. Mach. Learn. Res.*, vol. 7, pp. 2329–2367, Nov. 2006.
- [38] A. Brooks, A. Makarenkoa, S. Williamsa, and H. Durrant-Whytea, "Parametric POMDPs for planning in continuous state spaces," *Robot. Auton. Syst.*, no. 11, pp. 887–897, Nov. 2006.
- [39] E. Brunskill, L. Kaelbling, T. Lozano-Perez, and N. Roy, "Continuous-state POMDPs with hybrid dynamics," in *Proc. Tenth Int. Symp. Artificial Intell. Math.*, Fort Lauderdale, FL, 2008.
- [40] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*. New York: Oxford University Press, 2001.
- [41] J. Marschak and R. Radner, *Economic Theory of Teams*. New Haven, CT: Yale University Press, 1972.
- [42] Y.-C. Ho, "Team decision theory and information structures," *Proc. IEEE*, vol. 68, pp. 644–654, 1980.
- [43] H. S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proc. IEEE*, vol. 59, pp. 1557–1566, Nov. 1971.
- [44] H. S. Witsenhausen, "Some remarks on the concept of state," in *Directions in Large-Scale Systems*, Y. C. Ho and S. K. Mitter, Eds. New York: Plenum, 1976, pp. 69–75.
- [45] R. J. Aumann, "Agreeing to disagree," *Ann. Stat.*, no. 4, pp. 1236–39, 1976.
- [46] E. B. Dynkin and A. A. Yushkevich, *Controlled Markov Processes*. : Springer-Verlag, 1975, A Series of Comprehensive Studies in Mathematics.
- [47] P. Whittle, *Optimization Over Time*. New York: Wiley, 1983, vol. 2, Wiley Series in Probability and Mathematical Statistics.
- [48] A. Arapostathis, V. S. Borkar, E. Fernandez-Gaucherand, M. K. Ghosh, and S. I. Marcus, "Discrete-time controlled Markov processes with average cost criterion – A survey," *SIAM J. Contr. Optim.*, vol. 31, no. 2, pp. 282–344, Mar. 1993.
- [49] O. Hernández-Lerma, *Adaptive Markov Control Processes*. New York: Springer-Verlag, 1989.
- [50] Q. Zhang and S. A. Kassam, "Finite-state Markov model for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 1688–1692, 1999.
- [51] G. Forney Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inf. Theory*, vol. 18, pp. 363–378, 1972.
- [52] D. S. Bernstein, S. Zilberstein, and N. Immerman, "The complexity of decentralized control of Markov decision processes," in *Proc. 16th Int. Conf. Uncertainty Artif. Intell. (UAI)*, Stanford, CA, Jun. 2000, pp. 32–27.
- [53] W. S. Lovejoy, "A survey of partially observable Markov decision processes," *Ann. Oper. Res.*, vol. 28, no. 1, pp. 47–65, 1991.
- [54] R. Bansal and T. Basar, "Algorithms for partially observed Markov decision processes," *Automatica*, vol. 25, no. 5, pp. 679–694, 1989.
- [55] C. C. White, III and W. T. Scherer, "Solution procedures for partially observed Markov decision processes," *Oper. Res.*, vol. 37, no. 5, pp. 791–797, 1989.
- [56] C. C. White, III, "Partially observed Markov decision processes: A survey," *Ann. Oper. Res.*, vol. 32, 1991.
- [57] W. Zhang, "Algorithms for Partially Observed Markov Decision Processes," Ph.D. dissertation, Hong Kong University of Science and Technology, Hong Kong, 2001.
- [58] J. Rust, "Using randomization to break the curse of dimensionality," *Econometrica*, vol. 65, no. 3, pp. 487–516, 1997.
- [59] J. Rust, "Dealing with the complexity of economic calculations," in *Fundamental Limits to Knowledge in Economics*. Santa Fe, NM: Santa Fe Institute, 1996.
- [60] Y. Polyanskiy, H. V. Poor, and S. Verdú, "New channel coding achievability bounds," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 6–10, 2008, pp. 1763–1767.
- [61] L. Wang, R. Colbeck, and R. Renner, "Simple channel coding bounds," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, South Korea, Jul. 3, 2009, pp. 1804–1808.

**Aditya Mahajan** (S'06–M'09) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Kanpur, India in 2003 and the M.S. and Ph.D. degrees in electrical engineering and computer science from the University of Michigan, Ann Arbor, MI, in 2006 and 2008.

He is currently a Postdoctoral Associate in Electrical Engineering at Yale University, New Haven, CT. His research interests include decentralized stochastic control, team theory, real-time communication, information theory, and discrete event systems.

**Demosthenis Teneketzis** (M'87–SM'97–F'00) received the diploma in electrical engineering from the University of Patras, Patras, Greece, and the M.S., E.E., and Ph.D. degrees, all in electrical engineering, from the Massachusetts Institute of Technology, Cambridge, in 1974, 1976, 1977, and 1979, respectively.

He is currently Professor of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor. In winter and spring 1992, he was a Visiting Professor at the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland. Prior to joining the University of Michigan, he worked for Systems Control, Inc., Palo Alto, CA, and Alphatech, Inc., Burlington, MA. His research interests are in stochastic control, decentralized systems, queueing and communication networks, stochastic scheduling, and resource allocation problems, mathematical economics, and discrete-event systems.