On Globally Optimal Encoding, Decoding, and Memory Update for Noisy Real-Time Communication Systems

Aditya Mahajan and Demosthenis Teneketzis

May 9, 2008

Abstract

The design of optimal joint source-channel communication strategies for a real-time communication system, i.e., a sequential communication system in which information must be transmitted and decoded withing a fixed-finite delay, is considered. First, a system which runs for a finite horizon and consists of a first-order Markov source, a real-time encoder, a memoryless noisy channel, a real-time decoder with finite memory, and distortion metric that accepts zero delay, is considered. The design of optimal real-time communication strategies is formulated as a decentralized stochastic optimization problem. There is no existing solution methodology to solve general decentralized stochastic optimization problems over finite and infinite horizon. This paper develops a systematic methodology, based on the notions of information structure and information state, to sequentially obtain globally optimal real-time encoding, decoding, and memory update strategies. Such a sequential decomposition results in a set of nested optimality equations whose solution determines an optimal communication strategy. This methodology is extended to two classes of infinite-horizon systems, where optimal communication strategies are determined by the solution of an appropriate functional equation. The methodology is also extended to systems where distortion metric accepts a fixed-finite delay, to systems with higher-order Markov sources, and to systems with channels with memory. Thus, this paper develops a comprehensive method to study different variations of real-time communication.

Index Terms: Real-time communication, zero-delay communication, joint source-channel coding, dynamic teams, information state, non-classical information structures

1. Introduction

1.1. Motivation

In many controlled informationally-decentralized systems such as 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, the delay incurred in transmission of information has to be bounded. In order to understand how to design the above described 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.

The authors are with the department of EECS at the University of Michigan, Ann Arbor, MI 48109-2122, USA. (email: {adityam,teneket}@eecs.umich.edu)

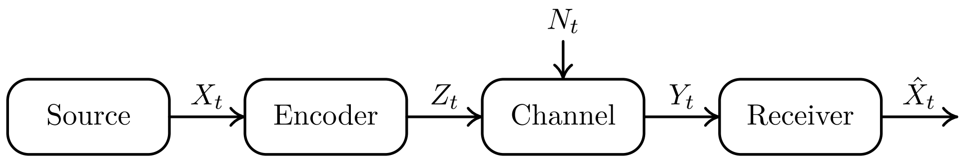

Figure 1: A point-to-point real-time communication system

In this paper we consider the simplest instance of a real-time communication system: a point-to-point real-time communication system shown in Figure 1. Consider a dynamic first-order Markov source whose outputs need to be sequentially transmitted over a noisy channel to a receiver. The encoding and decoding must be done with zero-delay. A distortion metric between the source outputs and its reconstruction measures the per symbol quality of reconstruction. The choice of encoding and decoding rules for each time instant is called a communication strategy. We want to choose a communication strategy that minimizes 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.

1.2. Conceptual difficulties

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, transmission rate, and channel capacity are asymptotic concepts; the performance bounds that it provides are tight only for asymptotically large values of delay. Real-time communication is not asymptotic. Hence, the concepts and results from information theory are not appropriate for real-time communication. In particular, separate source and channel coding is not optimal and joint source-channel coding strategies must be considered.

Real-time communication can be considered as a decentralized multi-agent sequential stochastic optimization problem. The system has two agents—the encoder and the decoder. Due to the noise in the communication channel, the encoder does not know the information available at the decoder and vice-versa; thus, the two agents have different decentralized information. Due to this decentralization of information, solving the real-time communication problem as an optimization problem is outside the domain of Markov decision theory [2] which is only applicable to stochastic optimization problems with centralized information.

As a stochastic optimization problem, real-time communication can be classified as a *team problem*; teams are a subclass of decentralized multi-agent stochastic optimization problems where all agents have a common objective (which in this case is minimizing an expected total distortion). An optimal solution of team problems can, in principle, be obtained by a brute force search. However, a brute force search is computationally intractable for medium to large sized problems. So we need a systematic method for finding an optimal solution. But a systematic method is not easy to find because team problems have a complex interdependence among the agents' decision rules; this interdependence results in optimization problems that are non-convex in strategy space. There is no known solution methodology to solve general team problems.

A solution concept for team problems is *sequential decomposition* which breaks the one-shot (brute-force) optimization problem into a sequence of nested optimization problems, usually one sub-problem for each time an agent acts. This decomposition exponentially simplifies the search for an optimal strategy. A key step in obtaining a sequential decomposition is identifying an information state appropriate for performance evaluation for each agent. To the best of our knowledge, there is only one known methodology for obtaining appropriate information states for team problems—Witsenhausen's standard form [3]. However, the standard form is applicable to only finite horizon problems; we are interested in both finite and infinite horizon problems, so we cannot use the standard form.

In this paper we identify information states sufficient for performance evaluation of the finite- and infinite-horizon real-time communication problem. We first present properties that such an information state must satisfy and then identify an information state that satisfies these properties. We then show how to obtain a sequential decomposition using these information states.

1.3. Literature Overview

The work on real-time communication can be classified on the basis of the source model (memoryless source, Markov source, stationary source, or individual sequence) and the channel model (noiseless channel, noisy channel with no feedback, noisy channel with noiseless feedback, noisy channel with noisy feedback). Zero-delay and finite-delay source coding problems were considered in [4]–[9]. A weaker constraint of *causal source coding* were investigated in [10]–[13]. Properties of optimal systems for real-time communication over noisy channels were obtained in [14]–[17]. Bounds on performance of communication systems with a real-time or finite-delay constraint on the information transmission were obtained via different methods (e.g., mathematical programming, forward flow of information, conditional mutual information, determination of non-anticipatory rate distortion function, randomizing over a family of encoders-decoders) in [18]–[25]. Real-time encoding and decoding of individual sequences over noisy channels with infinite memory at the encoder and the decoder was considered in [26].

Properties of real-time decoders for noisy observations of a Markov source were considered in [27, 28]. Properties of real-time encoders for transmitting Markov sources through noiseless channels were investigated in [29, 30]. 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 [31, 32]. Applications of results developed in [32] appeared in [33, 34]. 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 [35]. The structural properties of optimal real-time encoders and decoders, and a methodology for determining optimal communication strategies for point-to-point communication systems with a noisy channel and noisy feedback were investigated in [36]. In this paper we use the structural results of [35] to obtain a sequential decomposition of obtaining *globally optimal* of real-time encoding, decoding, and memory update strategies.

1.4. Contributions

The key contribution of this paper is the presentation of a systematic methodology for the design of globally optimal communication strategies for point-to-point real-time communication systems. We treat the design of an optimal real-time communication strategy as a decentralized multi-agent stochastic optimization problem and obtain a sequential decomposition for the problem. For the finite horizon case, our methodology converts the search of an optimal communication strategy into a sequence of nested optimality equations; for the infinite horizon cases, our methodology transforms the search for an optimal communication strategy into finding the unique fixed point of an appropriate functional equation. We also extend our methodology to distortion metrics accepting a fixed-finite delay, to higher-order Markov sources, and to channels with memory. This methodology drastically simplifies the search for an optimal communication strategy; in spite of this simplification, numerically solving the resultant optimality equations remains a formidable task. We are not aware, at this point, of the existence of good approximation techniques for solving these optimality equations. 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.

To the best of our knowledge, this paper is the first example of a sequential decomposition that works for both finite and infinite horizon versions of a decentralized stochastic optimization problem with non-classical information structure.

1.5. Organization

The rest of this paper is organized as follows. In Section 2 we formulate the finite horizon problem, present its salient features, and show how the structural results of [35] simplify the problem. In Section 3 we explain the notion of an information state, and the properties that an information state should satisfy. We then identify variables that satisfy these properties, and show how they can be used to sequentially obtain globally optimal design. In Section 4 we extend the results of the finite horizon problem to infinite horizon problems with expected discounted distortion and average distortion per unit time criteria. In Section 5 we consider fixed-finite delay communication problem; we convert it into a zero-delay communication problem and obtain a sequential decomposition for both finite and infinite horizon variations. In Section 6 we consider the problem with higher-order Markov source; we convert it into a first-order Markov source and obtain a sequential decomposition for both finite and infinite horizon variations. In Section 7 we consider channels with memory and show how to obtain a sequential decomposition in this case. In Section 8 we discuss the salient features of computing optimal designs using the methodology presenting in this paper. In Section 9 we compare the philosophy of our approach with the philosophy of information theory and coding theory. In Section 10 we present some concluding remarks and future directions.

1.6. 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 $(\mathscr{C}, \mathscr{G}, \mathscr{L})$ represent family of functions and Gothic letters $(\mathfrak{F}, \mathfrak{E}, \mathfrak{R})$ represent σ–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 $\mathbb{P}\{\mathcal{X}\}$ denotes the space of all PMF (probability mass functions) on $\mathcal{X}$. In order to denote that the expectation of a random variable or the probability of an event depends on a function φ, we use $\mathbb{E}\{\cdot \mid \varphi\}$ and $\Pr(\cdot \mid \varphi)$, respectively. This slightly unusual notation is chosen since we want to keep track of all functional dependencies and the conventional notation of $\mathbb{E}^{\varphi}\{\cdot\}$ and $\Pr^{\varphi}(\cdot)$ is too cumbersome to use.

2. The Finite Horizon Problem

2.1. Problem Formulation

We first consider the finite horizon version of the problem. Consider a discrete time communication system shown in Figure 1. A first-order Markov source produces a random sequence $\{X_t,\ t =$

$1,\ldots,T\}$. 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 can transmit a symbol Z_t taking values in a finite alphabet $\mathcal{Z}$. This encoded symbol is causally generated in real-time using the source outputs until that time according to an encoding rule c_t, i.e.,

$$Z_t = c_t(X_1,\ldots,X_t), \quad t = 1,\ldots,T, \tag{1}$$

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 can be described as

$$Y_t = h_t(Z_t, N_t), \tag{2}$$

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

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. We model this 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.,

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

The objective of the decoder is to generate 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.,

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

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

The collection $C := (c_1,\ldots,c_T)$ of encoding rules for the entire horizon is called an *encoding strategy*. Similarly, the collection $G := (g_1,\ldots,g_T)$ of decoding rules is called a *decoding strategy* and the collection $L := (l_1,\ldots,l_T)$ of memory update rules is called a *memory update strategy*. Further, 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

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

We are interested in the following optimization problem:

Problem 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.,*

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

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

In Problem 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 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 the one with the best performance. The number of possibly time-varying communication strategies are exponential in the size of the time horizon and the cardinality of the alphabets which makes a brute force search for an optimal solution intractable. So, 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 size of the time horizon and exponential complexity in the cardinality of the alphabets. In the next section we present an example for a real-time communication system.

2.2. 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 $\mathcal{M} = \{0, 1, \ldots, 7\}$. Suppose the source statistics are

$$P_{X_1} = \begin{bmatrix} 0.4 & 0.6 \end{bmatrix}, \qquad P_{X_2|X_1} = P_{X_3|X_2} = \begin{bmatrix} 1.0 & 0.0 \\ 0.1 & 0.9 \end{bmatrix} \tag{7}$$

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

$$h_t(Z_t, N_t) = Z_t \cdot N_t, \qquad P_{N_1} = P_{N_2} = P_{N_3} = \begin{bmatrix} 0.1 & 0.9 \end{bmatrix}. \tag{8}$$

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

$$Z_1 = c_1(X_1), \qquad Z_2 = c_2(X_2), \qquad Z_3 = c_3(X_3), \tag{9a}$$
$$\hat{X}_1 = g_1(Y_1), \qquad \hat{X}_2 = g_2(Y_2), \qquad \hat{X}_3 = g_3(Y_3). \tag{9b}$$

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

$$Z_1 = c_1(X_1), \qquad Z_2 = c_2(X_1, X_2), \qquad Z_3 = c_3(X_1, X_2, X_3), \tag{10a}$$
$$\hat{X}_1 = g_1(Y_1), \qquad \hat{X}_2 = g_2(Y_1, Y_2), \qquad \hat{X}_3 = g_3(Y_1, Y_2, Y_3). \tag{10b}$$

Observe that there are $(2^2 \times 3)^2 = 144$ memoryless communication strategies of the form (9), 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 (10).

x_1	x_2	x_3	z_1	z_2	z_3
0	0	0	1	1	0
0	0	1	1	1	1
0	1	0	1	0	0
0	1	1	1	0	1
1	0	0	0	1	0
1	0	1	0	1	1
1	1	0	0	0	0
1	1	1	0	0	1

(a) Encoder

y_1	y_2	y_3	$\hat{x}_1$	$\hat{x}_2$	$\hat{x}_3$
0	0	0	1	1	0
0	0	1	1	1	1
0	1	0	1	0	0
0	1	1	1	0	1
1	0	0	0	1	0
1	0	1	0	1	1
1	1	0	0	0	0
1	1	1	0	0	1

(b) Decoder

Table 1: Best memoryless encoding and decoding strategies for the example of Section 2.2. The performance of this scheme is 673/5000.

x_1	x_2	x_3	z_1	z_2	z_3
0	0	0	1	1	1
0	0	1	1	1	1
0	1	0	1	1	1
0	1	1	1	1	1
1	0	0	0	1	1
1	0	1	0	1	1
1	1	0	0	0	1
1	1	1	0	0	0

(a) Encoder

y_1	y_2	y_3	$\hat{x}_1$	$\hat{x}_2$	$\hat{x}_3$
0	0	0	1	1	1
0	0	1	1	1	0
0	1	0	1	0	0
0	1	1	1	0	0
1	0	0	0	0	0
1	0	1	0	0	0
1	1	0	0	0	0
1	1	1	0	0	0

(b) Decoder

Table 2: Best real-time encoding and decoding strategies for the example of Section 2.2. The performance of this scheme is 141/2500.

The best performance (i.e., total expected distortion) of a memoryless communication strategy is 0.1346 which is achieved by the strategy shown in Table 1; on the other hand, the best performance of a real-time communication strategy is 0.0564 which is achieved by the strategy shown in Table 2. Thus, for this example, real-time communication strategies provide 58% performance improvement over memoryless communication strategies.

It was shown in [16, 17] 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 is not 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 time horizon. In the above example, which is one of the simplest real-time communication system, there are 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 time horizon 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 time horizon (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.

2.3. Primitive random variables

In this paper we will be working with conditional probabilities, probability measures of probability measures, and σ-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.

Let χ_t a $|\mathcal{X}|$-dimensional random vector defined as follows: for $x \in \mathcal{X}$,

$$\chi_t(x) := \mathbb{1}\left[X_t = x\right] \quad \text{and} \quad \chi_t := [\chi_t(1), \dots, \chi_t(|\mathcal{X}|)]$$

There is a one to one relation between X_t and χ_t and we can use χ_t to have a martingale representation (stochastic difference equation) for the Markov chain $\{X_t,\, t = 1, \dots, T\}$ (see [37]) given by

$$\chi_{t+1} = P^T_{X_{t+1}|X_t}\chi_t + \theta_t,$$

where $\{\theta_t,\, t = 1, \dots, T\}$ is a sequence of independent zero-mean random vectors. Since we have assumed that the noise in the forward channel is independent of the source output, the random variables $(\chi_1, \theta_1, \dots, \theta_T, N_1, \dots, N_T)$ 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 can be 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 $(\Omega, \mathfrak{F}, P)$.

2.4. Problem classification

Problem 1 is a sequential stochastic optimization problem as defined in [38]. To understand the sequential nature of the problem, we need to refine the notion of time. We call each step of the system a *stage*. For each stage we consider three time instances[1]: ${}^1t$, ${}^2t$ and ${}^3t$. We can 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 ${}^1t$, ${}^2t$, and ${}^3t$, 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 1 is a sequential problem. The sequential ordering of the system variables is shown in Figure 2 (some of these variables will be defined later).

In Problem 1, all agents have the same objective given by (5). Multi-agent problems in which all agents have the same objective are called teams [39], and are further classified as static 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, while in dynamic teams, in general, an agent’s information depends on the functional form of the decision rules of other agents. In Problem 1 the receiver’s information depends on the functional form of the encoding rule. Thus Problem 1 is a dynamic team. Dynamic teams are, in general, functional optimization problems having a complex interdependence among the decision rules [40]. This interdependence leads to non-convex (in policy space) optimization problems that are hard to solve.

[1] The actual values of these time instances is irrelevant; we need three values in increasing order.

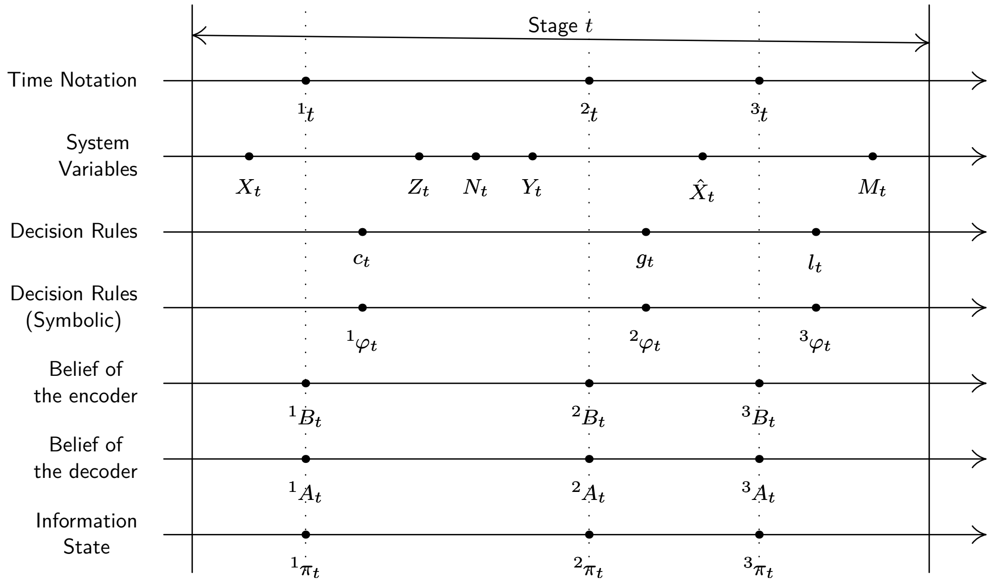

Figure 2: The sequential ordering of the variables of the real time communication system. ^1t, ^2t, and ^3t are refinements of stage t.

2.5. Information fields

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

$$^1\varphi^{t-1} := (c^{t-1}, g^{t-1}, l^{t-1}), \qquad ^1\varphi_t := c_t, \tag{11a}$$

$$^2\varphi^{t-1} := (c^{t}, g^{t-1}, l^{t-1}), \qquad ^2\varphi_t := g_t, \tag{11b}$$

$$^3\varphi^{t-1} := (c^{t}, g^{t}, l^{t-1}), \qquad ^3\varphi_t := l_t, \tag{11c}$$

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

Definition 1 (Encoder's Information). *Let ${}^iE_t$ denote the observation and ${}^i\mathfrak{E}_t$ denote the information field at the encoder's site at time ${}^it$, $i = 1, 2, 3$. Then*

$$^1E_t := (X^t, Z^{t-1}), \qquad ^2E_t := (X^t, Z^t), \qquad ^3E_t := {}^2E_t \tag{12a}$$

and

$$^i\mathfrak{E}_t := \sigma({}^iE_t; {}^i\varphi^{t-1}), \quad i = 1, 2, 3. \tag{12b}$$

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

Definition 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*

$$^1R_t := M_{t-1}, \qquad ^2R_t := (Y_t, M_{t-1}), \qquad ^3R_t := {}^2R_t \tag{13a}$$

and

$$ {}^i\mathfrak{R}_t := \sigma({}^iR_t; {}^i\varphi^{t-1}), \quad i = 1, 2, 3. \tag{13b} $$

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. Non-classical information structure.

The information fields at the encoder's and the receiver's sites are non-compatible, i.e., ${}^i\mathfrak{E}_t \not\subseteq {}^i\mathfrak{R}_t$ and ${}^i\mathfrak{E}_t \not\supseteq {}^i\mathfrak{R}_t$. 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 1 has a non-classical 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., ${}^1\mathfrak{E}_t \subseteq {}^2\mathfrak{E}_t \subseteq {}^3\mathfrak{E}_t \subseteq {}^1\mathfrak{E}_{t+1}$ and so on. On the other hand, the receiver has finite memory. As a result, the information fields at the receiver are not-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 we have ${}^3\mathfrak{R}_t \not\subseteq {}^1\mathfrak{R}_{t+1}$, i.e., at time ${}^3t$ the receiver sheds information.

2.6. 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 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 ${}^ir \in {}^i\mathcal{R}$,*

$$ {}^iB_t({}^ir) := \Pr\left({}^iR_t = {}^ir \,\middle|\, {}^i\mathfrak{E}_t\right) \tag{14} $$

Let ${}^i\mathcal{B} := \mathbb{P}\left\{{}^i\mathcal{R}\right\}$ denote the space of realizations of ${}^iB_t$

Definition 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 ${}^ie_t \in {}^i\mathcal{E}_t$,*

$$ {}^iA_t({}^ie_t) := \Pr\left({}^iE_t = {}^ie_t \,\middle|\, {}^i\mathfrak{R}_t\right) \tag{15} $$

Let ${}^i\mathcal{A}_t := \mathbb{P}\left\{{}^i\mathcal{E}_t\right\}$ 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 $x \in \mathcal{X}$,

$$ \hat{A}_t(x) := \Pr\left(X_t = x \,\middle|\, {}^2\mathfrak{R}_t\right) \tag{16} $$

The sequential ordering of the beliefs is shown in Figure 2. For any particular realization ${}^1e_t$ of ${}^1E_t$, and any arbitrary (but fixed) choice of ${}^1\varphi^{t-1}$, 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 $\mathbb{P}\{\mathcal{M}\}$, 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 1 (Evolution of the encoder's beliefs). *For each stage t, there exist deterministic functions ${}^1F_t$ and ${}^3F$ such that*

$$
\begin{aligned}
{}^2B_t &= {}^1F_t({}^1B_t, Z_t) && \text{(17a)}\\
{}^3B_t &= {}^2B_t && \text{(17b)}\\
{}^1B_{t+1} &= {}^3F({}^3B_t, l_t) && \text{(17c)}
\end{aligned}
$$

This is proved in Appendix A.

2.7. Structural Properties

In this section we summarize the structural results of [35]. We will later uses these structural results to develop a methodology for determining globally optimal communication strategies.

The structural properties of optimal real-time encoders were derived in [35, Section II-B]. We restate the main result below.

Theorem 1 (Structure of optimal real-time encoders). *Consider Problem 1 for any arbitrary (but fixed) decoding and memory update strategies, $G = (g_1, \dots, g_T)$ and $L = (l_1, \dots, l_T)$, respectively. Then there is no loss in optimality in restricting attention to encoding rules of the form*

$$Z_t = c_t(X_t, {}^1B_t), \quad t = 2, \dots, T. \tag{18}$$

The structural properties of optimal real-time decoders were derived in [35, Section II-E]. We restate the main result below.

Theorem 2 (Structure of optimal real-time decoders). *Consider Problem 1 for any arbitrary (but fixed) encoding and memory update strategies, $C = (c_1, \dots, c_T)$ and $L = (l_1, \dots, l_T)$, respectively. Then there is no loss in optimality in restricting attention to decoding rules of the form*

$$\hat{X}_t = \hat{g}_t(\hat{A}_t) := \arg\min_{\hat{x} \in \hat{\mathcal{X}}} \sum_{x \in \mathcal{X}} \rho_t(x, \hat{x}) \hat{A}_t(x). \tag{19}$$

2.8. Implication of the structural results

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

Corollary 1. *The optimal performance $\mathcal{J}_T^*$ given by* (6) *can be determined by*

$$\mathcal{J}_T^* := \inf_{\substack{C \in \hat{\mathscr{C}}^T \\ G \in \mathscr{G}^T \\ L \in \mathscr{L}^T}} \mathcal{J}_T(C, G, L), \tag{20}$$

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

Hence, in Problem 1 rather than choosing a communication strategy (C^*, G^*, L^*) belonging to $(\mathscr{C}^T \times \mathscr{G}^T \times \mathscr{L}^T)$ to minimize (6) we can choose a communication strategy (C^*, G^*, L^*) belonging to $(\hat{\mathscr{C}}^T \times \mathscr{G}^T \times \mathscr{L}^T)$ to minimize (20). Notice that the domain of an encoding rule belonging to $\mathscr{C}_t$ increases with t, while the domain of an encoding rule belonging to $\hat{\mathscr{C}}$ does not depend on t. Hence, using the structural results of Theorem 1 we can reformulate Problem 1 such that the encoding rules at each time have to be chosen from a time-invariant space. The reformulated problem is as follows:

Problem 2. *Under the assumptions of Problem 1, choose a communication strategy* (C^*, G^*, L^*) *belonging to* $(\hat{\mathscr{C}}^T \times \mathscr{G}^T \times \mathscr{L}^T)$ *that is optimal with respect to the performance criterion of* (20).

In the sequel we will concentrate on Problem 2.

3. Global Optimization

In this section we provide a methodology for sequential decomposition of Problem 2. The class of problems consisting of Problem 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 what is known as "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.

3.1. Information structures and information state

In a multi-agent 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. Multi-agent systems can be classified according to their information structures. In [41] three classes of information structures are defined: classical, quasi-classical, and non-classical 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 a change of variables can convert the information structure into a classical information structure. A system that has neither classical nor quasi-classical information structure is said to have a (strictly) *non-classical information structure*.

Markov decision theory explains how to obtain a sequential decomposition of problems with a classical information structure. Witsenhausen's standard form [38] explains how to obtain a sequential decomposition of a subclass of problems with a non-classical information structure. There is no known methodology to obtain a sequential decomposition of a general problem with a non-classical information structure. Problem 1 and its infinite horizon extensions have a non-classical information structure so they *cannot be solved by Markov decision theory*; the infinite horizon extensions of Problem 1 belong to a subclass that *cannot be solved using the standard form*. So, we need to develop a new methodology for sequential decomposition for these problems.

A critical step in obtaining a sequential decomposition for problems with non-classical information structures is identifying an information state sufficient for performance evaluation. 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 non-classical information structure. The idea behind an information state is best described in [42]: "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:

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 ${}^{i}t$: nature has produced (x^t, n^t), the designer has chosen ${}^{i}\varphi^{t-1}$ 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 ${}^{i}\varphi_t^T$ to minimize the expected future distortion $\mathbb{E}\left\{\sum_{s=t}^{T} \rho_s(X_s, \hat{X}_s) \,\middle|\, {}^{i}\varphi^{t-1}, {}^{i}\varphi_t^T\right\}$.

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 ${}^{i}\varphi^{t-1,(1)}$ and ${}^{i}\varphi^{t-1,(2)}$ are equivalent, denoted by ${}^{i}\varphi^{t-1,(1)} \sim {}^{i}\varphi^{t-1,(2)}$, if for any choice of future decision rules ${}^{i}\varphi_t^T$

$$\mathbb{E}\left\{\sum_{s=t}^{T} \rho_s(X_s, \hat{X}_s) \,\middle|\, {}^{i}\varphi^{t-1,(1)}, {}^{i}\varphi_t^T\right\} = \mathbb{E}\left\{\sum_{s=t}^{T} \rho_s(X_s, \hat{X}_s) \,\middle|\, {}^{i}\varphi^{t-1,(2)}, {}^{i}\varphi_t^T\right\}$$

Assume that the designer has already chosen ${}^{i}\varphi^{t-1}$ and wants to choose ${}^{i}\varphi_t^T$ to minimize the expected future cost. If ${}^{i}\varphi^{t-1,(1)} \sim {}^{i}\varphi^{t-1,(2)}$ then the optimal future communication rules will be the 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 ${}^{i}\pi_t : {}^{i}\Phi^{t-1} \to {}^{i}\Pi_t$ is a function such that for any ${}^{i}\varphi^{t-1,(1)}, {}^{i}\varphi^{t-1,(2)} \in {}^{i}\Phi^{t-1}$ if ${}^{i}\pi_t({}^{i}\varphi^{t-1,(1)}) = {}^{i}\pi_t({}^{i}\varphi^{t-1,(2)})$, then ${}^{i}\varphi^{t-1,(1)} \sim {}^{i}\varphi^{t-1,(2)}$. 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 [43], 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 non-classical 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, we want ${}^{2}\pi_t({}^{2}\varphi^{t-1})$ to be a function of ${}^{1}\pi_t({}^{1}\varphi^{t-1})$ and ${}^{1}\varphi_t$ (recall that ${}^{1}\varphi^{t-1} = ({}^{1}\varphi^{t-1}, {}^{1}\varphi_t)$), ${}^{3}\pi_t({}^{3}\varphi^{t-1})$ to be a function of ${}^{2}\pi_t({}^{2}\varphi^{t-1})$ and ${}^{2}\varphi_t$, and ${}^{1}\pi_{t+1}({}^{1}\varphi^{t})$ to be a function of ${}^{3}\pi_t({}^{3}\varphi^{t-1})$ and ${}^{3}\varphi_t$.

Any sequence $\{{}^i\pi_t,\ i=1,2,3,\ t=1,\dots,T\}$ that have 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.

P3. Time invariant domain

We want to identify functions ${}^i\pi_t : {}^i\Phi^{t-1} \to {}^i\Pi$ 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.

P4. Minimality.

If more than one appropriate information state exist we want to work with the "smallest" information state in order to compute an optimal solution in the most efficient manner. However, we have not been able to establish a good way of comparing information states (especially, information states of infinite horizon problems) so that we can always compare any two appropriate information states. So, in the rest of the paper, we will not consider minimality.

For a given communication rule (C,G,L), we call ${}^i\pi_t({}^i\varphi^{t-1})$ as the information state at time ${}^it$ and denote it by ${}^i\pi_t$. In summary, these information states should satisfy:

S1. The information state is a summary of past information.

Thus, ${}^1\pi_t$ should be a function of ${}^1\varphi^{t-1}$, ${}^2\pi_t$ should be a function of ${}^2\varphi^{t-1}$ and ${}^3\pi_t$ should be a function of ${}^3\varphi^{t-1}$.

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

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

This means that

$$\begin{aligned}\mathbb{E}\left\{\sum_{s=t}^{T}\rho_s(X_s,\hat{X}_s)\,\middle|\, C,G,L\right\} &= \mathbb{E}\left\{\sum_{s=t}^{T}\rho_s(X_s,\hat{X}_s)\,\middle|\, {}^3\pi_{t-1}, c_t^T, g_t^T, l_{t-1}^T\right\}\\ &= \mathbb{E}\left\{\sum_{s=t}^{T}\rho_s(X_s,\hat{X}_s)\,\middle|\, {}^1\pi_t, c_t^T, g_t^T, l_t^T\right\}\\ &= \mathbb{E}\left\{\sum_{s=t}^{T}\rho_s(X_s,\hat{X}_s)\,\middle|\, {}^2\pi_t, c_{t+1}^T, g_t^T, l_t^T\right\} \qquad (21)\end{aligned}$$

By (S2) this is equivalent to

$$
\begin{aligned}
\mathbb{E}\left\{\rho_t(X_t,\hat{X}_t)\,\middle|\,C,G,L\right\} &= \mathbb{E}\left\{\rho_t(X_t,\hat{X}_t)\,\middle|\,{}^3\pi_{t-1}, l_{t-1}, c_t, g_t\right\} \\
&= \mathbb{E}\left\{\rho_t(X_t,\hat{X}_t)\,\middle|\,{}^1\pi_t, c_t, g_t\right\} \\
&= \mathbb{E}\left\{\rho_t(X_t,\hat{X}_t)\,\middle|\,{}^2\pi_t, g_t\right\} \qquad (22)
\end{aligned}
$$

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 non-classical 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 5. *Define ${}^1\pi_t$, ${}^2\pi_t$, and ${}^3\pi_t$ as follows:*

$$
{}^1\pi_t \coloneqq \Pr\left(X_t, {}^1B_t \,\middle|\, {}^1\varphi^{t-1}\right), \qquad (23a)
$$

$$
{}^2\pi_t \coloneqq \Pr\left(X_t, {}^2B_t \,\middle|\, {}^2\varphi^{t-1}\right), \qquad (23b)
$$

$$
{}^3\pi_t \coloneqq \Pr\left(X_t, {}^3B_t \,\middle|\, {}^3\varphi^{t-1}\right). \qquad (23c)
$$

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 ${}^i\varphi^{t-1}$ of past decision rules for agent i, $i = 1,2,3$, the beliefs ${}^iB_t$ are $\mathfrak{F}$–measurable. Thus, for any choice of ${}^i\varphi^{t-1}$, $(X_t, {}^iB_t)$ 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 2. *For encoding rules of the form* (18)*, ${}^1\pi_t$, ${}^2\pi_t$, 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*

$$
{}^2\pi_t = {}^1Q_t(c_t)\,{}^1\pi_t \qquad (24a)
$$

$$
{}^3\pi_t = {}^2\pi_t \qquad (24b)
$$

$$
{}^1\pi_{t+1} = {}^3Q_t(l_t)\,{}^3\pi_t \qquad (24c)
$$

2. *The expected instantaneous cost can be expressed as*

$$
\mathbb{E}\left\{\rho_t(X_t,\hat{X}_t)\,\middle|\,{}^2\varphi^{t-1}, g_t\right\} = \hat{\rho}_t({}^2\pi_t, g_t) \qquad (25)
$$

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

3.2. 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 $\hat{\mathscr{C}}$, $\mathscr{G}$, and $\mathscr{L}$. The system dynamics are given by (24) 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 $\hat{\mathscr{C}}$, ${}^2\Delta_t$ is a function from ${}^2\Pi$ to $\mathscr{G}$, and ${}^3\Delta_t$ is a function from ${}^3\Pi$ to $\mathscr{L}$. Thus the system equations (24) can be written as

$$c_t = {}^1\Delta_t({}^1\pi_t), \qquad {}^2\pi_t = {}^1Q(c_t)\,{}^1\pi_t, \tag{26a}$$

$$g_t = {}^2\Delta_t({}^2\pi_t), \qquad {}^3\pi_t = {}^2\pi_t, \tag{26b}$$

$$l_t = {}^3\Delta_t({}^3\pi_t), \qquad {}^1\pi_{t+1} = {}^3Q(l_t)\,{}^3\pi_t. \tag{26c}$$

The initial state ${}^1\pi_1 = P_{X_1}$ is given. An instantaneous cost $\hat{\rho}({}^2\pi_t, g_t)$ is incurred at each stage. 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 Δ^T. The performance of a meta–strategy is given by the total cost incurred by that meta-strategy, i.e.,

$$\mathscr{J}_T(\Delta^T \,|\, {}^1\pi_1) = \sum_{t=1}^{T} \hat{\rho}({}^2\pi_t, g_t). \tag{27}$$

Now consider the following optimization problem:

Problem 3. *Consider the dynamic system* (26) *with known transformations* ${}^1Q$ *and* ${}^3Q$*. The initial state* ${}^1\pi_1$ *is given. Determine a meta–strategy* Δ^T *to minimize the total cost given by* (27)*.*

Given any meta-strategy Δ^T, the time evolution of ${}^i\pi_t$ is deterministic; ${}^i\pi_t$ and the corresponding ${}^i\varphi_t$ can be determined from (26). Thus, for a given initial states ${}^1\pi_t$, there is a communication strategy corresponding to any choice of meta-strategy. Further, we can rewrite the performance criterion of (5) as

$$\begin{aligned}
\mathscr{J}_T(C, G, L) &= \mathbb{E}\left\{ \sum_{t=1}^{T} \rho(X_t, \hat{X}_t) \,\middle|\, C, G, L \right\} \\
&\overset{(a)}{=} \sum_{t=1}^{T} \mathbb{E}\left\{ \rho(X_t, \hat{X}_t) \,\middle|\, {}^2\varphi^{t-1}, g_t \right\} \\
&\overset{(b)}{=} \sum_{t=1}^{T} \hat{\rho}({}^2\pi_t, g_t) \\
&=: \mathscr{J}_T(\Delta^T \,|\, {}^1\pi_1)
\end{aligned} \tag{28}$$

where (a) follows from the sequential ordering of system variables and (b) follows from Lemma 2. Thus, if $\Delta^{*,T}$ is an optimal meta-strategy for Problem 3, and (C^*, G^*, L^*) is the communication strategy corresponding to $\Delta^{*,T}$, then (C^*, G^*, L^*) is an optimal communication strategy for Problem 2 and thereby also for Problem 1. Hence, Problem 3 is equivalent to Problems 1 and 2. Now we provide an algorithm to determine an optimal meta-strategy for Problem 3.

3.3. The global optimization algorithm

Problem 3 can be formulated as a classical centralized optimization problem by considering the information state ${}^i\pi_t$ is the "controlled state" at time ${}^it$, the communication rule ${}^i\varphi_t$ (c_t, g_t, or l_t

depending on i) as the "control action" (or decision) at time ${}^it$, and the meta-function ${}^i\Delta_t$ as the "control law" at time ${}^it$. Hence, an optimal meta-strategy for Problem 3 is given by the optimal "control strategy" of the centralized optimization problem and can be determined as follows:

Theorem 3 (Global optimization algorithm). *An optimal meta-strategy $\Delta^{*,T}$ for Problem 3 (and consequently an optimal communication strategy for Problem 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*

$$ {}^1V_{T+1}({}^1\pi) = 0, \tag{29a} $$

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

$$ {}^1V_t({}^1\pi) = \inf_{c_t \in \hat{\mathscr{C}}} {}^2V_t\big({}^1Q_t(c_t)\,{}^1\pi\big), \tag{29b} $$

$$ {}^2V_t({}^2\pi) = \min_{g_t \in \mathscr{G}} \Big[\hat{\rho}_t({}^2\pi, g_t) + {}^3V_t({}^2\pi)\Big], \tag{29c} $$

$$ {}^3V_t({}^3\pi) = \min_{l_t \in \mathscr{L}} {}^1V_{t+1}\big({}^3Q_t(l_t)\,{}^3\pi\big). \tag{29d} $$

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 ${}^it$ onwards. These value functions can be determined iteratively by moving backwards in time. The optimal performance of Problem 3 (and Problem 1) is given by*

$$ \mathscr{J}_T^* = {}^1V_1({}^1\pi_1). \tag{30} $$

For any ${}^it$ 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. Relations* (26) *can be used to determine optimal communication strategy for Problem 1.*

Proof. This is a standard result for a deterministic optimization problem, see [2, Chapter 2]. □

Observe that the three step T-stage sequential decomposition of (29) can be combined into a one-step T-stage sequential decomposition

$$ {}^1V_t({}^1\pi) = \inf_{\substack{C \in \hat{\mathscr{C}}^T \\ G \in \mathscr{G}^T \\ L \in \mathscr{L}^T}} \Big[\hat{\rho}_t\big({}^1Q_t(c_t)\,{}^1\pi, g_t\big) + {}^1V_{t+1}\Big(\big({}^3Q_t(l_t) \circ {}^1Q_t(c_t)\big)\,{}^1\pi\Big)\Big] \tag{31} $$

which is a deterministic dynamic program in function spade. We present a finer decomposition in Theorem 3 which corresponds to the refinement of time presented in Section 2.4; the decomposition given by (29) has a smaller search space than the decomposition given in (31).

3.4. 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 1 is time-homogeneous, some of the results derived in the previous section can be simplified. The function ${}^1F_t$ in Lemma 1 does not depend on t; the transformations ${}^1Q_t$, ${}^3Q_t$ and the function $\hat{\rho}_t$ of Lemma 2 also do not depend on t; thus, we can drop the subscripts t and simply denote them by ${}^1F$, ${}^1Q$, ${}^3Q$ and $\hat{\rho}$, respectively. Hence Problem 3 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. Hence we can simplify Theorem 3 as follows.

Corollary 2. *If the system of Problem 1 is time-homogeneous, the nested optimality equations* (29) *can be written as*

$$ {}^1V_{T+1}({}^1\pi) = 0, \tag{32a} $$

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

$$ {}^1V_t({}^1\pi) = \inf_{c_t \in \hat{\mathscr{C}}} {}^2V_t\big({}^1Q(c_t)\,{}^1\pi\big), \tag{32b} $$

$$ {}^2V_t({}^2\pi) = \min_{g_t \in \mathscr{G}} \Big[\hat{\rho}({}^2\pi, g_t) + {}^3V_t({}^2\pi)\Big], \tag{32c} $$

$$ {}^3V_t({}^3\pi) = \min_{l_t \in \mathscr{L}} {}^1V_{t+1}\big({}^3Q(l_t)\,{}^3\pi\big). \tag{32d} $$

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

4. The infinite horizon time-homogeneous problem

In this section we extend the time-homogeneous model of Section 3.4 to an infinite horizon ($T \to \infty$) 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:

1. *The expected discounted cost criteria*

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

   $$ \mathscr{J}^{\beta}(C, G, L) := \mathbb{E}\left\{ \sum_{t=1}^{\infty} \rho(X_t, \hat{X}_t) \,\middle|\, C, G, L \right\} \tag{33} $$

   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

   $$ \bar{\mathscr{J}}(C, G, L) := \limsup_{T \to \infty} \mathbb{E}\left\{ \sum_{t=1}^{T} \rho(X_t, \hat{X}_t) \,\middle|\, C, G, L \right\}. \tag{34} $$

   We take the lim sup rather than the lim as for some communication strategies (C, G, L) the limit may not exist.

Ideally, while implementing a solution for infinite horizon problems, we would like to use time-invariant communication strategy. This motivates the following definitions.

Definition 6 (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 7 (Stationary meta-strategy). *A meta-strategy* $\tilde{\Delta}^{\infty} = (\tilde{\Delta}_1, \tilde{\Delta}_2, \cdots)$, *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, there is no loss in optimality in restricting attention to stationary strategies (see [2]). 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 non-classical information structures. (Recall that Problem 1 has a non-classical information structure.) In this section we show that for the time-homogeneous infinite-horizon extensions of Problem 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.*

4.1. The expected discounted distortion problem

Consider a time-homogeneous infinite-horizon problem with the expected discounted distortion criterion of (33). With a slight modification to the proof of [35] one can show that the structural results of Section 2.7 are valid for this case, hence we can restrict attention to encoders belonging to $\hat{\mathscr{C}}$. Consider ${}^1\pi_t$, ${}^2\pi_t$, and ${}^3\pi_t$ as in Definition 5: they satisfy the properties of Lemma 2; further, 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 Γ denote the space $\hat{\mathscr{C}} \times \mathscr{G} \times \mathscr{L}$. We can combine (26) as

$$ {}^1\pi_{t+1} \tilde{Q}(\gamma_t)\, {}^1\pi_t, \qquad \gamma_t = \tilde{\Delta}_t({}^1\pi_t) \tag{35} $$

where

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

and

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

and 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 (33) is equivalent to the following deterministic optimization problem.

Problem 4. *Consider a deterministic system with state space ${}^1\Pi$ and action space Γ. The system dynamics are given by*

$$ {}^1\pi_{t+1} = \tilde{Q}(\gamma)\, {}^1\pi_t, \qquad \gamma_t = \tilde{\Delta}_t({}^1\pi_t) \tag{36} $$

where $\tilde{Q}$ is a known transformation and $\tilde{\Delta} : \Pi \to \Gamma$ 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 $\tilde{\Delta}^\infty = (\tilde{\Delta}_1, \tilde{\Delta}_2, \ldots)$ so as to minimize the discounted infinite horizon total distortion given by

$$ \mathscr{J}^\beta(\tilde{\Delta}^\infty) = \sum_{t=1}^{\infty} \beta^{t-1} \tilde{\rho}({}^1\pi_t, \gamma_t) \tag{37} $$

Problem 4 is a standard deterministic time-invariant infinite horizon problem with total discounted distortion (cost) criterion. Since we have assumed $\rho(\cdot)$ to be uniformly bounded, $\hat{\rho}$ and $\tilde{\rho}$ are also uniformly bounded, and therefore an optimal meta-strategy is guaranteed to exist and we have the following result:

Theorem 4. *For Problem 4 and consequently for the infinite horizon expected discounted cost problem with the performance criterion given by* (33) *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*

$$\mathcal{J}^\beta(\tilde{\Delta}^{*,\infty}) = V({}^1\pi_1), \tag{38}$$

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

$$V({}^1\pi) = \min_{\gamma \in \Gamma} \big\{ \tilde{\rho}({}^1\pi, \gamma) + \beta V\big(\tilde{Q}(\gamma)({}^1\pi)\big) \big\}, \tag{39}$$

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

$$V({}^1\pi) = \tilde{\rho}\big({}^1\pi, \tilde{\Delta}^*({}^1\pi)\big) + \beta V\big(\tilde{Q}\big(\tilde{\Delta}^*({}^1\pi)\big)({}^1\pi)\big). \tag{40}$$

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

$$(c_t^*, g_t^*, l_t^*) =: \gamma_t^* = \tilde{\Delta}^*({}^1\pi_t). \tag{41}$$

Proof. This is a standard result, see [44, Chapter 6]. □

Observe that the fixed point equation (39) can be decomposed into its "natural" sequential form as

$$
\begin{aligned}
{}^1V({}^1\pi) &= \inf_{c \in \hat{\mathscr{C}}} {}^2V({}^1Q(c)\,{}^1\pi) && \text{(42a)}\\
{}^2V({}^2\pi) &= \min_{g \in \mathscr{G}} \tilde{\rho}({}^2\pi, g) + \beta\, {}^3V({}^2\pi) && \text{(42b)}\\
{}^3V({}^3\pi) &= \min_{l \in \mathscr{L}} {}^1V({}^3Q(l)\,{}^3\pi) && \text{(42c)}
\end{aligned}
$$

These equations are the infinite horizon analogue of (29).

4.2. The average distortion per unit time problem

Consider the time-homogeneous infinite horizon problem with the average distortion per unit time criterion of (34). Using the arguments similar to the first paragraph of Section 4.1, this problem is equivalent to the following deterministic problem:

Problem 5. *Consider a deterministic system with state space* ${}^1\Pi$ *and action space* Γ*. The system dynamics are given by*

$${}^1\pi_{t+1} = \tilde{Q}(\gamma)\,{}^1\pi_t, \qquad \gamma_t = \tilde{\Delta}_t({}^1\pi_t) \tag{43}$$

where $\tilde{Q}$ *is a known transformation and* $\tilde{\Delta} : \Pi \to \Gamma$ *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*

$$\bar{\mathcal{J}}(\tilde{\Delta}^{\infty}) = \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \tilde{\rho}({}^1\pi_t, \gamma_t) \tag{44}$$

Problem 5 *cannot* be solved by taking the limit $\beta \to 1$ in the result of Theorem 4. Such a result is valid only if the problem has finite state and action space (see [45, Theorem 31.5.2]) which is not the case here. See [46] for a survey of various results connecting the expected discounted cost problem with the average cost per unit time problem.

For Problem 5 an optimal meta-strategy may not exist. However, under suitable conditions, we can guarantee the existence of meta-strategies that are arbitrarily close to optimal. Specifically, we have the following result:

Theorem 5. *For Problem 5 and correspondingly for the infinite horizon average cost per unit time problem with the performance criterion given by* (34)*, assume*

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

$$v({}^1\pi) = \min_{\gamma\in\Gamma} v\big(\tilde{Q}(\gamma)\,{}^1\pi\big) = v\big(\tilde{Q}\big(\tilde{\Delta}^*({}^1\pi)\big)\,{}^1\pi\big), \tag{45}$$

and

$$\tilde{\rho}({}^1\pi, \tilde{\Delta}^*({}^1\pi)) + r\big(\tilde{Q}\big(\tilde{\Delta}^*({}^1\pi)\big)\,{}^1\pi\big) \le v({}^1\pi) + r({}^1\pi) \le \min_{\gamma\in\Gamma}\Big\{\tilde{\rho}({}^1\pi,\gamma) + r(\tilde{Q}(\gamma)\,{}^1\pi)\Big\} + \epsilon. \tag{46}$$

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

$$\mathcal{J}_T(\tilde{\Delta}^{*,T}) \le r({}^1\pi_1) + Tv({}^1\pi_1) \le \mathcal{J}_T(\tilde{\Delta}^T) + T\epsilon \tag{47}$$

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

$$\overline{\mathcal{J}}(\tilde{\Delta}^{*,\infty}) \le v({}^1\pi_1) \le \underline{\mathcal{J}}(\tilde{\Delta}^{\infty}) + \epsilon \tag{48}$$

where

$$\overline{\mathcal{J}}(\tilde{\Delta}^{*,\infty}) := \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \tilde{\rho}\big({}^1\pi_t, \Delta^*(\pi_t)\big) \tag{49}$$

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

$$\underline{\mathcal{J}}(\tilde{\Delta}^{\infty}) := \liminf_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \tilde{\rho}\big({}^1\pi_t, \tilde{\Delta}_t(\pi_t)\big) \tag{50}$$

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

$$(c_t^*, g_t^*, l_t^*) =: \gamma_t^* = \tilde{\Delta}^*({}^1\pi_t). \tag{51}$$

Proof. This is a standard result, see [44, Chapter 7]. □

Conditions that guarantee that assumption (A1) of Theorem 5 is satisfied are fairly technical and do not provide much insight into the properties of the plant, the channel, and the cost functions that will guarantee the existence of such policies. The interested reader may look at [44, Chapter 7,

§10]. It may be possible to extend the sufficiency conditions of [47]–[49] to uncountable action spaces.

4.3. Discussion of the results

The discussion of Section 3.2 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.

5. Finite Delay

For many applications where communication delay is important, the acceptable delay is finite and fixed but non-zero. In this section we consider the case when the distortion metric tolerates a fixed-finite delay δ, 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 modelled by modifying the model of Section 2.1 as follows:

M1. The variables $\hat{X}_1, \dots, \hat{X}_\delta$ are simply not generated; the receiver spends the first δ 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), $C := (c_1, \dots, c_T)$, $G := (g_{\delta+1}, \dots, g_T)$, $L := (l_1, \dots, l_T)$ is given by

$$\mathcal{J}_T(C, G, L) := \mathbb{E}\left\{ \sum_{t=\delta+1}^{T} \rho_t(X_{t-\delta}, \hat{X}_t) \,\middle|\, C, G, L \right\}. \tag{52}$$

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

Problem 6. *Consider the model of Problem 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* (52)*.*

5.1. Transformation to a zero-delay problem

In this section we show how to convert the fixed-finite delay problem into a zero-delay problem. One method to such a conversion is the sliding window repackaging of the source presented in [29, 35]. In this paper we provide an alternative method. We believe that this new method leads to a simpler formulation of the global optimization problem.

We now show that Problem 6 can be converted into a zero-delay problem by using a *sliding window* repackaging of the source, similar to the sliding window repackaging in [29, 35]. Define the process $\{\bar{X}_t,\, t = 1, \dots, T\}$ as follows

$$\bar{X}_t := \begin{cases} (X_1, \dots, X_t), & t \leq \delta; \\ (X_{t-\delta}, \dots, X_t), & t > \delta. \end{cases} \tag{53}$$

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 $c_t \in \mathscr{C}_t$ we can find a $\bar{c}_t : \bar{\mathcal{X}}^t \to \mathcal{Z}$ such that

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

Let $\bar{\mathscr{C}}_t$ denote the collection of all $\bar{c}_t$ corresponding to all $c_t \in \mathscr{C}_t$. 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 $\bar{\mathscr{G}}_t = \emptyset$ for $t \leq \delta$ and $\bar{\mathscr{G}}_t = \mathscr{G}$. Further, we can define a modified distortion function $\bar{\rho}_t$ as follows:

$$\bar{\rho}_t(\bar{X}_t, \hat{X}_t) := \begin{cases} 0, & t \leq \delta \\ \rho_t(X_{t-\delta}, \hat{X}_t), & t > \delta \end{cases} \tag{54}$$

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

$$\begin{aligned} \mathcal{J}_T(C, G, L) &= \mathbb{E}\left\{ \sum_{t=\delta+1}^{T} \rho_t(X_{t-\delta}, \hat{X}_t) \,\middle|\, C, G, L \right\} \\ &= \mathbb{E}\left\{ \sum_{t=1}^{T} \bar{\rho}_t(\bar{X}_t, \hat{X}_t) \,\middle|\, \bar{C}, \bar{G}, L \right\} \\ &=: \bar{\mathcal{J}}_T(\bar{C}, \bar{G}, L). \end{aligned} \tag{55}$$

Hence, Problem 6 is equivalent to the following problem:

Problem 7. *Consider Problem 6 with the sliding window repackaging of the source given by* (53)*. Choose a communication strategy* $(\bar{C}^*, \bar{G}^*, L^*)$ *that is optimal with respect to the performance criterion of* (55)*, i.e.,*

$$\bar{\mathcal{J}}_T(\bar{C}^*, \bar{G}^*, L^*) = \bar{\mathcal{J}}_T^* := \min_{\substack{\bar{C} \in \bar{\mathscr{C}}^T \\ \bar{G} \in \bar{\mathscr{G}}^T \\ L \in \mathscr{L}^T}} \mathcal{J}_T(\bar{C}, \bar{G}, L), \tag{56}$$

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

5.2. Agents' beliefs and structural results

Problem 7 is a zero-delay real-time communication problem. So, the analysis and results of Sections 2–4 can be applied to this problem. We can define the encoder's and receiver's beliefs as in Definitions 3 and 4. We need to modify $\hat{A}_t(x)$ in Definition 4 as follows:

$$\hat{A}_t(x) := \Pr\left(X_{t-\delta} = x \,\middle|\, {}^2\mathfrak{R}_t\right), \quad t = \delta+1, \dots, T. \tag{57}$$

The structural properties of optimal encoders can be restated as follows:

Theorem 6 (Structure of optimal fixed-finite delay encoder). *Consider Problem 7 for any arbitrary (but fixed) decoding and memory update strategies, $\bar{G} = (g_1, \dots, g_T)$ and $L = (l_1, \dots, l_T)$, respectively. Then there is no loss in 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

$$Z_t = c_t(X_1, \dots, X_t, B_t), \qquad t = 2, \dots, \delta; \tag{58a}$$

$$Z_t = c_t(X_{t-\delta}, \dots, X_t, B_t), \quad t = \delta+1, \dots, T. \tag{58b}$$

The structural properties of optimal receivers can be restated as follows:

Theorem 7 (Structure of optimal fixed-finite delay receiver). *Consider Problem 7 for any arbitrary (but fixed) encoding and memory update strategies, $\bar{C} = (\bar{c}_1, \dots, \bar{c}_T)$ and $L = (l_1, \dots, l_T)$, respectively. Then there is no loss in optimality in restricting attention to decoding rules of the form*

$$\hat{X}_t = \hat{g}_t(\hat{A}_t) := \arg\min_{\hat{x} \in \hat{\mathcal{X}}} \sum_{x \in \mathcal{X}} \rho_t(x, \hat{x}) \hat{A}_t(x). \tag{59}$$

where $\hat{A}_t$ is defined in (57)*.*

5.3. Global optimization

We can obtain globally optimal communication strategies for Problem 7 along the lines of Section 3. For that matter, we define information states ${}^1\bar{\pi}$, ${}^2\bar{\pi}$, and ${}^3\bar{\pi}$ as follows:

$${}^i\bar{\pi}_t := \Pr\left(\bar{X}_t, {}^iB_t\right) = \begin{cases} \Pr\left(X_1, \dots, X_t, {}^iB_t\right), & t \le \delta; \\ \Pr\left(X_{t-\delta}, \dots, X_t, {}^iB_t\right), & t > \delta. \end{cases} \tag{60}$$

These information states are related in the same manner as Lemma 2. So, we can formulate an equivalent optimization problem as in Section 3.2 whose solution is given along the lines of the nested optimality equations in Section 3.3. Observe that in Problem 7, $\bar{\rho}_t(\bar{X}_t, \hat{X}_t) = 0$ for $t \le \delta$, so we can simplify the global optimization algorithm as follows.

Theorem 8 (Global optimization problem for fixed-finite delay communication). *An optimal meta-strategy $\Delta^{*,T}$ for Problem 7 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: let ${}^i\bar{\pi}_t \in {}^i\bar{\Pi} := \mathbb{P}\left\{\mathcal{X}^{\delta+1} \times {}^i\mathcal{B}\right\}$ and*

$${}^iV_{T+1}({}^1\pi) = 0 \tag{61a}$$

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

$$^1V_t(^1\bar{\pi}) = \inf_{\bar{c}_t \in \hat{\bar{\mathscr{C}}}} {}^2V_t\big({}^1Q_t(\bar{c}_t)\,{}^1\bar{\pi}\big), \tag{61b}$$

$$^2V_t(^2\bar{\pi}) = \min_{g_t \in \mathscr{G}} \Big[\hat{\rho}_t(^2\bar{\pi}, g_t) + {}^3V_t(^2\bar{\pi})\Big], \tag{61c}$$

$$^3V_t(^3\bar{\pi}) = \min_{l_t \in \mathscr{L}} {}^1V_{t+1}\big({}^3Q_t(l_t)\,{}^3\bar{\pi}\big). \tag{61d}$$

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

$$^1V_t(^1\bar{\pi}_t) = \inf_{\bar{c}_t \in \hat{\bar{\mathscr{C}}}} {}^2V_t\big({}^1Q_t(\bar{c}_t)\,{}^1\bar{\pi}_t\big), \tag{62a}$$

$$^2V_t(^2\bar{\pi}_t) = {}^3V_t(^2\bar{\pi}_t), \tag{62b}$$

$$^3V_t(^3\bar{\pi}_t) = \min_{l_t \in \mathscr{L}} {}^1V_{t+1}\big({}^3Q_t(l_t)\,{}^3\bar{\pi}_t\big). \tag{62c}$$

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.*

5.4. 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 4. Observe that a time-homogeneous model for Problem 6 *does not* imply that the transformed zero delay problem is time-homogeneous. This is because for the first δ steps $\bar{X}_t$ takes values in a space that is increasing with time. Thus, for the first δ 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 δ 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 9. *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 5.1. Therefore, we can use the results on Theorem 4 to find the value function V *and a stationary meta-strategy* Δ^* *which is optimal for* $\delta + 1$ *onwards.* V *is given by the unique fixed point of* (39) *and* Δ^* *satisfies* (40).

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* ${}^i\bar{\pi}_t \in {}^i\bar{\Pi}_t := \mathbb{P}\left\{\mathcal{X}^t \times {}^i\mathcal{B}\right\}$, *we have*

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

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

$${}^{1}V_t({}^{1}\bar{\pi}_t) = \inf_{\bar{c}_t \in \hat{\bar{\mathscr{C}}}} {}^{2}V_t\big({}^{1}Q_t(\bar{c}_t)\,{}^{1}\bar{\pi}_t\big), \tag{63a}$$

$${}^{2}V_t({}^{2}\bar{\pi}_t) = {}^{3}V_t({}^{2}\bar{\pi}_t), \tag{63b}$$

$${}^{3}V_t({}^{3}\bar{\pi}_t) = \min_{l_t \in \mathscr{L}} {}^{1}V_{t+1}\big({}^{3}Q_t(l_t)\,{}^{3}\bar{\pi}_t\big). \tag{63c}$$

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 the result of Theorem 5 to find an optimal infinite-horizon meta-strategy for $\delta + 1$ onwards, and use any policy for the initialization phase.

6. Higher-order Markov sources

In many applications the source statistics are higher-order Markov rather than first-order Markov. Such applications can be modelled by making the following modification to the model of Section 2.1:

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

$$\Pr\big(X_{t+1} = x_{t+1} \,\big|\, X^t = x^t\big) = \Pr\big(X_{t+1} = x_{t+1} \,\big|\, X^t_{t-k+1} = x^t_{t-k+1}\big)$$

This model was considered in [35, Section III-B] and it was shown that

1. There is no loss of optimality in restricting attention to encoders of the form

$$Z_t = c_t(X_{t-k+1}, {}^{1}B_t), \qquad t = k+1, \dots, T. \tag{64}$$

2. The structure of the optimal receiver is the same as that of the model of Section 2.1.

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 was transformed into a zero-delay problem. For that matter, we define the process $\{\bar{X}_t,\, t = 1, \dots, T\}$ as follows:

$$\bar{X}_t := \begin{cases} (X_1, \dots, X_t), & t \le k; \\ (X_{t-k+1}, \dots, X_t), & t > k. \end{cases} \tag{65}$$

Observe that $\{\bar{X}_t,\, t = 1, \dots, T\}$ is a first-order Markov process. The structural results of [35] state that we can restrict attention to encoders of the form

$$Z_t = c_t(\bar{X}_t, {}^{1}B_t) \tag{66}$$

Let $\bar{\mathscr{C}}_t$ denote the class of all encoders at time t of the form (66).

The channel and the receiver are the same as in the model of Section 2.1. We defined a modified distortion function $\bar{\rho}_t$ as follows:

$$\bar{\rho}_t(\bar{X}_t, \hat{X}_t) := \rho_t(X_t, \hat{X}_t). \tag{67}$$

With these modifications we can formulate an optimization problem similar to Problem 2. The finite horizon problem can be solved in the same manner as Section 3. 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, and
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 5.3.

7. Channels with Memory

So far in this paper, we have assumed that we have a memoryless channel. In a realistic scenario, the channel has memory arising due to either a changing physical environment (e.g., wireless fading channel) or the present output depending on past inputs (e.g. an ISI channel) 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.

7.1. Problem Formulation

We consider the same problem as in Section 2 with one change. Instead of assuming the channel to be memoryless, we assume that it has memory in the form of a *state*. Let S_t denote the state of the channel at time t. We assume that the state belongs to a finite set $\mathcal{S}$. The channel can be described as

$$Y_t = h_t(Z_t, N_t, S_{t-1}) \tag{68}$$

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 $t-1$. We 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}. We 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

$$S_t = \hat{h}_t(Z_t, N_t, S_{t-1}) \tag{69}$$

where $\hat{h}_t(\cdot)$ is the channel update function. We 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 2.1. The sequential ordering of the variables is shown in Figure 3. We are interested in the following optimization problem.

Problem 8. *Consider the real-time communication system of Section 2.1 with a channel with memory given by* (68) *and* (69)*. Choose a communication strategy* (C^*, G^*, L^*) *that is optimal with respect to the performance criterion of* (5)*, i.e.,*

$$\mathcal{J}_T(C^*, G^*, L^*) = \mathcal{J}_T^* := \min_{\substack{C \in \mathscr{C}^T \\ G \in \mathscr{G}^T \\ L \in \mathscr{L}^T}} \mathcal{J}_T(C, G, L), \tag{70}$$

where $\mathscr{C}^T$*,* $\mathscr{G}^T$*, and* $\mathscr{L}^T$ *are defined in Problem 1.*

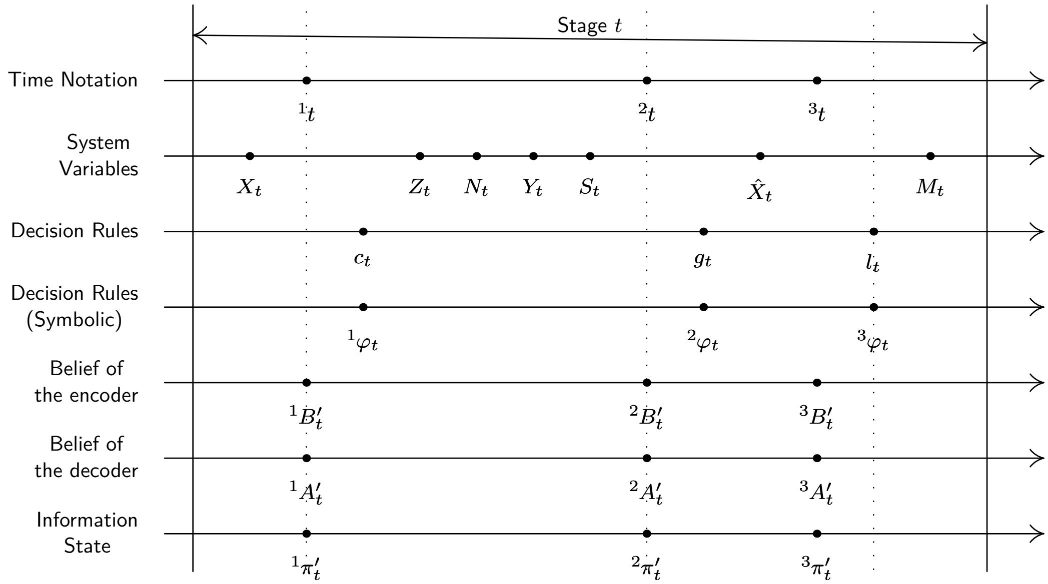

Figure 3: The sequential ordering of the variables of the real time communication system for channels with memory. ^1t, ^2t, and ^3t are refinements of stage t.

7.2. Agents' beliefs and their evolution

For this model, we define information fields at the encoder and the receiver as in Definitions 1 and 2. For ease of notation, we define ${}^iS_t$ as the state of the channel at time ${}^it$, i.e.,

$$ {}^1S_t = S_{t-1}, \qquad {}^2S_t = S_t, \qquad {}^3S_t = S_t. \tag{71}$$

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

Definition 8 (Encoder's Beliefs). *Let ${}^iB'_t$ denote the encoder's belief about the receiver's observation and the state of the channel at time ${}^it$, $i = 1, 2, 3$, $t = 1, \dots, T$. Then for ${}^ir \in {}^i\mathcal{R}$, and ${}^is \in \mathcal{S}$*

$$ {}^iB'_t({}^ir, {}^is) := \Pr\left({}^iR_t = {}^ir, {}^iS_t = {}^is \,\middle|\, {}^i\mathfrak{E}_t\right). \tag{72}$$

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

Definition 9 (Receiver's Beliefs). *Let ${}^iA'_t$ denote the receiver's belief about the encoder's observations and the channel state at time ${}^it$, $i = 1, 2, 3$, $t = 1, \dots, T$. Then for ${}^ie_t \in {}^i\mathcal{E}_t$ and ${}^is \in \mathcal{S}$*

$$ {}^iA'_t({}^ie_t, {}^is) := \Pr\left({}^iE_t = {}^ie_t, {}^iS_t = {}^is \,\middle|\, {}^i\mathfrak{R}_t\right). \tag{73}$$

Let ${}^i\mathcal{A}'_t := \mathbb{P}\left\{{}^i\mathcal{E}_t \times \mathcal{S}\right\}$ 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 ${}^2t$, given by (16)*.*

The beliefs of the encoder evolve in a manner similar to Lemma 1.

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

$$ {}^2B'_t = {}^1F'_t({}^1B'_t, Z_t) \tag{74a} $$
$$ {}^3B'_t = {}^2B'_t \tag{74b} $$
$$ {}^1B'_{t+1} = {}^3F'({}^3B'_t, l_t) \tag{74c} $$

This is proved in Appendix C.

7.3. 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 10 (Structure of optimal encoders for channels with memory). *Consider Problem 8 with arbitrary but fixed decoding and memory update strategies, $G := (g_1, \dots, g_T)$, and $L := (l_1, \dots, l_T)$, respectively. Then there is no loss of optimality in restricting attention to encoding rules of the form*

$$ Z_t = c_t(X_t, {}^1B'_t), \qquad t = 2, \dots, T. \tag{75} $$

We follow the methodology of the alternative proof of the structural results in [35].

Proof. We 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 3 imply that the process $\{(X_t, {}^1B'_t),\ t = 1, \dots, T\}$ is a controlled Markov chain 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 ${}^1\varphi^t$

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

where (a) follows from the Markov nature of the source and Lemma 3. 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.

Further, the conditional expected instantaneous cost can be written as

$$
\begin{aligned}
\mathbb{E}\left\{\rho_t(X_t, \hat{X}_t) \,\middle|\, {}^2\mathfrak{C}_t\right\} &= \mathbb{E}\left\{\rho_t(X_t, \hat{X}_t) \,\middle|\, X^t, Z^t; c^t, g^t, l^{t-1}\right\} \\
&= \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \rho_t(X_t, g_t(y_t, m_{t-1})) \Pr\left(Y_t = y_t, M_{t-1} = m_{t-1} \,\middle|\, X^t, Z^t; c^t, g^t, l^{t-1}\right) \\
&= \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \rho_t(X_t, g_t(y_t, m_{t-1}))\, {}^3B'_t(y_t, m_{t-1}) \\
&\overset{(b)}{=} \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \rho_t(X_t, g_t(y_t, m_{t-1}))\, {}^1F_t({}^1B'_t, Z_t)(y_t, m_{t-1}) \\
&=: \tilde{\rho}_t(X_t, {}^1B'_t, Z_t, g_t)
\end{aligned}
$$

where (b) follows from Lemma 3. Thus, the total expected cost can be written as

$$\begin{aligned}\mathbb{E}\left\{\sum_{t=1}^{T}\rho_t(X_t,\hat{X}_t)\,\middle|\, C,G,L\right\} &= \mathbb{E}\left\{\sum_{t=1}^{T}\mathbb{E}\left\{\rho_t(X_t,\hat{X}_t)\,\middle|\,{}^2\mathfrak{E}_t\right\}\,\middle|\, C,G,L\right\}\\ &= \mathbb{E}\left\{\sum_{t=1}^{T}\tilde{\rho}_t(X_t,{}^1B'_t,Z_t,g_t)\,\middle|\, C,G,L\right\} \qquad (76)\end{aligned}$$

Hence, from the encoder's point of view, we have a perfectly observed controlled Markov process $\{(X_t,{}^1B'_t),\, t=1,\dots,T\}$ 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 [2, Chapter 6] we know that there is no loss of optimality in restricting attention to encoding rules of the form (75). $\square$

Theorem 11 (Structure of optimal decoders for channels with memory). *Consider Problem 8 for any arbitrary but fixed encoding and memory update strategies $C := (c_1,\dots,c_T)$ and $L := (l_1,\dots,l_T)$, respectively. Then there is no loss of optimality in restricting attention to decoding rules of the form*

$$\hat{X}_t = \hat{g}_t(\hat{A}_t) := \arg\min_{\hat{x}\in\hat{\mathcal{X}}}\sum_{x\in\mathcal{X}}\rho_t(x,\hat{x})\hat{A}_t(x). \qquad (77)$$

Proof. We 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 $\mathbb{E}\left\{\rho_t(X_t,\hat{X}_t)\,\middle|\,{}^2\mathfrak{R}_t\right\}$ for each time t. This conditional expected instantaneous distortion can be written as

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

and is minimized by the decoding rule given in (77). $\square$

We can use the structural results of Theorem 10 to choose encoding rules from a space of functions that is not changing with time. Let $\hat{\mathscr{C}}'$ denote the space of functions from $\mathcal{X}\times{}^1\mathcal{B}'$ to $\mathcal{Z}$, and $\hat{\mathscr{C}}'^{,t}$ denote $\hat{\mathscr{C}}'\times\cdots\times\hat{\mathscr{C}}'$ (T-times). Then, the result of Theorem 10 implies that

Corollary 3. *For Problem 8, the optimal performance $\mathcal{J}_T^*$ can be determined by*

$$\mathcal{J}_T^* := \inf_{\substack{C\in\hat{\mathscr{C}}'^{,T}\\ G\in\mathscr{G}^T\\ L\in\mathscr{L}^T}}\mathcal{J}_T(C,G,L). \qquad (78)$$

Consequently, we can reformulate Problem 8 in a manner similar to the reformulation presented in Section 2.8.

Problem 9. *Under the assumptions of Problem 8 choose a communication strategy (C^*,G^*,L^*) belonging to $(\hat{\mathscr{C}}'^{,T}\times\mathscr{G}^T\times\mathscr{L}^T)$ that is optimal with respect to the performance criterion of* (78).

7.4. Global optimization

We use the structural results of Section 7.3 to identify information states to obtain a sequential decomposition of Problem 9. This decomposition determines a globally optimal communication strategy for Problem 9 (and consequently for Problem 8). We follow the philosophy and approach of Sections 3 and 4. For that matter, we define the following.

Definition 10. *Define ${}^1\pi'$, ${}^2\pi'$, and ${}^3\pi'$ as follows:*

$$ {}^i\pi'_t = \Pr\left(X_t, {}^iB'_t \,\middle|\, {}^i\varphi^{t-1}\right), \qquad i = 1,2,3. \tag{79}$$

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

We can show that these probability measures are related to one another in a manner similar to Lemma 2. Consequently, we can reformulate Problem 9 in the same manner as Problem 2 is reformulated to Problem 3; we can further obtain a sequential decomposition similar to the nested optimality equations of Theorem 3. This sequential decomposition can be extended to the infinite horizon along the lines of the results of Section 4. All these results can be derived along the lines of the analysis in Sections 3 and 4, so we omit the details here. Next we present a special case of a channel with memory and show that for this case the structure of optimal encoders is simplified and the complexity of the information states is reduced.

7.5. A special case: the ISI channel

In this section, we consider a special case of channel with memory—the ISI (inter-symbol interference) channel. In an ISI channel the channel state is a function of the previous state and the channel input, i.e.,

$$S_t = \hat{h}_t(Z_t, S_t).$$

Thus, the channel state information is available at the encoder, and the encoder's belief simplify as follows:

Lemma 4. *For $i = 1,2,3$, and $t = 1,\dots,T$*

$$ {}^iB'_t({}^ir, s) = {}^iB_t({}^ir)\,\mathbb{1}\left[{}^iS_t = s\right] \tag{80}$$

where ${}^iB_t$ are as defined in Definition 3.

Proof. Consider ${}^1r \in {}^1\mathcal{R}$, $s \in \mathcal{S}$, and ${}^1e_t \in {}^1\mathcal{E}_t$. Then,

$$\begin{aligned}
{}^1b'_t({}^1r, s) &= \Pr\left({}^1R_t = {}^1r, S_{t-1} = s \,\middle|\, {}^1E_t = {}^1e_t; {}^1\varphi^{t-1}\right) \\
&= \Pr\left(S_{t-1} = s \,\middle|\, {}^1R_t = {}^1r, {}^1E_t = {}^1e_t; {}^1\varphi^{t-1}\right) \Pr\left({}^1R_t = {}^1r \,\middle|\, {}^1E_t = {}^1e_t; {}^1\varphi^{t-1}\right) \\
&= \mathbb{1}\left[s = \hat{h}_{t-1}\big(c_{t-1}({}^1e_{t-1}), \hat{h}_{t-2}(\dots)\big)\right] {}^1b_t({}^1r) \\
&= \mathbb{1}\left[S_{t-1} = s\right] {}^1b_t({}^1r)
\end{aligned} \tag{81}$$

Similar arguments hold for $i = 2$ and 3. □

Using the decomposition of Lemma 4, the structural results of Theorem 10 and the information states of Definition 10 can be simplified as follows:

Corollary 4. *For an ISI channel, there is no loss of optimality in restricting attention to encoding rules of the form*

$$Z_t = c_t(X_t, {}^1B_t, S_{t-1}). \tag{82}$$

Furthermore, the information states for global optimization simplify to

$$
{}^{i}\pi'_t = \Pr\left(X_t, {}^{i}B_t, {}^{i}S_t\right). \tag{83}
$$

The information states of (83) can be used to obtain a globally optimal communication strategies in a manner similar to Theorem 3. Optimal communication strategies for infinite horizon variations can be determined along the lines of Section 4.

8. Computational issues

As mentioned earlier, all the variations of real-time communication problems considered in this paper are decentralized stochastic optimization problems consequently, cannot be solved by classical Markov decision theory. However, the results presented in this paper show that an appropriate choice of information state 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 informations state, in contrast to the problems with classical information structure (which include Markov decision problems) where the objective is to choose an optimal action for each choice of information state. This difference is the key reason for the difficulty in numerically solving these problems. We would like to assert that the computational complexity of the solution is not a shortcoming of our approach; it is an intrinsic feature of all decentralized stochastic optimization problems. It was shown in [52] that decentralized stochastic optimization problems are NEXP complete, i.e., they *provably* do not admit a polynomial time solution.

The result of this paper shows that we can write sufficient conditions for finding optimal meta-strategies (the nested optimality conditions of Theorem 3 and the fixed point equations of Theorems 4 and 5) 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 modelling 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

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

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

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

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

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.

9. 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 of 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.

The objective of the design of a communication system is to find communication strategies that perform nearly optimally and are easy to implement. For communication systems with 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.

This approach works even for communication systems with finite but sufficiently large delay constraints. However, this approach fails for communication systems with small delay constraints because for information theoretic bounds are not tight for small values of delay and consequently, fundamental limits of performance are not known. As a result, there is no benchmark for performance evaluation of communication strategies, and we cannot determine whether or not a 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 the existence of good approximation techniques for solving the optimality equations of Section 3 and 4. (As pointed out in Section 8, the approximation techniques for POMDP are an obvious candidate, but we do not know if they are provably good approximations for the optimality equations of Sections 3 and 4). 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.

10. 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 [38]. In our solution, we exploit the structural results of [35] to identify information states that belong to a space that does not increase with time. This is in contrast to the information states in [38] 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.

Acknowledgements

This research was supported in part by ONR Grant N00014-03-1-0232, NSF Grant CCR-0325571, and NASA Grant NNX06AD47G. We are grateful to A. Anastasopoulos and S. Pradhan for insightful discussions. We are also grateful to the reviewers for their detailed and careful comments which helped us to significantly improve the presentation of the paper.

Appendix A. Relation between the beliefs

Proof of Lemma 1. We prove the three results separately.

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

$$
\begin{aligned}
{}^2b_t({}^2r_t) &= \Pr\left({}^2R_t = {}^2r_t \,\middle|\, {}^2E_t = {}^2e_t; {}^2\varphi^{t-1}\right) \\
&= \Pr\left(Y_t = y_t, M_{t-1} = m_{t-1} \,\middle|\, {}^1E_t = {}^1e_t, Z_t = z_t; {}^1\varphi^{t-1}, c_t\right) \\
&= \Pr\left(Y_t = y_t \,\middle|\, M_{t-1} = m_{t-1}, {}^1E_t = {}^1e_t, Z_t = z_t; {}^1\varphi^{t-1}, c_t\right) \\
&\quad \times \Pr\left(M_{t-1} = m_{t-1} \,\middle|\, {}^1E_t = {}^1e_t, Z_t = z_t; {}^1\varphi^{t-1}, c_t\right) \\
&\stackrel{(a)}{=} P_{N_t}\big(n_t \in \mathcal{N} : y_t = h_t(z_t, n_t)\big) \\
&\quad \times \Pr\left(M_{t-1} = m_{t-1} \,\middle|\, {}^1E_t = {}^1e_t; {}^1\varphi^{t-1}\right) \\
&= P_{N_t}\big(n_t \in \mathcal{N} : y_t = h_t(z_t, n_t)\big) \times {}^1b_t(m_{t-1}) \\
&=: {}^1F_t({}^1b_t, z_t)(y_t, m_{t-1}) = {}^1F_t({}^1b_t, z_t)({}^2r_t)
\end{aligned} \tag{84}
$$

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 ${}^3\varphi^{t-1} = ({}^2\varphi^{t-1}, g_t)$. Recall that ${}^3E_t = {}^2E_t$ and ${}^3R_t = {}^2R_t$. Then,

$$
\begin{aligned}
{}^3b_t({}^3r_t) &= \Pr\left({}^3R_t = {}^3r_t \,\middle|\, {}^3E_t = {}^3e_t; {}^3\varphi^{t-1}\right) \\
&= \Pr\left({}^2R_t = {}^3r_t \,\middle|\, {}^2E_t = {}^3e_t; {}^2\varphi^{t-1}, g_t\right) \\
&\stackrel{(b)}{=} \Pr\left({}^2R_t = {}^3r_t \,\middle|\, {}^2E_t = {}^3e_t; {}^2\varphi^{t-1}, g_t\right) \\
&= \Pr\left({}^2R_t = {}^3r_t \,\middle|\, {}^2E_t = {}^3e_t; {}^2\varphi^{t-1}\right) \\
&= {}^2b_t({}^3r_t)
\end{aligned} \tag{85}
$$

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 ${}^1\varphi^t = ({}^3\varphi^{t-1}, l_t)$. Then

$$
\begin{aligned}
{}^1b_{t+1}({}^1r_{t+1}) &= \Pr\left({}^1R_{t+1} = {}^1r_{t+1} \,\middle|\, {}^1E_{t+1} = {}^1e_{t+1}; {}^1\varphi^t\right) \\
&= \Pr\left(M_t = m_t \,\middle|\, {}^3E_t = {}^3e_t, X_{t+1} = x_{t+1}; {}^3\varphi^{t-1}, l_t\right) \\
&= \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \Pr\left(Y_t = y_t, M_t = m_t, M_{t-1} = m_{t-1} \,\middle|\, {}^3E_t = {}^3e_t, X_{t+1} = x_{t+1}; {}^3\varphi^{t-1}, l_t\right) \\
&= \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \Pr\left(M_t = m_t \,\middle|\, Y_t = y_t, M_t = m_{t-1}, {}^3E_t = {}^3e_t, X_{t+1} = x_{t+1}; {}^3\varphi^{t-1}, l_t\right) \\
&\qquad \times \Pr\left(Y_t = y_t, M_{t-1} = m_{t-1} \,\middle|\, {}^3E_t = {}^3e_t, X_{t+1} = x_{t+1}; {}^3\varphi^{t-1}, l_t\right) \\
&\stackrel{(c)}{=} \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \mathbb{1}\left[m_t = l_t(y_t, m_{t-1})\right] \Pr\left(Y_t = y_t, M_{t-1} = m_{t-1} \,\middle|\, {}^3E_t = {}^3e_t; {}^3\varphi^{t-1}\right) \\
&= \sum_{\substack{y_t \in \mathcal{Y} \\ m_{t-1} \in \mathcal{M}}} \mathbb{1}\left[m_t = l_t(y_t, m_{t-1})\right] {}^3b_t(y_t, m_{t-1}) \\
&=: {}^3F({}^3b_t, l_t)(m_t) = {}^3F({}^3b_t, l_t)({}^1r_{t+1})
\end{aligned} \tag{86}
$$

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

Appendix B. Relation between information states

Proof of Lemma 2. We prove the three results separately.

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

$$\begin{aligned}
{}^2\pi_t(x_t, {}^2b_t) &= \Pr\left(X_t = x_t, {}^2B_t = {}^2b_t \,\middle|\, {}^2\varphi^{t-1}\right) \\
&= \int_{{}^1b_t \in {}^1\mathcal{B}} \sum_{z_t \in \mathcal{Z}} \Pr\left(X_t = x_t, Z_t = z_t, {}^2B_t = {}^2b_t, {}^1B_t = {}^1b_t \,\middle|\, {}^1\varphi^{t-1}, c_t\right) d\,{}^1b_t \\
&= \int_{{}^1b_t \in {}^1\mathcal{B}} \sum_{z_t \in \mathcal{Z}} \Pr\left({}^2B_t = {}^2b_t \,\middle|\, X_t = x_t, Z_t = z_t, {}^1B_1 = {}^1b_1; {}^1\varphi^{t-1}, c_t\right) d\,{}^1b_t \\
&\quad \times \Pr\left(Z_t = z_t \,\middle|\, X_t = x_t, {}^1B_t = {}^1b_t; {}^1\varphi^{t-1}, c_t\right) \\
&\quad \times \Pr\left(X_t = x_t, {}^1B_t = {}^1b_t \,\middle|\, {}^1\varphi^{t-1}, c_t\right) d\,{}^1b_t \\
&\overset{(a)}{=} \int_{{}^1b_t \in {}^1\mathcal{B}} \sum_{z_t \in \mathcal{Z}} \mathbb{1}\left[{}^2b_t = {}^1F_t({}^1b_t, z_t)\right] \mathbb{1}\left[z_t = c_t(x_t, {}^1b_t)\right] \\
&\quad \times \Pr\left(X_t = x_t, {}^1B_t = {}^1b_t \,\middle|\, {}^1\varphi^{t-1}\right) d\,{}^1b_t \\
&= \int_{{}^1b_t \in {}^1\mathcal{B}} \sum_{z_t \in \mathcal{Z}} \mathbb{1}\left[{}^2b_t = {}^1F_t({}^1b_t, z_t)\right] \mathbb{1}\left[z_t = c_t(x_t, {}^1b_t)\right] {}^1\pi_t(x_t, {}^1b_t) d\,{}^1b_t \\
&=: \left({}^1Q(c_t)\,{}^1\pi_t\right)(x_t, {}^2b_t)
\end{aligned} \tag{87}$$

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

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

$$\begin{aligned}
{}^3\pi_t(x_t, {}^3b_t) &= \Pr\left(X_t = x_t, {}^3B_t = {}^3b_t \,\middle|\, {}^3\varphi^{t-1}\right) \\
&= \Pr\left(X_t = x_t, {}^2B_t = {}^3b_t \,\middle|\, {}^2\varphi^{t-1}, g_t\right) \\
&\overset{(b)}{=} \Pr\left(X_t = x_t, {}^2B_t = {}^3b_t \,\middle|\, {}^2\varphi^{t-1}\right) \\
&= {}^2\pi_t(x_t, {}^3b_t)
\end{aligned} \tag{88}$$

   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 ${}^1\varphi^t = ({}^3\varphi^{t-1}, l_t)$. Then a component of ${}^1\pi_{t+1}$ is given by

$$\begin{aligned}
{}^1\pi_{t+1}(x_{t+1}, {}^1b_{t+1}) &= \Pr\left(X_{t+1} = x_{t+1}, {}^1B_{t+1} = {}^1b_{t+1} \,\middle|\, {}^1\varphi^t\right) \\
&= \int_{{}^3b_t \in {}^3\mathcal{B}} \sum_{x_t \in \mathcal{X}_t} \Pr\left(X_{t+1} = x_{t+1}, X_t = x_t, {}^1B_{t+1} = {}^1b_{t+1}, {}^3B_t = {}^3b_t \,\middle|\, {}^3\varphi^{t-1}, l_t\right) d\,{}^3b_t \\
&\overset{(c)}{=} \int_{{}^3b_t \in {}^3\mathcal{B}} \sum_{x_t \in \mathcal{X}_t} P_{X_{t+1}|X_t}(x_{t+1} \mid x_t)\, \mathbb{1}\left[{}^1b_{t+1} = {}^3F({}^3b_t, l_t)\right] \\
&\quad \times \Pr\left(X_t = x_t, {}^3B_t = {}^3b_t \,\middle|\, {}^3\varphi^{t-1}\right) d\,{}^3b_t
\end{aligned}$$

$$= \int_{{}^3b_t \in {}^3\mathcal{B}} \sum_{x_t \in \mathcal{X}_t} P_{X_{t+1}|X_t}(x_{t+1} \mid x_t)\, \mathbb{1}\left[{}^1b_{t+1} = {}^3F({}^3b_t, l_t)\right] {}^3\pi_t(x_t, {}^3b_t) d\,{}^3b_t$$

$$=: \big({}^3Q_t(l_t)\,{}^3\pi_t\big)(x_{t+1}, {}^1b_{t+1}) \tag{89}$$

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

$$\mathbb{E}\left\{\rho_t(X_t, \hat{X}_t) \,\middle|\, {}^2\varphi^{t-1}, g_t\right\} = \sum_{\substack{x_t \in \mathcal{X} \\ {}^2r_t \in {}^2\mathcal{R}_t}} \rho_t\big(x_t, g_t({}^2r_t)\big) \Pr\left(X_t = x_t, {}^2R_t = {}^2r_t \,\middle|\, {}^2\varphi^{t-1}, g_t\right). \tag{90}$$

Now consider

$$\begin{aligned}
\Pr\left(X_t = x_t, {}^2R_t = {}^2r_t \,\middle|\, {}^2\varphi^{t-1}, g_t\right) &= \int_{{}^2b_t \in {}^2\mathcal{B}} \Pr\left(X_t = x_t, {}^2R_t = {}^2r_t, {}^2B_t = {}^2b_t \,\middle|\, {}^2\varphi^{t-1}, g_t\right) d\,{}^2b_t \\
&= \int_{{}^2b_t \in {}^2\mathcal{B}} \Pr\left({}^2R_t = {}^2r_t \,\middle|\, X_t = x_t, {}^2B_t = {}^2b_t; {}^2\varphi^{t-1}, g_t\right) \\
&\quad \times \Pr\left(X_t = x_t, {}^2B_t = {}^2b_t \,\middle|\, {}^2\varphi^{t-1}, g_t\right) d\,{}^2b_t \\
&\overset{(d)}{=} \int_{{}^2b_t \in {}^2\mathcal{B}} {}^2b_t({}^2r_t) \Pr\left(X_t = x_t, {}^2B_t = {}^2b_t \,\middle|\, {}^2\varphi^{t-1}\right) d\,{}^2b_t \\
&= \int_{{}^2b_t \in {}^2\mathcal{B}} {}^2b_t({}^2r_t)\, {}^2\pi_t(x_t, {}^2b_t) d\,{}^2b_t.
\end{aligned} \tag{91}$$

Substituting the result of (91) in (90) we get,

$$\begin{aligned}
\mathbb{E}\left\{\rho_t(X_t, \hat{X}_t) \,\middle|\, {}^2\varphi^{t-1}, g_t\right\} = \sum_{\substack{x_t \in \mathcal{X} \\ {}^2r_t \in {}^2\mathcal{R}_t}} \rho_t\big(x_t, g_t({}^2r_t)\big) \times &\int_{b_t \in \mathcal{B}_t} {}^2b_t({}^2r_t)\, {}^2\pi_t(x_t, {}^2b_t) d\,{}^2b_t \\
&=: \hat{\rho}_t({}^2\pi_t, g_t).
\end{aligned} \tag{92}$$

Observe that (87) and (89) imply that the transformations ${}^1Q_t$ and ${}^3Q_t$ are linear in the sense that if ${}^1\pi_t^{(1)}, {}^1\pi_t^{(2)} \in {}^1\Pi_t$, $c_t \in \hat{\mathscr{C}}$ and $\lambda \in [0, 1]$, then

$${}^1Q_t(c_t)\big(\lambda\, {}^1\pi_t^{(1)} + (1-\lambda)\, {}^1\pi_t^{(2)}\big) = \lambda\, {}^1Q_t(c_t)\, {}^1\pi_t^{(1)} + (1-\lambda)\, {}^1Q_t(c_t)\, {}^1\pi_t^{(2)} \tag{93}$$

and similar relation holds for ${}^3Q_t$. □

Appendix C. Relation between the beliefs for channels with memory

Proof of Lemma 3. 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 ${}^2\varphi^{t-1} = ({}^1\varphi^{t-1}, c_t)$. Then

$$
\begin{aligned}
{}^2b'_t({}^2r_t, s_t) &= \Pr\left({}^2R_t = {}^2r_t, {}^2S_t = s_t \,\middle|\, {}^2E_t = {}^2e_t; {}^2\varphi^{t-1}\right) \\
&= \Pr\left(Y_t = y_t, M_{t-1} = m_{t-1}, S_t = s_t \,\middle|\, {}^1E_t = {}^1e_t, Z_t = z_t; {}^1\varphi^{t-1}, c_t\right) \\
&= \sum_{z_{t-1}\in\mathcal{S}} \Pr\left(Y_t = y_t, M_{t-1} = m_{t-1}, S_t = s_t, S_{t-1} = s_{t-1} \,\middle|\, {}^1E_t = {}^1e_t, Z_t = z_t; {}^1\varphi^{t-1}, c_t\right) \\
&= \sum_{z_{t-1}\in\mathcal{S}} \Pr\left(Y_t = y_t, S_t = s_t \,\middle|\, {}^1E_t = {}^1e_t, Z_t = z_t, S_{t-1} = s_{t-1}, M_{t-1} = m_{t-1}; {}^1\varphi^{t-1}, c_t\right) \\
&\qquad \times \Pr\left(M_{t-1} = m_{t-1}, S_{t-1} = s_{t-1} \,\middle|\, {}^1E_t = {}^1e_t, Z_t = z_t; {}^1\varphi^{t-1}, c_t\right) \\
&\overset{(a)}{=} \sum_{z_{t-1}\in\mathcal{S}} P_{N_t}\big(n_t \in \mathcal{N} : y_t = h_t(z_t, n_t, s_{t-1}), s_t = \hat{h}_t(z_t, n_t, s_{t-1})\big) \\
&\qquad \times \Pr\left(M_{t-1} = m_{t-1}, S_{t-1} = s_{t-1} \,\middle|\, {}^1E_t = {}^1e_t, Z_t = z_t; {}^1\varphi^{t-1}\right) \\
&= \sum_{z_{t-1}\in\mathcal{S}} P_{N_t}\big(n_t \in \mathcal{N} : y_t = h_t(z_t, n_t, s_{t-1}), s_t = \hat{h}_t(z_t, n_t, s_{t-1})\big) \\
&\qquad \times {}^1b'_t(m_{t-1}, s_{t-1}) \\
&=: {}^1F'_t({}^1b'_t, z_t)(y_t, m_{t-1}) = {}^1F'_t({}^1b'_t, z_t)({}^2r_t). \qquad (94)
\end{aligned}
$$

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 ${}^3\varphi^{t-1} = ({}^2\varphi^{t-1}, g_t)$. Recall that ${}^3E_t = {}^2E_t$, ${}^3R_t = {}^2R_t$, and ${}^3S_t = {}^2S_t$. Then,

$$
\begin{aligned}
{}^3b'_t({}^3r_t, s_t) &= \Pr\left({}^3R_t = {}^3r_t, {}^3S_t = s_t \,\middle|\, {}^3E_t = {}^3e_t; {}^3\varphi^{t-1}\right) \\
&= \Pr\left({}^2R_t = {}^3r_t, {}^2S_t = s_t \,\middle|\, {}^3E_t = {}^3e_t; {}^2\varphi^{t-1}, g_t\right) \\
&\overset{(b)}{=} \Pr\left({}^2R_t = {}^3r_t, {}^2S_t = s_t \,\middle|\, {}^3E_t = {}^3e_t; {}^2\varphi^{t-1}\right) \\
&= {}^2b'_t({}^3r_t, s_t). \qquad (95)
\end{aligned}
$$

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 ${}^1\varphi^t = ({}^3\varphi^{t-1}, l_t)$. Recall that ${}^1S_{t+1} = {}^3S_t = S_t$. Then

$$
\begin{aligned}
&{}^1b'_{t+1}({}^1r_{t+1}, s_t) = \Pr\left({}^1R_{t+1}, {}^1S_{t+1} = s_t \,\middle|\, {}^1E_{t+1} = {}^1e_{t+1}; {}^1\varphi^t\right) \\
&= \Pr\left(M_t = m_t, {}^3S_t = s_t \,\middle|\, {}^3E_t = {}^3e_t, X_{t+1} = x_{t+1}; {}^3\varphi^{t-1}, l_t\right) \\
&= \sum_{\substack{y_t\in\mathcal{Y}\\ m_{t-1}\in\mathcal{M}}} \Pr\left(Y_t = y_t, M_t = m_t, M_{t-1} = m_{t-1}, {}^3S_t = s_t \,\middle|\, {}^3E_t = {}^3e_t, X_{t+1} = x_{t+1}; {}^3\varphi^{t-1}, l_t\right) \\
&= \sum_{\substack{y_t\in\mathcal{Y}\\ m_{t-1}\in\mathcal{M}}} \Pr\left(M_t = m_t \,\middle|\, Y_t = y_t, M_{t-1} = m_{t-1}, {}^2S_t = s_t, {}^3E_t = {}^3e_t, X_{t+1} = x_{t+1}; {}^3\varphi^{t-1}, l_t\right) \\
&\qquad \times \Pr\left(Y_t = y_t, M_{t-1} = m_{t-1}, {}^2S_t = s_t \,\middle|\, {}^3E_t = {}^3e_t, X_{t+1} = x_{t+1}; {}^3\varphi^{t-1}, l_t\right) \\
&\overset{(c)}{=} \sum_{\substack{y_t\in\mathcal{Y}\\ m_{t-1}\in\mathcal{M}}} \mathbb{1}\left[m_t = l_t(y_t, m_{t-1})\right] \Pr\left(Y_t = y_t, M_{t-1} = m_{t-1}, {}^2S_t = s_t \,\middle|\, {}^3E_t = {}^3e_t; {}^3\varphi^{t-1}\right) \\
&= \sum_{\substack{y_t\in\mathcal{Y}\\ m_{t-1}\in\mathcal{M}}} \mathbb{1}\left[m_t = l_t(y_t, m_{t-1})\right] {}^3b'_t(y_t, m_{t-1}, s_t) \\
&=: {}^3F'({}^3b'_t, l_t)(m_t) = {}^3F'({}^3b'_t, l_t)({}^1r_{t+1})
\end{aligned}
$$

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

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 22, pp. 379-423, Jul. 1948.

[2] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation Identification and Adaptive Control*, Prentice Hall, 1986.

[3] H. S. Witsenhausen, "On the sequences of pairs of dependent random variables," *SIAM Journal of Applied Mathematics*, vol. 28, pp. 100-113, 1975.

[4] N. T. Gaarder and D. Slepian, "On optimal finite-state digital transmission systems," in *International Symposium on Information Theory*, 1979, Grignano, Italy.

[5] ——, "On optimal finite-state digital transmission systems," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 167–186, 1982.

[6] T. Linder and G. Lugosi, "A zero-delay sequential scheme for lossy coding of individual sequences," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2533–2538, 2001.

[7] N. Merhav and I. Kontoyiannis, "Source coding exponents for zero-delay coding with finite memory," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 609–625, 2003.

[8] T. Weissman and N. Merhav, "On limited-delay lossy coding and filtering of individual sequences," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 721–733, 2002.

[9] A. Gyorgy, T. Linder, and G. Lugosi, "Efficient adaptive algorithms and minimax bounds for zero-delay lossy source coding," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2337–2347, 2004.

[10] S. P. Lloyd, "Rate versus fidelity for binary source," *Bell System Technical Journal*, vol. 56, pp. 427-437, 1977.

[11] P. Piret, "Causal sliding block encoders with feedback," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 237–240, Mar. 1979.

[12] D. L. Neuhoff and R. K. Gilbert, "Causal source codes," *IEEE Trans. Inf. Theory*, vol. 28, no. 6, pp. 701-713, Sep. 1982.

[13] T. Linder and R. Zamir, "Causal source coding for stationary sources with high resolution," in *International Symposium on Information Theory*, 2001, Washington, D.C..

[14] T. Fine, "Properties of an optimum digital system and applications," *IEEE Trans. Inf. Theory*, vol. 10, no. 4, pp. 287- 296, Oct. 1964.

[15] B. McMillan, "Communication systems which minimize coding noise," *Bell System Technical Journal*, vol. 48, no. 9, pp. 3091-3113, Sep. 1969.

[16] T. Ericson (1979). Delayless information transmission. Internal Pulication LiTH-ISY-I-0260, Department of EE, Linköping University, Sweden.

[17] ——, "A result on delayless information tranmission," in *International Sympnosium on Information Theory*, 1979, Grignano, Italy.

[18] H. S. Witsenhausen, "Informational aspects of stochastic control," in *Proceedings of the Oxford Conference on Stochastic Optimization*, 1978.

[19] D. Teneketzis, "Communication in Decentralized Control," Ph.D. Thesis, Department of EECS, MIT, Cambridge, MA., Sep. 1979.

[20] G. Munson, "Causal Information Transmission with Feedback," Ph.D. Thesis, Department of Electrical Engineering, Cornell University, Ithaca, NY, 1981.

[21] A. Gorbunov and P. Pinsker, "Non-anticipatory and prognostic epsilon entropies and message generation rates," *Problems in Information Transmission*, vol. 9, pp. 1840191, 1973.

[22] ——, "Prognostic epsilon entropy of a Gaussian message and a Gaussian source," *Problems in Information Transmission*, vol. 10, pp. 184-191, 1974.

[23] P. Pinsker and A. Gorbunov, “Epsilon entropy with delay with small mean-square reproduction error,” *Problems in Information Transmission*, vol. 23, no. 2, pp. 3-8, April-June 1987.

[24] S. Engell, “New results on the real-time transmission problem,” *IEEE Trans. Inf. Theory*, vol. 33, no. 2, pp. 210–218, Mar. 1987.

[25] Y.-C. Ho, M. Kastner, and E. Wong, “Teams, signaling, and information theory,” *IEEE Trans. Autom. Control*, vol. 23, no. 2, pp. 305–312, 1978.

[26] S. Matloub and T. Weissman, “Universal zero-delay joint source-channel coding,” *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5240-5250, Dec. 2006.

[27] A. Drake (1962). Observation of a Markov process through a noisy channel. Sc.D. Thesis, Department of EE, MIT, Cambridge, MA.

[28] J. Devore, “A note on the observation of a Markov source through a noisy channel,” *IEEE Trans. Inf. Theory*, vol. 20, no. 6, pp. 762–764, 1974.

[29] H. S. Witsenhausen, “On the structure of real-time source coders,” *Bell System Technical Journal*, vol. 58, no. 6, pp. 1437-1451, July-August 1979.

[30] V. S. Borkar, S. K. Mitter, and S. Tatikonda, “Optimal sequential vector quantization of Markov sources,” *SIAM Journal of Optimal Control*, vol. 40, no. 1, pp. 135–148, Jan. 2001.

[31] J. C. Walrand and P. Varaiya, “Optimal causal coding—decoding problems,” *IEEE Trans. Inf. Theory*, vol. 29, no. 6, pp. 814-820, Nov. 1983.

[32] R. Lipster and A. Shiryayev, *Statistics of Random Processes, Vol. II:Applications*, Springer-Verlag, 1977.

[33] T. Basar and R. Bansal, “Optimal design of measurement channels and control policies for linear-quadratic stochastic systems,” *European Journal of Operational Research*, vol. 93, pp. 226–236, 1994.

[34] ——, “Simultaneous design of measurement channels and control stategies for stochastic systems with feedback,” *Automatica*, vol. 25, pp. 679-694, 1989.

[35] D. Teneketzis, “On the structure of optimal real-time encoders and decoders in noisy communication,” *IEEE Trans. Inf. Theory*, pp. 4017-4035, Sep. 2006.

[36] A. Mahajan and D. Teneketzis, “On-time diagnosis in discrete event systems,” *submitted to the 9th International Workshop on discrete event systems*, Jan. 2008.

[37] A. Segall, “Stochastic processes in estimation theory,” *IEEE Trans. Inf. Theory*, pp. 275–286, May 1976.

[38] H. S. Witsenhausen, “A standard form for sequential stochastic control,” *Mathematical Systems Theory*, vol. 7, no. 1, pp. 5-11, 1973.

[39] J. Marschak and R. Radner, *Economic Theory of Teams*, Yale University Press, New Haven, 1972.

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

[41] H. S. Witsenhausen, “Separation of estimation and control for discrete time systems,” *Proc. IEEE*, vol. 59, no. 11, pp. 1557–1566, Nov. 1971.

[42] ——, “Some remarks on the concept of state,” in *Directions in Large-Scale Systems*, ed. Y. C. Ho and S. K. Mitter, 1976, Plenum, pp. 69-75.

[43] R. J. Aumann, “Agreeing to disagree,” *Annals of Statistics*, pp. 1236-39, 1976.

[44] E. B. Dynkin and A. A. Yushkevich, *Controlled Markov Processes*, ser. A Series of Comprehensive Studies in Mathematics, Springer-Verlag, 1975.

[45] P. Whittle, *Optimization Over Time*, vol. 2, ser. Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, 1983.

[46] A. Arapostathis, V. S. Borkar, E. Fernandez-Gaucherand, M. K. Ghosh, and S. Marcus, “Discrete-time controlled Markov processes with average cost criterion - a survey,” *SIAM Journal of Control and Optimization*, vol. 31, no. 2, pp. 282–344, Mar. 1993.

[47] L. I. Sennott, *Stochastic dynamic programming and the control of queueing systems*, Wiley-Interscience, New York, NY, USA, 1999.

[48] ——, "The computation of average optimal policies in denumerable state Markov decision chains," *Advances in Applied Probability*, vol. 29, pp. 114-137, 1997.

[49] ——, "On computing average cost optimal policies with application to routing parallel queues," *ZOR Mathematical Methods in Operations Research*, vol. 45, pp. 45-62, 1997.

[50] Q. Zhang and S. Kassam, "Finite-state Markov model for Rayleigh fading channels," *IEEE Transactions on Communications*, vol. 47, no. 11, 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, no. 3, pp. 363–378, 1972.

[52] D. S. Bernstein, S. Zilberstein, and N. Immerman, "The complexity of decentralized control of markov decision processes," in *Proceedings of the 16th International Conference on Uncertainty in Artificial Intelligence (UAI)*, 2000, Stanford, CA, pp. 32-27.

[53] W. S. Lovejoy, "A survey of partially observable Markov decision processes," *Annals of Operation Research*, 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," *Operations Research*, vol. 37, no. 5, pp. 791–797, 1989.

[56] C. C. White III, "Partially observed Markov decision processes: a survey," *Annals of Operation Research*, vol. 32, 1991.

[57] W. Zhang, "Algorithms for partially observed Markov decision processes," Ph.D. Thesis, Hong Kong University of Science and Technology, 2001.

[58] J. Rust, "Using randomization to break the curse of dimensionality," *Econometrica*, vol. 65, no. 3, pp. 487–516, 1997.

[59] ——, "Dealing with the complexity of economic calculations," in *Fundamental Limits to Knowledge in Economics*, 1996.