THE INTERPLAY BETWEEN INFORMATION AND CONTROL THEORY WITHIN INTERACTIVE DECISION-MAKING PROBLEMS

BY

SIVA KUMAR GORANTLA

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Adjunct Assistant Professor Todd Coleman, Chair
Professor Tamer Basar
Professor Douglas Jones
Assistant Professor Negar Kiyavash
Adjunct Professor Sean Meyn

ABSTRACT

The context for this work is two-agent team decision systems. An *agent* is an intelligent entity that can measure some aspect of its environment, process information and possibly influence the environment through its action. In a colloborative two-agent team decision system, the agents can be coupled by noisy or noiseless interactions and cooperate to solve problems that are beyond the individual capabilities or knowledge of each agent.

This thesis focuses on using *stochastic control* and *information theoretic* tools hand-in-hand in solving and analyzing an interactive two-agent sequential decision-making problem. *Stochastic control* techniques can help in identifying optimal strategies for sequential decision making based on observations. *Information-theoretic* tools address the fundamental limit of performance between two agents with noisy interaction - in the context of communication and rate-distortion. The motivation for this work comes from the quest for using stochastic control tools in identifying optimal policies for a two-agent team decision system with an objective of *maximizing the information rate*. The resulting policies, if they exist, will involve decision making at each step based on observations, in contrast to existing communication schems that decide what to transmit over a long time-horizon, at the start of communication. However, there are many questions that have to be addressed:

- How should we formulate a stochastic-control problem to capture information gains?
- Suppose we can formulate such a control problem. Can we solve for explicit, non-random, optimal strategies that operate on sufficient statistics (thus resulting in a simple structure for optimal policies)?
- Further, do these control-theory based policies assure reliability of communication in an information-theoretic sense?

Consider a different problem where a third person has knowledge of the optimal policies of the two interacting agents, but is unaware of the cost function they are colloboratively optimizing.

- Can he deduce what the two agents are trying to achieve based on their policies?

In this thesis, we focus on addressing these questions using perspectives from both information and control theory. We consider an interacting two-agent decision-making problem consisting of a Markov source process, a causal encoder with feedback, and a causal decoder. We augment the standard formulation by considering general alphabets and a non-trivial cost function operating on current and previous symbols; this enables us to introduce the 'sequential information gain cost' function that can capture information gains accumulated at each time step. We emphasize how this problem formulation leads to a different style of coding scheme with a control-theoretic flavor. Further, we solve for structural results on these optimal policies using dynamic programming principles. We then demonstrate another interplay between information theory and control theory, at the level of reliability of message-point communication schemes, by establishing a relationship between *reliability* in feedback communication to the *stability* of the posterior belief's nonlinear filter.

We also consider the two-agent inverse optimal control (IOC) problem, where a fixed policy satisfying certain statistical conditions is shown to be optimal for some cost function, using probabilistic matching.

We provide examples of the applicability of this framework to communication with feedback, hidden Markov models and the nonlinear filter, decentralized control, brain-machine interfaces, and queuing theory.

To my brother and parents, for their love and support

ACKNOWLEDGMENTS

This thesis would not have been possible without the constant support from my adviser, Prof. Todd Coleman. I am grateful to him for giving the right direction to this work. He has been a great source of inspiration as a researcher and as a person.

Besides my adviser, I would like to thank my committee - Prof. Tamer Basar, Prof. Douglas Jones, Prof. Sean Meyn and Prof. Negar Kiyavash – for their insights and excellent advice.

Thanks are due to my friends and fellow graduate students at the Coordinated Science Laboratory - Sreekanth, Jay, Hemant, Nilesh, Ram, Sachin, Mavis, Xun, Ray, Sanggyun, Sreeram, and Chris to name a few - for the numerous discussions on information theory, communications, academic gossip and life in general. I am thankful to my friends Varun, Sandeep, Shilpa, Meghna, Shastry, Theja, Rohit and Anshuman for all those moments of fun in this small town of Champaign-Urbana, without which life would mostly be too dull and boring.

I would like to thank my Dad and Mom, to whom this thesis is dedicated, for their hard work and sacrifices that have gone into shaping my life. I would like to thank my brother for always supporting me in every issue.

Last, but not least, I would like to thank God for making every coin I tossed more than fair for me.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Many current and future societal problems involve designing and understanding networks of sequential decision-making entities cooperating in an uncertain environment. Some of these entities may be physical/biological agents, whereas others might be computerized systems. For example, cyber-physical systems feature interacting networks of physical processes that are noisily sensed and actuated by computational algorithms. The mammalian brain is another example, where the cooperative goals of sensing, perception, learning, and eliciting behavior are achieved via coupled neural systems that interact via signaling across a noisy biological medium.

From an engineering system designer vantage point, obtaining optimal coordination strategies for a network of interacting decision-makers is in general computationally intractable [2]. For a class of small networks (e.g. comprising a specific interaction structure between an encoder and a decoder), and an asymptotic performance objective, fundamental limits of performance can be addressed using the information theoretic concepts of communication and rate-distortion [3]. Identifying optimal strategies for sequential decision-making under uncertainty for a single agent, on the flipside, is traditionally addressed with control theoretic-principles of Markov decision theory [4].

From a scientific vantage point, the joint statistical dynamics between interacting decision-makers can provide insight into the cost or utility they are cooperatively optimizing. For small networks (e.g. an encoder and decoder) with a limited statistical dynamics interaction structure, this has been addressed with the information-theoretic principle of source-channel probabilistic matching [5]. Inverse optimal control theory theory identifies cost functions for which a fixed strategy of one decision-maker is optimal [6] and has been used in neural [7, 8] and cognitive science [9] applications.

It appears evident that understanding this class of problems for more general objectives and interaction structures can utilize insights from both information and control theory, but the differences in their philosophical starting points is striking, even for a two-agent problem.

1.1 Traditional Approaches in Information and Control Theory

Conventional theoretic approaches in information and control theory have striking differences in terms of their objectives, decision making and decision variables.

Control Theory: Markov decision theory problems typically involve observations of state variables' whose future statistics are impacted by their current values and the current 'decision variable' that is under causal control of a decision-maker. The alphabet size of observations and decision variables are typically *unrelated* to the time horizon n of the problem. Moreover, at each time step, a decision must be made based upon causal information up to that time. Lastly, the performance objective is to minimize an expected sum of costs, each of which operates on current state, observation, and decision variables. Structural results are typically desirable in such settings because they develop conditions relating the existence of explicit, non-random strategies that operate on sufficient statistics. Succinctly, we can state this as follows:

(a) *time horizon-independent alphabets*

(b) *decisions made sequentially based on causal information*

(c) *performance objective: sum of costs operating on current observations and decision variables*

Information Theory problems, traditionally specify large but *fixed* time horizon n for which some decisions are not made until this terminal point. Even in problems where neither an observation nor a decision variable lies in a time-horizon dependent set (e.g. reproducing a source over a noisy channel with a fidelity criterion), Shannon's 'separation theorem' [10] shows that for very large n, it is sufficient to first decompose the problem into sub-problems, each of which contains some observations or decision variables with time horizon-dependent alphabet structure (e.g. of size 2^{nR}) and a performance objective pertaining to constrained extremizing of mutual information while assuring reliability. As such, traditional information theoretic problem formulations have the following starting point:

(a) *time horizon-dependent alphabets*

(b) *some decisions made at final stage of long time horizons*

(c) *performance objective: extremize mutual information(information rate) and assure reliability*

So these two philosophies have striking differences. Consider the class of 'causal coding/decoding' problem that further demonstrates this:

1.1.1 The Causal Coding/Decoding Problem

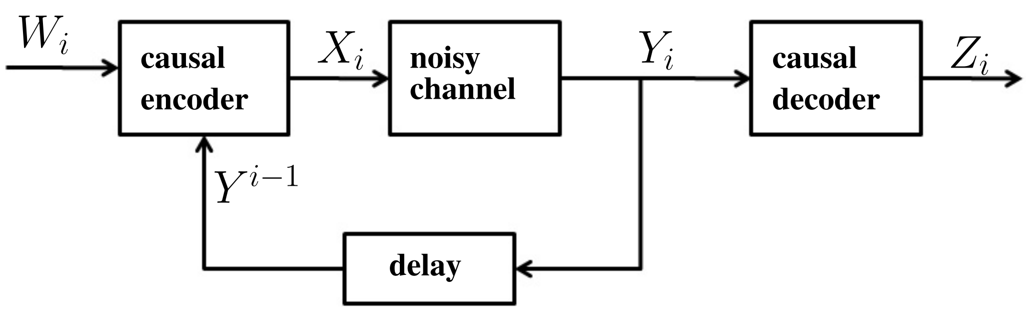

Figure 1.1: Basic problem setup: an optimal causal coding/decoding problem.

At each time step i, the causal encoder's decision variable is the input $X_i \in \mathsf{X}$ to a noisy channel that is a causal function of source inputs $(W_1, \ldots, W_i)$ and the noisy channel outputs $Y_1, \ldots, Y_{i-1}$: $X_i = e_i(W^i, Y^{i-1})$. The causal decoder's decision variable is a 'source estimate' $Z_i \in \mathsf{Z}$ that is a causal function of channel outputs $(Y_1, \ldots, Y_i)$: $Z_i = d_i(Y^i)$. They jointly design their strategies $\pi = (e, d)$ to minimize a function $J_{n,\pi}$ pertaining to an expected sum of costs:

$$J_{n,\pi} = \mathbb{E}_{e,d}\left[\sum_{i=1}^{n} g(W_i, Z_i)\right] \tag{1.1}$$

Some aspects of the problem appear to make it amenable to a control theoretic analysis: (a) the source alphabet W is unrelated to the time horizon n and (b) the sequential decision-making and additive costs , (c) the performance objective (1.1) operates additively on observations/decision variables in the vicinity of each time

i as compared to only at the final time horizon n. The presence of the noisy channel in the loop possibly makes it amenable to an information-theoretic analysis: mutual information could plausibly provide tight bounds on attainable costs. On the flipside, neither agent's observations at any time point are a nested version of the other's and so they have a 'non-classical' information structure [11] - making this a 'hard' control problem. Analogously, the 'hard' delay constraint pertaining to causal decoding and typical 'hard decision' assumption of W,Z being in discrete, time-horizon independent alphabets typically render information-theoretic techniques irrelevant to the understanding of these 'real-time' problems [12, 13].

1.2 Communication with Feedback - Motivation for Control-Theoretic Analysis

We now consider the traditional feedback communication model and how its assumptions - along with traditional 'real-time' problem assumptions - can be modified so that fundamental limits are unchanged but the frameworks align.

Feedback Information Theory: It is well known that [3], for a point-to-point communication system, the capacity does not increase with feedback. Hence the same encoding-decoding schemes that are used for communication without feedback can be used to extremize mutual information and assure reliability. Yet, feedback can be useful in (a) developing *sequential* encoding-decoding schemes, and (b) improving reliability performance (improving the rate at which error in decoding goes to zero - error exponents). The latter is not of interest in our analysis in this dissertation. The exciting part of feedback information theory is the possibility of time-horizon independent alphabets and sequential decision making while having a performance objective of a traditional information theory problem. Recently, a development by Shayevitz and Feder [14, 15, 16] has re-visited a philosophically different way to frame the feedback communication model - dating back to the 1960s [17, 18, 19] - that has a more dynamical systems and control theoretic flavor. Succintly, feedback information theoretic problems will have the following starting point:

(a) *possibility to build time horizon-independent alphabets*

(b) *decisions can made sequentially based on causal information*

(c) *performance objective: extremize mutual information(information rate) and assure reliability*

This leads us to ask the following questions:

- How should a control problem be posed to solve the equivalent of a feedback communication problem with the objective of maximizing mutual information?
  - What changes should be made to a traditional control theoretic setup to capture information gains?
  - What will be the structure of the optimal encoding-decoding policies?
- Does the control-theory based optimal encoding-decoding policies assure *reliability* in communication?

We will answer the first question in Chapter 3 where we introduce the concept of *sequential information gain cost*. 'Sequential information gain cost' captures information gains and can be computed sequentially in terms of current observations and decision variables; thus it has the flavors of both theoretic setups. We will also show two crucial generalizations that are required to be able to capture information gains: first, the decision variables should be beliefs rather than estimates which can take values over an alphabet, and second, the additive cost function should be generalized over (1.1) to include previous decision variable Z_{i-1}. We obtain a structural result demonstrating the existence of optimal coordination strategies operating on sufficient statistics, capturing traditional results [20] as a special case. This is used to obtain the structural results that can aid the design of optimal and 'user-friendly' coordination strategies for brain-machine interfaces [21]. As a consequence of using sequential information gain cost, we show that

- the posterior matching scheme (PM Scheme) [16] is an optimal coordination strategy for the information gain cost and source model $W_i = W_{i-1} = W_0$ and $\mathsf{W} = [0, 1]$.
- Under a particular constraint, the hidden Markov model and nonlinear filter [22] are an optimal coordination strategy for the sequential information gain cost with $\mathsf{W} = \mathsf{X}$ (ref Fig 1.1),

In Chapter 5, we further investigate the interplay between control and feedback-information theory, this time from the lens of reliability of message-point communication schemes. We establish necessary and sufficient conditions for when a message point feedback communication achieves reliable communication by providing an equivalence between non-linear filter stability and reliable feedback communications.

1.3 Inverse Optimal Control in Interactive Decision Making Problems

The joint statistical dynamics between interacting decision-makers can provide insight into the cost or utility they are cooperatively optimizing. For a two-agent problem with a limited statistical dynamics interaction structure, this has been addressed with the information-theoretic principle of source-channel probabilistic matching [5]. Inverse optimal control theory identifies cost functions for which a fixed strategy of one decision-maker is optimal [6].

Reconsider the causal coding-decoding problem in Fig 1.1. Though dynamic programming (DP) technique provides a general methodology to solve this team decision problem, this involves performing dynamic programming over the space of probability measures, which is a hard problem.

- While solving the team decision problem, is it possible to bypass the step involving dynamic programming over the space of probability beliefs?

In Chapter 4, we focus on an alternate approach - the 'inverse optimal control' approach, that can help in bypassing the dynamic programming step in certain cases. In this approach, we identify a fixed strategy of the agents and verify if it is optimal. Verification is done by identifying a set of "easy-to-describe" cost functions for which this fixed strategy is optimal using our inverse optimal control result. If the actual cost function falls in this set, then we know the policies we started with are in fact optimal, and there is no further need to perform the dynamic programming step. One downside in this approach is the guesswork involved in identifying the policies at the start.

In Section 2.2, we provide a short summary of earlier results in inverse optimal control literature. In Section 4.1, we provide our inverse optimal control result for a two-agent team decision system based on an information-theoretic approach. We identify a set of "easy-to-describe" cost functions - through the variational equations for rate-distortion and capacity-cost functions - for which a fixed policy is optimal. As a consequence of this result, we were able to make an interesting connection of inverse optimal control with time reversibility as discussed below.

Time reversibility plays an important role in disciplines concerning dynamical systems, e.g. in physics (conservation laws); statistical mechanics (in terms of equilibrium states); stochastic processes (e.g. queuing networks [23, 24] and convergence rates of Markov chains [25, ch 20]); and biology (e.g. trans paths in ion channels [26]). However, its use in providing information-theoretic fundamental limits appears to be somewhat limited.[1] The following observation drives us to investigate its impact on information-theoretic limits further:

In queuing systems, the celebrated Burke's theorem [24, 29] uses time reversibility to show that, in a certain stochastic dynamical system - an M/M/1 queue in steady-state - *the state of the system (queue) at time t is* ***independent*** *of all outputs (departures) before time t.* This observation has been used in proving achievability theorems for queuing timing channels [30], [31],[32]. In feedback information theory, statistical independence of one random variable at time t from others up to and including time t plays an important role - for example, tightness conditions in the converse to the channel coding with feedback.[2]

These observations lead us to investigate the role of time reversibility in characterizing information-theoretic fundamental limits of stochastic systems with dynamics via statistical independence. In Section 4.2, we extend our inverse optimal control result, to show that if a fixed coordination policy elicits reversibly feasible dynamics and a condition on time-reversibility, then it is a sufficient condition for the policy to be optimal. We then look at the following examples in Section 4.4, show that they are inverse-control optimal and deduce the cost functions for which the schemes are optimal.

- Gauss-Markov source, AGN channel pair - pertains to the decentralized control problems in [34, Ch 6],[35] with quadratic state cost and squared error distortion,

- Markov counting-function source, Z channel pair - pertains to the $\cdot/M/1$ queue for timing channels [30, 31],

[1] Although Mitter et al. [27, 28] have related Markov chain reversibility to entropy flow and equilibrium states in thermodynamic systems

[2] Which is part of the "posterior matching principle" for optimal communication with feedback over a DMC [16, Sec III],[33]

- Markov counting-function source, ‘inverted E’ channel pair - pertains to Blackwell’s trapdoor communication channel [36, 37, 38].

This is presented in the logical flow in Fig 1.2.

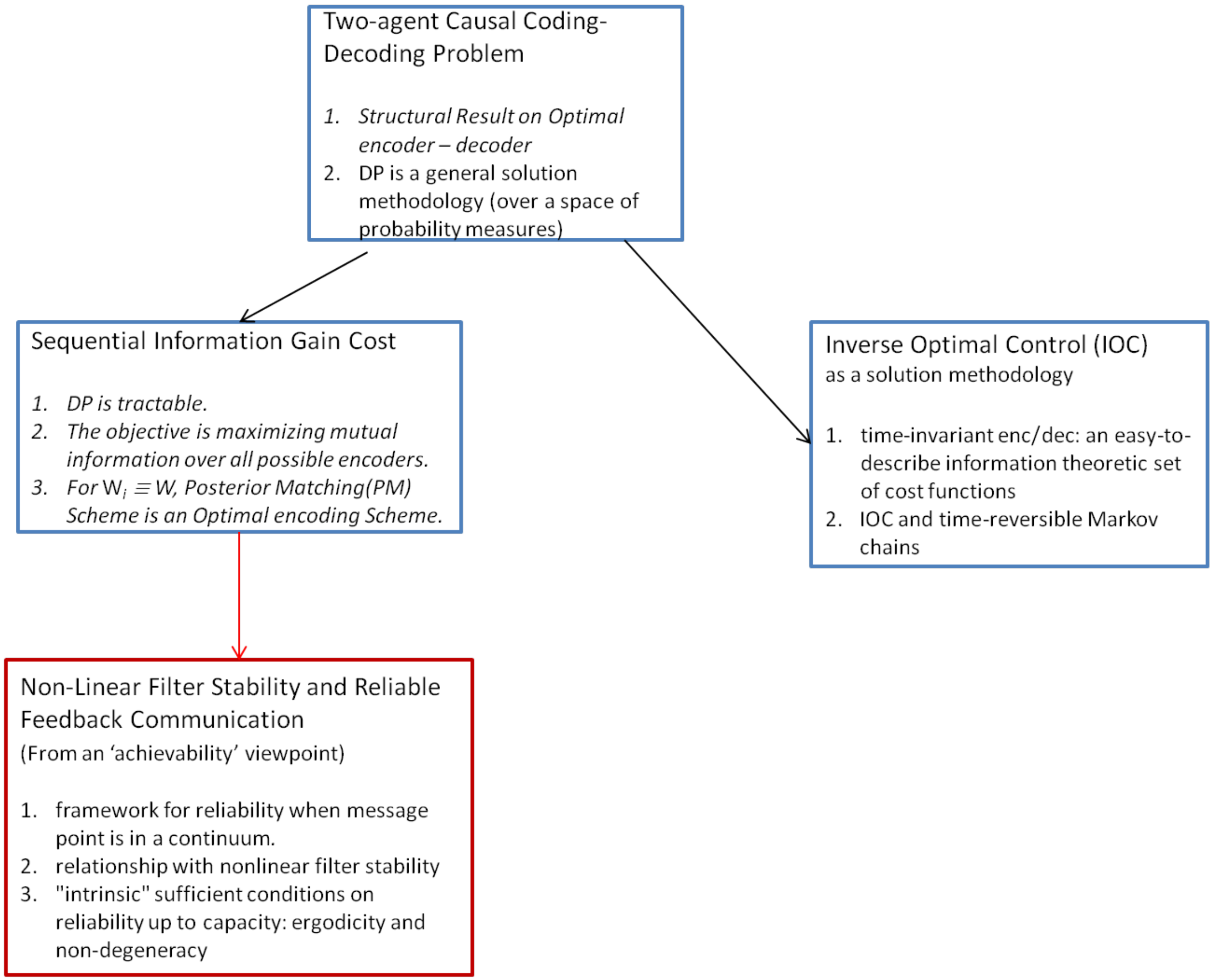

Figure 1.2: Logical flow of the dissertation.

CHAPTER 2

NOTATIONS, LITERATURE REVIEW AND PROBLEM SETUP

2.1 Definitions and Notations

Probabilistic Notation

- For a sequence $a_1, a_2, \ldots$, denote a_i^j as $(a_i, \ldots, a_j)$ and $a^j \triangleq a_1^j$.
- Denote the probability space with sample space Ω, sigma-algebra $\mathcal{F}$, and probability measure $\mathbb{P}$ as $(\Omega, \mathcal{F}, \mathbb{P})$.
- For a given $(\Omega, \mathcal{F})$ and a Borel space $(\mathsf{V}, \mathcal{B}(\mathsf{V}))$, denote any measurable function $X : \Omega \to \mathsf{V}$ as a random object. If $\mathsf{V} = \mathbb{R}$, then X is termed a random variable.
- Upper-case letters V represent random objects and lowercase letters $v \in \mathsf{V}$ represent their realizations.
- For any two sigma-algebras $\mathcal{A}$ and $\mathcal{B}$, define $\mathcal{A} \vee \mathcal{B}$ to be the smallest σ-algebra containing both.
- Denote $\sigma(Y)$ as the sigma-algebra generated by random variable Y and

  $$\mathcal{F}_{k,n}^{Y} \triangleq \bigvee_{m=k}^{n} \sigma(Y_m)$$

  as the sigma-algebra generated by $(Y_m : k \leq m \leq n)$.
- For two probability measures $\mathbb{P}$ and $\mathbb{Q}$ on $(\Omega, \mathcal{F})$, we say that $\mathbb{P}$ is absolutely continuous with respect to $\mathbb{Q}$ (denoted by $\mathbb{P} \ll \mathbb{Q}$) if $\mathbb{Q}(A) = 0$ implies $\mathbb{P}(A) = 0$ for all $A \in \mathcal{F}$. If $\mathbb{P} \ll \mathbb{Q}$, denote the Radon-Nikodym derivative

as any random variable $\frac{d\mathbb{P}}{d\mathbb{Q}} : \Omega \to \mathbb{R}$ that satisfies

$$\mathbb{P}(A) = \int_{\omega \in A} \frac{d\mathbb{P}}{d\mathbb{Q}}(\omega)\mathbb{Q}(d\omega), \quad A \in \mathcal{F}.$$

- Denote $\mathcal{P}(\mathsf{V})$ as the space of probability measures on $(\mathsf{V}, \mathcal{B}(\mathsf{V}))$. For any random object $V : \Omega \to \mathsf{V}$, denote

$$P_V(A) \triangleq \mathbb{P}(V \in A) \triangleq \mathbb{P}(\{\omega : V(\omega) \in A\}), \quad A \in \mathcal{B}(\mathsf{V}).$$

- Denote μ as the Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

- Denote the conditional probability distribution of one random object V given that another U takes on u as

$$P_{V|U=u}(A) \triangleq \mathbb{P}(V \in A|U = u), \quad A \in \mathcal{B}(\mathsf{V}).$$

- For a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ and a sigma-algebra $\mathcal{G} \subset \mathcal{F}$, denote $\mathbb{Q}_{|\mathcal{G}}$ to be the probability measure restricted to $\mathcal{G}$, i.e. the probability measure on $(\Omega, \mathcal{G})$ for which $\mathbb{Q}_{|\mathcal{G}}(A) = \mathbb{Q}(A)$ for all $A \in \mathcal{G}$.

- Denote $L_1(\mathbb{P})$ to be the set of all $\mathcal{F}$-measurable functions f for which $|f|$ is $\mathbb{P}$-integrable.

Markov Chains Notation:

- A random process $V = (V_i : i \geq 1)$ is a Markov chain if

$$P_{V_{i+1}|V^i=v^i}(A) = P_{V_{i+1}|V_i=v_i}(A), \quad A \in \mathcal{B}(\mathsf{V}). \tag{2.1}$$

  It is *time-homogenous* if $P_{V_{i+1}|V_i=v_i}(A) = Q(A|v_i)$.

- A Markov chain is *time-reversible* if the forward and reverse time processes are statistically indistinguishable:

$$(V_j : 1 \leq j \leq n) \overset{d}{=} (V_{n-j+1} : 1 \leq j \leq n) \tag{2.2}$$

  where $\overset{d}{=}$ denotes equivalence in distribution.

- A hidden Markov model $(\tilde{W}, Y) = (\tilde{W}_i, Y_i)_{i \geq 1}$ is a random process where:

– $\tilde{W}$ is a Markov chain

– Y satisfies, for all $A \in \mathcal{B}(\mathsf{Y})$:

$$\mathbb{P}\left(Y_n \in A | \mathcal{F}_{1,\infty}^{\tilde{W}} \vee \mathcal{F}_{1,n-1}^{Y}\right) = P_{Y|\tilde{W}}(A|\tilde{W}_n). \tag{2.3}$$

Information Theoretic Notation:

- Given two probability measures $P, Q \in \mathcal{P}(\mathsf{V})$, define the *Kullback-Leibler divergence* as

$$D(P\|Q) \equiv \begin{cases} \int_{\mathsf{V}} \log \frac{dP}{dQ}(v) P_V(dv), & \text{if } P \ll Q \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.4}$$

- Given two sets of conditional distributions $(P_{V|U=u}, P'_{V|U=u} \in \mathcal{P}(\mathsf{V}) : u \in \mathcal{U})$ and a distribution $P_U \in \mathcal{P}(\mathsf{U})$, define the *conditional divergence* as

$$D\left(P_{V|U} \| P'_{V|U} | P_U\right) \triangleq \int_{\mathcal{U}} D\left(P_{V|U=u} \| P'_{V|U=u}\right) P_U(du). \tag{2.5}$$

- Consider a set of conditional distributions $(P_{V|U=u}, \in \mathcal{P}(\mathsf{V}) : u \in \mathcal{U})$ and a distribution $P_U \in \mathcal{P}(\mathsf{U})$. This induces a marginal distribution $P_V \in \mathcal{P}(\mathsf{V})$. The mutual information is given by

$$I(P_{V|U}, P_U) \triangleq I(V;U) \triangleq D\left(P_{V|U} \| P_V | P_U\right). \tag{2.6}$$

  U and V are independent if and only if $I(V;U) = 0$.

- The conditional mutual information is given by

$$I(W;Y_2|Y_1) = D\left(P_{W|Y_1,Y_2} \| P_{W|Y_1} | P_{Y_1,Y_2}\right). \tag{2.7}$$

- The chain rule for mutual information is given by

$$\begin{aligned} I(W;Y^n) &= \sum_{i=1}^{n} I(W;Y_i|Y^{i-1}). \\ \Rightarrow I(W^n;Y^n) &= \sum_{i=1}^{n} I(W^n;Y_i|Y^{i-1}). \end{aligned} \tag{2.8}$$

- Consider a memoryless channel $P_{Y|X} = (Q_{Y|X=x} \in \mathcal{P}(\mathsf{Y}) : x \in \mathsf{X})$, a cost function $\eta : \mathsf{X} \to \mathbb{R}_+$, and an upper bound $L \in \mathbb{R}_+$. Define the capacity-cost function as $C\left(\eta, P_{Y|X}, L\right)$ [39] and its maximizing distribution $P_X^*(\eta, P_{Y|X}, L)$ as:

$$
\begin{aligned}
P_X^*(\eta, P_{Y|X}, L) &\triangleq \underset{P_X \in \mathcal{P}(\mathsf{X}) s.t. \mathbb{E}[\eta(X)] \leq L}{\arg\max} I(P_X, P_{Y|X}) && (2.9)\\
C\left(\eta, P_{Y|X}, L\right) &\triangleq I\left(P_X^*(\eta, P_{Y|X}, L), P_{Y|X}\right). && (2.10)
\end{aligned}
$$

2.2 Markov Decision Processes and Inverse Optimal Control

2.2.1 Markov Decision Process

Markov decision processes (MDPs), named after Andrey Markov, provide a mathematical framework for modeling decision-making in situations where outcomes are partly random and partly under the control of a decision maker. MDPs are useful for studying a wide range of optimization problems solved via dynamic programming and reinforcement learning. The description of a simple MDP problem is given in Figure 2.1 in terms of (state,control,cost).

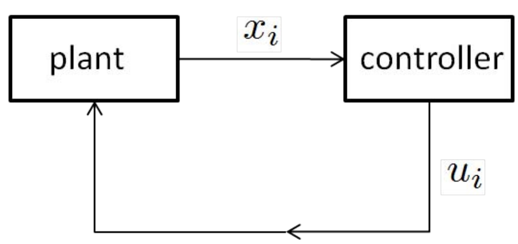

Figure 2.1: Simple MDP with plant and controller.

State Dynamics There are several aspects that determine how the dynamics of the state can be represented. Firstly, the state changes can occur in continuous time or discrete time. Secondly, the state dynamics can be completely deterministic or partly random, given the control of the decision maker.

$$\text{CT: } dx(t) = f(x(t), u(t)) + \tilde{f}(x(t), u(t))dw(t)$$
$$\text{DT: } x_{t+1} \sim p(\cdot|x_t = x, u_t = u)$$

where w_t is a brownian motion process. We define $p_{ij}(u)$, the transition law, as

$$p_{ij}(u) = p(x_{t+1} = j|x_t = i, u_t = u).$$

Control Law The controller acts as a decision maker but his actions are restricted to take some structure.

$$u_t = u(x_t, t).$$

Cost There is a cost associated for every action and the current state as

$$\begin{array}{ll} \text{DT:} & J(x_0) = \sum_{t=1}^{T} L(x_t, u_t) + D(x_T) \\ \text{CT:} & J(x_0) = \int_{t=0}^{T} L(x_t, u_t) + D(x_T) \end{array}$$

$L(x, u)$ is the cumulative cost added at each time.

A simple example with linear state dynamics and quadratic cost is given by

$$x_{t+1} = Ax_t + Bu_t + w_t$$
$$L(x, u) = \frac{1}{2}x^T Qx + \frac{1}{2}u^T Ru.$$

2.2.2 Solving a MDP- Forward vs. Inverse Problem

The Forward Problem Given the state dynamics, the cost function, the objective is to find a control law that can minimize the cost $J(x_0)$ where x_0 is the initial state of the system

The Inverse Problem Given the state dynamics, the control law, the objective is to find the family of cost functions for which the given control law is optimal.

Example: The LQG Case - Forward vs Inverse Problem:

Consider the deterministic system dynamics given by

$$\begin{aligned}\dot{x} &= Ax + Bu \\ u(t) &= -kx(t) \\ L(x,u) &= \frac{1}{2}x^T Q x + \frac{1}{2}u^T u \ (R = I \text{ W.L.O.G })\end{aligned}$$

Note that the optimal control law is stationary. This is possible only if the time-horizon is infinite ($T = \infty$).

Solution to the forward LQG problem is given by the following theorem (Given $f \equiv (A, B)$ and $L \equiv (Q, R)$), find $u \equiv k$)

Lemma 2.2.1. *There **always** exists a non-negative definite matrix P^* s.t.*

$$P^*A + A^T P^* - P^*BB^T P^* + Q^T Q = 0 \tag{2.11}$$

and the optimal control law is given by

$$k^* = P^*B$$

and the optimal cost is

$$J^*(x_0) = \frac{1}{2}x_0^T P^* x_0.$$

Remark 1. *Note that the equation* (2.11) *is the Ricatti equation for infinite horizon. This is an equivalent of HJB equation with linear dynamics and $T = \infty$.*

In general the state dynamics and the control law (f, u) are related to the cost function L through value function V for the optimal control law, in the Gaussian case $(f \equiv (A, B), u \equiv k)$ are related to $L \equiv Q$ through the matrix P^*. Kalman [40] gives necessary and sufficient conditions for optimality without involving P^* as follows, consequently providing one of the first results of inverse optimal control.

Lemma 2.2.2. *k is an optimal and stable control law if and only if*

$$|1 + k^T \Phi(i\omega)B|^2 > 1$$

where $\Phi(s) = (sI - A)^{-1}$. *And there* ***always*** *exist a* Q *s.t.*

$$|1 + k^T\Phi(i\omega)B|^2 = 1 + ||Q\Phi(i\omega)B||^2.$$

2.2.3 General Approach in Solving the Forward Problem

MDPs can be solved by dynamic programming. This involves finding the Value function $V(x,t)$ given by

$$V(x,t) = \min_u \{L(x,u) + V(f(x,t))\}.$$

The above equation is called the Bellman equation. The continuous-time version of it is called the Hamilton-Jacobi-Bellman equation given by

$$\dot{V}(x,t) - \min_u \{\nabla_x V(x,t) \cdot f + L\} = 0$$

with the boundary condition $V(x,T) = D(x_T)$.

The HJB equation is a PDE whose solution is the Bellman value function $V(x,t)$. The HJB equation provides sufficient conditions for an optimum, and this condition must be satisfied over the whole of the state space.

The optimum cost to the forward problem is given by

$$J^*(x_0) = V(x_0, 0).$$

2.2.4 General Approaches in Solving the Inverse Problem - Inverse Optimal Control

Given the state dynamics (f) and the control law (u), find all the cost functions (L) for which the control law (u) is optimal.

Using a Dynamic Programming point of view:

Casti [41] considered deterministic (fixed) dynamics and provided the following necessary and sufficient conditions on the cost function L for a given control law u to be optimal using a dynamic programming point of view.

Lemma 2.2.3. *[41, Theorem 2.1] All functions L which are optimal relative to a given f and u must satisfy the differential equation*

$$0 \;=\; \frac{d}{dT}\{p(x,u)\} + \nabla_x[L - p\cdot f] \tag{2.12}$$

where

$$p(x,u) = (A)^{\#}\,\nabla_u L + (I - A^{\#}A)y \; A = \left(\frac{df}{du}\right)^T.$$

Remark 2. *The theorem provides an equivalent representation of HJB equation when assuming u is an optimal control law. Note that (f,u) and L are related through the value function V, which is a solution of the HJB equation. Let us try to eliminate V to provide a direct relationship between (f,u) and L.*

Control Lyapunov Function Approach:

Deng and Kristic [42] show that for every system with a 'stochastic control Lyapunov function' it is possible to construct a controller which is optimal with respect to a meaningful cost functional.

Now suppose for a deterministic continuous time system

$$dx = f(x,u) = a(x)dt + b(x)udt. \tag{2.13}$$

Definition 2.2.4. *A function $V : \mathbb{R}^n \to \mathbb{R}_+$ is a control Lyapunov function (clf) if*

$$inf_u\{\nabla_a V + \nabla_b V u\} < 0, \forall x \neq 0.$$

This is equivalent to saying

$$\forall x \neq 0, \exists u \textit{ s.t. } \dot{V}(x,u) < 0.$$

Now suppose we find a function V which is a control Lyapunov function for the system in (2.13). Treat V as an optimal value function and this leads us to find the cost and the optimal control law for which this cost is optimal.

Lemma 2.2.5. *[42, Theorem 3.1] If V is a clf to the system in* (2.13)*, then the* *control law*

$$u^* = -\frac{\beta}{2} R^{-1} (\nabla_b V)^T \frac{(\gamma')^{-1}(|\nabla_b V R^{-1/2}|)}{|\nabla_b V R^{-1/2}|}, \; \beta \geq 2$$

where R is an arbitrary matrix, $\beta \geq 2$ can be arbitrary and γ is an arbitrary function. This control law u^* solves the problem of inverse optimal control by minimizing the cost functional

$$L(x,u) = l(x) + \beta^2 \gamma \left(\frac{2}{\beta} |R^{1/2} u| \right)$$

where

$$l(x) = 2\beta \left[\gamma^{-1}(|\nabla_b V R^{-1/2}|) - \nabla_a V \right].$$

Remark 3. *Apart from the guesswork involved in designing control-Lyapunov function, this method is easier than solving the forward problem (solving HJB equation) because once we know the value function from the HJB equation, finding an explicit formula for cost function is known.*

2.2.5 Knowledge of Actions vs Knowledge of Policies - Inverse Reinforcement Learning

The above control-theoretic approaches require knowledge of the exact policies to deduce what the cost function is. Many times it is only possible to have access to the actual actions and not the complete policies. Inverse reinforcement learning (IRL) methods in machine learning rely on data in the form of state transitions obtained from an expert performing a task. The data is used to i) infer the cost function the expert is trying to optimize, ii) build a controller which mimics the expert (imitation learning).

As seen in the control-Lyapunov function approach Sec 2.2.4, once we have a value function (V) - it is possible to determine the cost function (L) and optimizing control law (u^*) for MDPs. The class of linear MDPs (LMDP) introduced by Dvijotham and Todorov [43] also have this property. But here, instead of guessing the value function, they [43] estimate the value function using the data provided. Once the value function is known, the cost function and the optimizing control law can be determined explicitly. LMDP formulation becomes handy because the IRL algorithms can be implemented much faster than the usual MDPs.

Dynamics of a Linear MDP process:

The dynamics of a linear MDP process are explained below: The state has passive dynamics given by

$$x_{i+1} \sim p(\cdot|x).$$

The controller can impose different dynamics

$$x_{i+1} \sim \pi(\cdot|x).$$

The cost has two components - first depending on the state (x), second depending on the action ($\pi(\cdot|x)$):

$$L(x,u) = L(x,\pi(\cdot|x)) = q(x) + D\left(\pi(\cdot|x) \| p(\cdot|x)\right).$$

Determining the cost function and optimal control law using value function $V(x)$:

Define the desirability function $z(x) = \exp(-V(x))$ where $V(x)$ is the optimal value function. It can be shown that the optimal control law is

$$\pi^*(x'|x) = \frac{p(x'|x)z(x')}{\sum_x' p(x'|x)z(x')}. \tag{2.14}$$

And the normalized Bellman equation is

$$\lambda z(x) = \exp(-q(x)) \left(\sum_x^{\prime} p(x'|x)z(x') \right). \tag{2.15}$$

Hence, once we know the value function $V(x)$, we can determine $\pi^*(x'|x)$ and the cost function $q(x)$ from the above equations.

Determining the Value function $V(x)$ using data:

We are provided the dataset of state transitions $\{x_n, x'_n\}_{n=1,\cdots,N}$ under an optimal control

$$x'_n \sim \pi^*(\cdot|x_n).$$

Assume that we know the passive dynamics $p(\cdot|x)$ and with this information we have to infer the value function $v(x)$. The inference method used is maximum likelihood.

Think of π^* as being parameterized by the value function $V(x)$ (2.14). Then the negative log-likelihood is

$$LL[v(\cdot)] = -\sum_n V(x'_n) + \sum_n \log\left(\sum_{x'} p(x'_n|x)e^{-V(x'_n)}\right).$$

Thus inverse reinforcement learning for LMDP framework reduces to unconstrained convex optimization of an easily computable function.

Once $\hat{V}$ is estimated, we can compute $\pi^*(x'|x)$ and the cost function $q(x)$ using (2.14) and (2.15).

Another way of inferring Value function using weights of features The value function can be inferred by assuming $V(x)$ is a sum of weighted features $f_i(x)$ and estimating the weights:

$$V(x) = \sum_i w_i f_i(x)$$

where $f_i(x)$ are given features and w_i are unknown weights. This approach is used in earlier IRL algorithms (Abbeel and Ng, 2004).

2.3 The Stability of Conditional Markov Processes and Filter Stability of Hidden Markov Models

2.3.1 Hidden Markov Models and Filter Stability

A hidden Markov model $(S, Y) = (S_n, Y_n)_{n\geq 1}$ is a pair of random sequences where the signal component S_n takes values in the space S and the observation component Y_n takes values in the space Y. See Figure 2.2.

Definition 2.3.1. *A HMM is defined as follows:*

(a) S is a Markov chain. Let ξ and ν denote the transition probability and the

prior on S such that for any $A \in \mathcal{B}(\mathsf{S})$

$$\xi(S_{n-1}, A) \triangleq \mathbb{P}\left(S_n \in A | \mathcal{F}^S_{1,n-1}\right) = \mathbb{P}\left(S_n \in A | \sigma(S_{n-1})\right) \qquad \mathbb{P}-a.s., \tag{2.16a}$$

$$\nu(A) \triangleq \mathbb{P}\left(S_0 \in A\right). \tag{2.16b}$$

(b) Y satisfies, for all $A \in \mathcal{B}(\mathsf{Y})$:

$$\mathbb{P}\left(Y_n \in A | \mathcal{F}^S_{1,\infty} \vee \mathcal{F}^Y_{1,n-1}\right) = P_{Y|S}(A|S_n). \tag{2.16c}$$

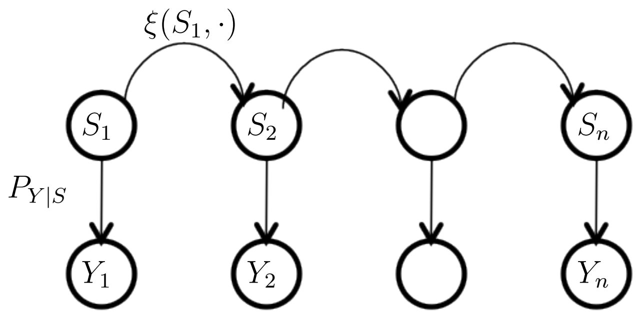

Figure 2.2: A hidden Markov model $(S_n, Y_n)_{n\geq 1}$ such that $(S_n)_{n\geq 1}$ is a Markov chain with transition probability ξ and $(Y_n)_{n\geq 1}$ is the observation process generated according to $P_{Y|S}$.

Let $\pi_n(\cdot) = \mathbb{P}\left(S_n \in \cdot | \mathcal{F}^Y_{1,n}\right)$ denote the posterior distribution on S_n after seeing observations $Y_1, \cdots, Y_n$ subject to $\pi_0 = \nu$. A regular version of π_n satisfies the update equation

$$\pi_n(\cdot) = \mathbb{P}\left(S_n \in \cdot | \mathcal{F}^Y_{1,n}\right) \tag{2.17}$$

$$\pi_n(du) = \frac{P_{Y|S}(Y_n|u) \int_s \xi(s, dx)\pi_{n-1}(ds)}{\int_x P_{Y|S}(Y_n|x) \int_s \xi(s, dx)\pi_{n-1}(ds)}. \tag{2.18}$$

The posterior update equation in (2.18) shows that π_n is a functional of π_{n-1} and Y_n; as such, a succinct characterization is the dynamical system $\pi_n(\cdot) = \Lambda(\pi_{n-1}, Y_n)(\cdot)$ under the initial condition $\pi_0 = \nu$.

Construction of measure $\mathbf{P}$ corresponding to HMM:

Let us work on the canonical path space $\Omega = \Omega^S \times \Omega^Y$, where $\Omega^S = \mathsf{S}^{\mathbb{Z}}$ and $\Omega^Y = \mathsf{Y}^{\mathbb{Z}}$. Denote by $\mathcal{F}$, the $\sigma-$algebra generated by Ω. Denote $\mathcal{F}_{k,n} = \mathcal{F}^S_{k,n} \vee \mathcal{F}^Y_{k,n}$. In order to construct the measure $\mathbf{P}$ corresponding to the definition 2.3.1 of HMM, we need the following ingredients:

- The probability kernel $\xi : \mathsf{S} \times \mathcal{B}(\mathsf{S}) \to [0,1]$ as defined in (2.16a).
- A probability measure ϖ on $(\mathsf{S}, \mathcal{B}(\mathsf{S}))$ such that

$$\int \xi(z, A)\varpi(dz) = \pi(A) \qquad \forall A \in \mathcal{B}(\mathsf{U}).$$

- The probability kernel $P_{Y|X} : \mathsf{S} \times \mathcal{B}(\mathsf{Y}) \to [0,1]$ as defined in (2.16c).

The HMM generative model is represented by

- The probability measure $\mathbf{P} : \mathcal{F} \to [0,1]$ and for every $n \in \mathbb{N}$, $\mathbf{P}|_{\mathcal{F}_{-n,n}} = \mathbf{P}^{(n)}$ where $\mathbf{P}^{(n)}$ is a probability measure on $\mathcal{F}_{-n,n}$ such that

$$\begin{aligned} \mathbf{P}^{(n)}(A) &= \int 1_{\{(s,y)\in A\}} P_{Y|S}(dy(n)|s(n)) \cdots P_{Y|S}(dy(-n)|s(-n)) \\ &\quad \times \xi(s(n-1), ds(n)) \cdots \xi(s(-n), ds(-n+1))\varpi(ds(-n)). \end{aligned}$$

- In addition to the probability measure $\mathbf{P}$, introduce a probability kernel $\mathbf{P}^{\cdot} : \mathsf{S} \times \mathcal{F}_{0,\infty} \to [0,1]$ such that $\mathbf{P}^z$ is the law of $(S_n, Y_n)_{n\geq 0}$ started at $S_0 = z$. For any probability measure ν on $(\mathsf{S}, \mathcal{B}(\mathsf{S})$, we define the probability measure

$$\mathbf{P}^{\nu}(A) = \int 1_{\{(s,y)\in A\}} \mathbf{P}^z(ds, dy)\nu(dz) \qquad \forall A \in \mathcal{F}_{0,\infty}.$$

Nonlinear Filter Stability of HMM:

The nonlinear filter for HMM (2.18) is termed 'stable' if the posterior belief is insensitive to initial conditions, i.e. for any $\nu \ll \overline{\nu}$

$$\lim_{n\to\infty} \mathbb{E}||\pi_n - \bar{\pi}_n|| = 0 \tag{2.19}$$

where π_n and $\bar{\pi}_n$ are posteriors subject to initial conditions $\pi_0 = \nu$ and $\bar{\pi}_0 = \overline{\nu}$.

The typical question of stability is under which conditions (in terms of ξ, $P_{Y|S}$) the filter is stable and satisfies (2.19). There have been different approaches to establish the conditions for the filter stability problem ([44, 45, 46, 47, 48, 49, 50, 51]). Van Handel et al. recently developed 'intrinsic methods' that provide necessary and sufficient conditions on filter stability [52]. Lemma 2.3.2 provides the necessary condition for filter stability to hold [52]:

Lemma 2.3.2. *[52, eq 1.10] For two probability measures $\nu \ll \overline{\nu}$, the filter is stable* (2.19) *if and only if*

$$\overline{\mathbb{E}}\left(\frac{d\nu}{d\overline{\nu}}(S_0)\big|\bigcap_{n\geq 0}\mathcal{F}^Y_{0,\infty}\vee\mathcal{F}^S_{n,\infty}\right) = \overline{\mathbb{E}}\left(\frac{d\nu}{d\overline{\nu}}(S_0)\big|\mathcal{F}^Y_{0,\infty}\right). \tag{2.20}$$

This condition involves a complicated interaction between the dynamics of the latent signal (ξ) and the structure of the channel likelihood ($P_{Y|S}$). One sufficient condition for stability to hold, that we will later show is also crucial for reliability in communication systems, is

Lemma 2.3.3. *[1, Thm 4.2] Suppose that process S is ergodic, and the observation process Y is generated according to a non-degenerate channel law. Then*

$$\bigcap_{n\geq 0}\mathcal{F}^Y_{0,\infty}\vee\mathcal{F}^S_{n,\infty} = \mathcal{F}^Y_{0,\infty} \qquad \overline{\mathbb{P}} - a.s. \tag{2.21}$$

This condition (2.21) serves as the "glue" between filter stability and reliability in communication.

A key element in the proof of sufficiency (lemma 2.3.3) is showing that the HMM model fits in the framework of a conditional Markov process. $\{S_n\}$ can be interpreted as a non-homegenous Markov process when conditioned upon the entire observation record $\{Y_n\}_{n\geq 0}$. And the ergodicity of the (unconditional) signal process $\{S_n\}$ along with non-degeneracy of the channel $P_{Y|S}$ is equivalent to the *weak ergodicity* of the conditional signal process $\{S_n\}$. Finally weak ergodicity of a conditional Markov process implies that (2.21) (and thus stability) condition holds. An important take away from Van Handel's [1] result is showing that if certain assumptions on a generative model (e.g., HMM in the above case) are equivalent to the weak-ergodicity of the corresponding conditional Markov process, then (2.21) holds for the generative model. (This condition (2.21), as we will see later, is also a condition for stability for other generative models of interest.)

With this, we look at the formal definition and some properties of conditional Markov processes.

2.3.2 Conditional Markov Processes

Consider the pair $(S, Y) = (S_n, Y_n)_{n \in \mathbb{Z}}$ where S_n takes values in the space S and Y_n takes values in the space Y. We realize these processes on the canonical path space $\Omega = \Omega^S \times \Omega^Y$ with $\Omega^S = \mathsf{S}^{\mathbb{Z}}$ and $\Omega^Y = \mathsf{Y}^{\mathbb{Z}}$. Denote by $\mathcal{F}, \mathcal{F}^S, \mathcal{F}^Y$ by the Borel σ-field on $\Omega, \Omega^S, \Omega^Y$ respectively. We now introduce a measure $\mathbf{P}$ on $(\Omega, \mathcal{F})$ which defines a conditional Markov process. To this end, consider the probability kernel of the form $P^S : \mathsf{S} \times \Omega^Y \times \mathcal{B}(\mathsf{S}) \to [0, 1]$. Define a stationary probability measure **P** such that the following holds a.s. for every $n \in \mathcal{Z}$:

$$\mathbf{P}(S_{n+1} \in A | \mathcal{F}^S_{-\infty,n} \vee \mathcal{F}^Y) = P^S(S_n, Y \circ \Theta^n, A).$$

Thus, S_n is interpreted as a Markov chain in a random environment: the environment is the entire sequence Y, and S_n is a nonhomogenous Markov process, for almost every path Y, under the regular conditional probability $\mathbf{P}(\cdot | \mathcal{F}^Y)$.

<u>Construction of measure $\mathbf{P}$ corresponding to Conditional Markov Processes</u>:

In order to construct $\mathbf{P}$, we need the following three ingredients:

- The probability kernel $P^S : \mathsf{S} \times \Omega^Y \times \mathcal{B}(\mathsf{S}) \to [0, 1]$.
- A probability kernel $\mu : \Omega^Y \times \mathcal{B}(\mathsf{S}) \to [0, 1]$ such that

$$\int P^S(z, y, A)\mu(y, dz) = \mu(\Theta y, A), \qquad \text{for all } y \in \Omega^Y, A \in \mathcal{B}(\mathsf{S}).$$

- A probability measure $\mathbf{P}^Y$ on $(\Omega^Y, \mathcal{F}^Y)$ which is invariant under the shift, that is, $\mathbf{P}^Y(Y \in A) = \mathbf{P}^Y(Y \circ \Theta \in A)$ for all $A \in \mathcal{F}^Y$.

The generative model is represented by

- Define a probability kernel $\mathbf{P}_\cdot : \Omega^Y \times \mathcal{F}^S_{-\infty,\infty} \to [0, 1]$ and $\mathbf{P}_y|_{\mathcal{F}^S_{-n,n}} = \mathbf{P}^{(n)}_y$ where

$$\mathbf{P}^{(n)}_y(A) = \int 1_{\{u \in A\}} P^S(u(n-1), \Theta^n y, du(n)) \cdots$$
$$\times P^S(u(-n), \Theta^{-n} y, du(-n+1))\mu(\Theta^{-n} y, du(-n)).$$

- The probability measure $\mathbf{P}$ on $(\Omega, \mathcal{F})$ by setting

$$\mathbf{P}(A) = \int 1_{\{(x,y)\in A\}} \mathbf{P}_y(dx)\mathbf{P}^Y(dy).$$

Define a process S_n which starts at $S_0 = z$ and has the probability kernel $\mathbf{P}_{\cdot,\cdot} : \mathsf{S} \times \Omega^Y \times \mathcal{F}^S_{0,\infty} \to [0,1]$ by setting for $A \in \mathcal{F}^S_{0,n}$ as

$$\begin{aligned}\mathbf{P}_{z,y}(A) = \int 1_{\{u\in A\}} & P^S(u(n-1), \Theta^n y, du(n)) \cdots \\ & \times P^S(u(1), \Theta y, du(2)) P^S(u(0), y, du(1)) \delta_z(du(0))\end{aligned}$$

where $\delta_z(A) = 1_{\{z\in A\}}$.

Lemma 2.3.4. *[[1] Lemma 4.1] Suppose that the conditional Markov Process satisfies weakly ergodicity*

$$||\mathbf{P}_{z,y}(S_n \in \cdot) - \mathbf{P}_{z',y}(S_n \in \cdot)||_{TV} \xrightarrow{n\to\infty} 0 \, for \; (\mu \otimes \mu)\mathbf{P}^Y - a.e.(z, z', y).$$

Then the following holds true:

$$\bigcap_{n\geq 0} \mathcal{F}^Y_{-\infty,\infty} \vee \mathcal{F}^S_{n,\infty} = \mathcal{F}^Y_{-\infty,\infty} \qquad \mathbf{P} - a.s. \tag{2.22}$$

Remark 4. *Note that this result Lemma 2.3.4 is not in itself of use in providing asympotic properties of HMM* (2.21)*, as the entire observation field $\mathcal{F}^Y_{-\infty,\infty}$ appears in the expression rather than $\mathcal{F}^Y_{0,\infty}$. But Van Handel [1] shows that ergodicity of the unconditional Markov process $\{S_n\}$ is sufficient to show that* (2.22) *implies* (2.21)*. To summarize, for the case of HMM, if the unconditional signal process $\{S_n\}$ is ergodic and the observations are generated by a non-degenerate law, then* (2.21) *holds (as seen in Lemma 2.3.3).*

2.3.3 Relation between Hidden Markov Models and Conditional Markov Processes

We now demonstrate a relationship between HMMs and conditional Markov processess. For every hidden Markov model defined by $(\xi_S, P_{Y|S})$, we can construct a corresponding conditional Markov process with $(P^S, \mu, \mathbf{P}^Y)$ given according to Lemma 2.3.5.

Lemma 2.3.5. *[1, Lemma 3.3] There exist probability kernels $P^S : \mathsf{S} \times \Omega^Y \times \mathcal{B}(\mathsf{S}) \to [0,1]$ and $\mu : \Omega^Y \times \mathcal{B}(\mathsf{S}) \to [0,1]$, and a probability measure $\mathbf{P}^Y$ on $(\Omega^Y, \mathcal{F}^Y)$, such that the conditions of controlled Markov processes (Section 2.3.2) are satisfied and the measure $\mathbf{P}$ constructed there coincides with the measure $\mathbf{P}$ of the HMM generative model. In particular,*

$$P^S(S_n, Y \circ \Theta^n, A) = P(S_{n+1} \in A | \mathcal{F}^S \vee \mathcal{F}^Y)\mathbf{P} - a.s., \tag{2.23a}$$
$$\mu(Y \circ \Theta^n, A) = P(S_n \in A | \mathcal{F}^Y)\mathbf{P} - a.s. \tag{2.23b}$$

for every $A \in \mathcal{B}(\mathsf{S})$ and $n \in \mathbb{Z}$, and $\mathbf{P}^Y = \mathbf{P}|_{\mathcal{F}^Y}$.

2.4 Message Point Communication Schemes - Applications and Reliability

2.4.1 A Message Point Communication System

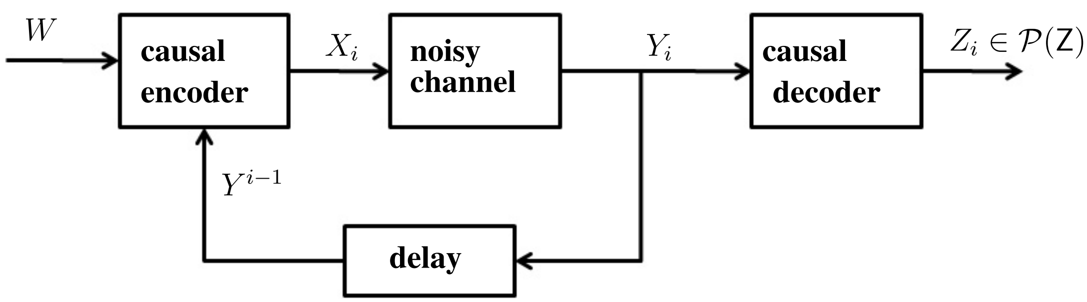

Figure 2.3: A message point communications system with feebdack: The message point W is communication over a memoryless channel in the presence of causal feedback. The causal decoder computes (and updates) the posterior belief on the message W by looking at the observation sequence $\{Y_i\}$. The objective is to maximize mutual information $I(W; Y^n)$ under specific constraints.

We consider communication of a message point $W \in \mathsf{W}$ over a memoryless channel with causal feedback. The message space W can be an arbitrary compact, uncountable subset of $\mathbb{R}^d$ with the following properties:

- W should be uncountable, so that an increasing finer set of quantizers of the form $(Q_n : \mathsf{W} \to \{1, 2, \cdots, 2^{nR} : n \geq 1, R \geq 0\}$ can be described, and

- W should be compact, so that every open cover (pertaining to quantization intervals) has a finite subcover (in particular at time n, there are 2^{nR} of them).

Applicability:

The 'continuous message point' problem formulation for reliable communication is pleasing for applications beyond traditional digital communications, such as biological communication, network control [53], and brain-machine interfaces [21]. The uncertain message is fundamentally in a continuum, because

- the alphabet of the message, W, is fixed irrespective of the number n of uses of the channel or any notion of rate of communication.
- A 'rate' R being achievable is defined in terms of whether or not the decoder's posterior belief converges quickly enough to a point mass.

Problem setup:

See Figure 2.3.

- The channel input $X_i \in \mathsf{X}$ is passed through a non-anticipative, memoryless channel to produce $Y_i \in \mathsf{Y}$:

$$\mathbb{P}\left(Y_n \in A | \mathcal{F}^X_{1,\infty} \vee \mathcal{F}^Y_{1,n-1} \vee \sigma(W)\right) = P_{Y|X}(A|X_n). \tag{2.24}$$

  We say that the channel is *non-degenerate* if for any distribution P_X, and any $x \in \mathsf{X}$, $y \in \mathsf{Y}$,

$$\frac{dP_{Y|X=x}}{dP_Y}(y) > 0. \tag{2.25}$$

- The encoder policy $(e_n : n \geq 1)$ specifies the next channel input based on the message and feedback

$$X_n = e_n(W, Y^{n-1}) \equiv \tilde{e}_n(W)$$

  where $\tilde{e} \equiv \tilde{e}(Y^{n-1})$ is a random object. One specific encoding policy is the posterior matching(PM) scheme which will be discussed in detail in Sec 2.4.3.

- Given an encoder policy e, the decoder computes its posterior belief on W:

$$\pi_n(\cdot) = \mathbb{P}\left(W \in \cdot | \mathcal{F}_{1,n}^Y\right). \tag{2.26}$$

$\pi_n(\cdot)$ can be updated as it satisfies the following recursive equation, the *nonlinear filter* [22]:

$$\pi_n(dw) = \frac{dP_{Y|X}\left(\cdot|\tilde{e}_n(\pi_{n-1}, w)\right)}{dP_\Lambda(\cdot|\pi_{n-1}, \tilde{e}_n)}(Y_n)\pi_{n-1}(dw) \tag{2.27}$$

$$P_\Lambda(dy|b, \tilde{e}) \triangleq \int_{\mathsf{W}} P_{Y|X}(dy|\tilde{e}(w', b))b(dw') \tag{2.28}$$

subject to $\pi_0 = \nu$, the initial belief about W in the absence of observations. Note that (2.27) is a manifestation of Bayes' rule: the numerator is the *likelihood*, the denominator (2.28) is a *normalization constant*, and the coefficient π_{n-1} is the *prior*. We denote $\bar{\pi}_n$ as the posterior satisfying (2.27) with initial condition ν replaced with $\overline{\nu}$. Thus:

$$\mathbb{P}\left(W \in A|\mathcal{F}_{1,n}^Y\right) \;=\; \int_{w\in A} \pi_n(dw) \tag{2.29a}$$

$$\overline{\mathbb{P}}\left(W \in A|\mathcal{F}_{1,n}^Y\right) \;=\; \int_{w\in A} \bar{\pi}_n(dw). \tag{2.29b}$$

- Objective: Given a cost constraint L, maximize the information rate

$$\max_{e_n : n\geq 1} \quad I(W;Y^n) \tag{2.30a}$$

$$\text{s.t.} \quad \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\eta(X_i)\right] \leq L. \tag{2.30b}$$

2.4.2 Reliable Communication for Message Point Communication Schemes

Our notion of reliable communication is somewhat non-traditional because the message point W lies in a compact set W. We now provide a formalism of achievability that is equivalent to the standard one [3].

Definition 2.4.1. *For any $k \geq 1$, partition W into 2^k equally spaced intervals pertaining to the uniform quantizer $Q_k : \mathsf{W} \to \left\{1, \ldots, 2^k\right\}$. Denote $\mathcal{G}_k \triangleq \sigma(Q_k(W))$ as the information about W given by the quantizer output. Denote $\overline{\nu} \in \mathcal{P}(\mathsf{W})$ as* *the uniform distribution on* W *and* $\nu^{k,W} \in \mathcal{P}(\mathsf{W})$ *as the* **random measure** *that is uniformly distributed over* ***one*** *of the* 2^k *partitions which contains* W*:*

$$\frac{d\overline{\nu}}{d\mu}(w) = 1 \tag{2.31}$$

$$\frac{d\nu^{k,W}}{d\mu}(l) = \begin{cases} 2^k, & Q_k(l) = Q_k(W) \\ 0, & \textit{otherwise} \end{cases}. \tag{2.32}$$

See Figure 2.4 for an example of $\overline{\nu}$ and $\nu^{k,W}$ when the message is a point on the [0,1] line, i.e., $\mathsf{W} = [0, 1]$.

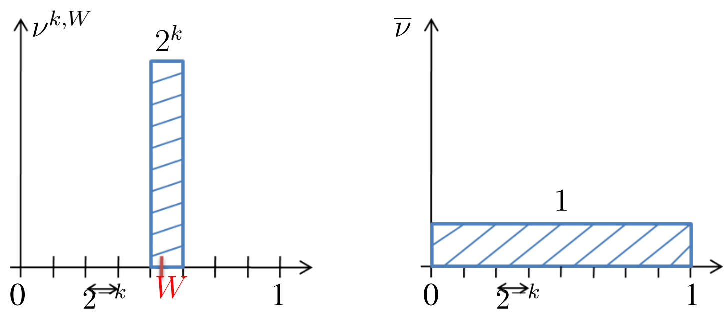

Figure 2.4: Two priors $\nu^{k,W}$ and $\overline{\nu}$. $\nu^{k,W}$ is a *random* measure: it is uniformly distributed on the interval of length 2^{-k} containing W. $\overline{\nu}$ is uniformly distributed on $\mathsf{W} = [0, 1]$.

Denote $\pi_n^{k,W}$ as the posterior measure satisfying the nonlinear filter equations (2.27) initialized with $\pi_0 = \nu^{k,W}$ and $\bar{\pi}_n$ as the analogous posterior initialized with $\bar{\pi}_0 = \overline{\nu}$.

Note that for any $0 < k < \infty$, $\nu^{k,W} \ll \overline{\nu}$. With this, we can recover the traditional notions of achievability:

Definition 2.4.2. *(*RELIABILITY*) An encoder* e *is reliable if for any* k*,*

$$\bar{\pi}_n\left(\{l : Q_k(l) = Q_k(W)\}\right) \xrightarrow{\bar{\mathbb{P}}} 1 \tag{2.33}$$

and it achieves rate $R > 0$ *if*

$$\bar{\pi}_n\left(\{l : Q_{nR}(l) = Q_{nR}(W)\}\right) \xrightarrow{\bar{\mathbb{P}}} 1. \tag{2.34}$$

Note that reliability means that any fixed number of bits can be decoded in the limit of large block length. Definition 2.4.2 is equivalent to the classical notion of rate: after n channel uses, one of an exponentially large number of hypotheses must be successfully distinguished from. Note that with prior $\nu = \nu^{k,W}$, for any n, it follows from (2.32) that $\pi_n^{k,W}(\{l : Q_k(l) = Q_k(W)\}) = 1$. Thus we have:

Lemma 2.4.3. *An encoder e is reliable iff for any k,*

$$D\left(\pi^{k,W}_{n|\mathcal{G}_k} \| \bar{\pi}_{n|\mathcal{G}_k}\right) \xrightarrow{\overline{\mathbb{P}}} 0. \tag{2.35}$$

Equivalently, An encoder e is reliable iff for any k,

$$\lim_{n\to\infty} \mathbb{E}||\pi^{k,W}_{n|\mathcal{G}_k} - \bar{\pi}_{n|\mathcal{G}_k}||_{TV} = 0 \qquad \overline{\mathbb{P}} - a.s. \tag{2.36}$$

Rate $R > 0$ is achievable if and only if

$$D\left(\pi^{nR,W}_{n|\mathcal{G}_{nR}} \| \bar{\pi}_{n|\mathcal{G}_{nR}}\right) \xrightarrow{\overline{\mathbb{P}}} 0. \tag{2.37}$$

Proof. See Figure 2.5 for the case $\mathsf{W} = [0, 1]$. $\pi^{nR,W}_{n|\mathcal{G}_{nR}}$ and $\bar{\pi}_{n|\mathcal{G}_{nR}}$ are equivalent to probability mass functions $p_n(j)$ and $\bar{p}_n(j)$ s.t.

$$p^W_n(j) = \pi^{nR,W}_n(\{l : Q_{nR}(l) = j\}) \quad 1 \le j \le 2^{nR} \tag{2.38}$$

and likewise for $\bar{p}_n(j)$, replacing $\pi^{nR,W}_n$ with $\bar{\pi}_n$. Thus

$$\begin{aligned} D\left(\pi^{nR,W}_{n|\mathcal{G}_{nR}} \| \bar{\pi}_{n|\mathcal{G}_{nR}}\right) &= \sum_{j=1}^{2^{nR}} p^W_n(j) \log \frac{p^W_n(j)}{\bar{p}_n(j)} \\ &= -\log \bar{\pi}_n(\{l : Q_{nR}(l) = Q_{nR}(W)\}) \end{aligned} \tag{2.39}$$

where (2.42) holds because $p^W_n(Q_{nR}(W)) = 1$. This is sufficient to prove (2.35) and (2.37).

To prove (2.36), define p^W_n and $\bar{p}^W_n$ as

$$p^W_n(j) = \pi^{k,W}_n(\{l : Q_k(l) = j\}) \quad 1 \le j \le 2^k \tag{2.40}$$

$$\bar{p}^W_n(j) = \bar{\pi}^{k,W}_n(\{l : Q_k(l) = j\}) \quad 1 \le j \le 2^k. \tag{2.41}$$

Hence,

$$
\begin{aligned}
\mathbb{E}||\pi^{k,W}_{n|\mathcal{G}_k} - \bar{\pi}_{n|\mathcal{G}_k}||_{\mathrm{TV}} &= \sum_{j=1}^{2^k} p^W_n(j)|p^W_n(j) - \bar{p}_n(j)| \\
&= \left|\bar{\pi}_n\left(\{l : Q_k(l) = Q_k(W)\}\right) - 1\right| \qquad (2.42)
\end{aligned}
$$

thus satisfying (2.33). $\square$

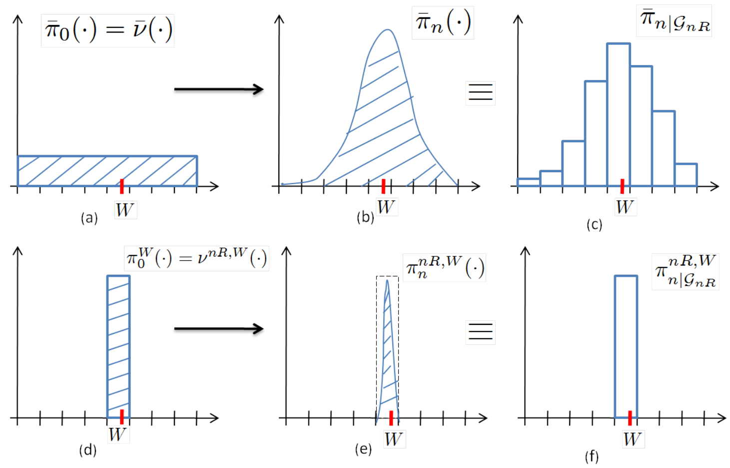

Figure 2.5: (a) represents prior $\overline{\nu}$, (b) represents the posterior $\bar{\pi}_n$ after observations y^n, and (c) represents $\bar{\pi}_n$ restricted to $\mathcal{G}_{nR}$. (d) represents prior $\nu^{nR,W}$, (e) represents the posterior $\pi_n^{nR,W}$ after the same y^n, and (f) represents $\pi_n^{nR,W}$ restricted to $\mathcal{G}_{nR}$. (c) and (f) are shown as PMFs over $\{1, 2, \cdots, 2^{nR}\}$. Rate $R > 0$ is achieveable iff the KL-divergence between (f) and (c) converges to 0 in $\overline{\mathbb{P}}$.

2.4.3 Posterior Matching Scheme - An Optimal Message-Point Communication Scheme with Feedback

In this section, we will introduce a simple yet optimal feedback-based encoding scheme and discuss its properties. We start by discussing the properties that any feedback encoding scheme should hold for it to be optimal, by looking at the converse to feedback communication problem. We then introduce the posterior matching-style encoding scheme and prove its optimality.

Motivation:

The converse to the point-to-point communication problem tells us about the properties for an encoding scheme $e(W, Y^n)$ to be optimal.

$$
\begin{aligned}
I(W;Y^n) &= H(Y^n) - H(Y^n|W) \\
&= \sum_{i=1}^{n} H(Y_i|Y^{i-1}) - H(Y_i|Y^{i-1}, W) && \text{(2.43a)} \\
&= \sum_{i=1}^{n} H(Y_i|Y^{i-1}) - H(Y_i|X_i, Y^{i-1}, W) && \text{(2.43b)} \\
&= \sum_{i=1}^{n} H(Y_i|Y^{i-1}) - H(Y_i|X_i) && \text{(2.43c)} \\
&\overset{(a)}{\leq} \sum_{i=1}^{n} H(Y_i) - H(Y_i|X_i) && \text{(2.43d)} \\
&= \sum_{i=1}^{n} I(X_i;Y_i) \overset{(b)}{\leq} nC && \text{(2.43e)}
\end{aligned}
$$

where (2.43b) follows from the structure of encoding policy $X_i = e(W, Y^{i-1})$, and (2.43c) is because $(W, Y^{i-1}) - X_i - Y_i$ forms a Markov chain. The equation holds with equality only if

(a) Y_i is independent of Y^{i-1}.

(b) X_i is drawn according to a capacity-achieving distribution P_X^*.

Hence any encoding scheme of the form $\{e_n(W, Y^n) : n \geq 1\}$ should satisfy conditions (a) and (b) to be optimal (maximize mutual information). We will now introduce the following posterior matching-style feedback based encoding scheme which satisfies the above conditions and is thus optimal. In addition, the PM scheme is also desirable for implementation because of the following reasons:

- There is no forward error correction - it simply adapts on the y and sequentially hands the decoder what is missing.
- The scheme admits a simple time-invariant dynamical system structure.

These properties of the PM scheme have made it amenable to implementation in real-world systems coupling computers with physical/biological systems that practically achieve fundamental limits [53, 21].

We now define the posterior matching-style coding scheme.

Definition 2.4.4. *Given a message point distribution P_W, a map $\phi : \mathsf{W} \to \mathsf{X}$, and a noisy channel $P_{Y|X}$, we say that the set of mappings $\{T_y : \mathsf{W} \to \mathsf{W}\}_{y \in \mathsf{Y}}$ is PM-compatible if and only if the encoder scheme with time-invariant dynamics given by*

$$\tilde{W}_0 = W \in \mathsf{W}, \qquad \tilde{W}_{i+1} = T_y(\tilde{W}_i) \tag{2.44a}$$

$$X_{i+1} = F_X^{-1}(\tilde{W}_{i+1}) \tag{2.44b}$$

satisfies the following properties (and hence satisfies conditions (a) and (b) above):

Property 2.4.5. *(a) $\tilde{W}_i$ is independent of Y^{i-1}, i.e., $\tilde{W}_i \perp\!\!\!\perp Y^{i-1}, \forall i \geq 1$.*

(b) $\tilde{W}_i \sim P_W^*, \forall i \geq 1$.Further the mapping ϕ is such that $\tilde{W}_i \sim P_W^* \Rightarrow X_i \sim P_X$.

(c) The mappings $\{T_y : \mathsf{W} \to \mathsf{W}\}_{y \in \mathsf{Y}}$ and $\phi : \mathsf{W} \to \mathsf{X}$ are invertible.

Define $I \triangleq I(X, Y)$ s.t. $X \sim P_X^*$, and $P_{Y|X}$ is the channel law given by (2.51).

Example 1. *When* $\mathsf{W} = [0, 1]$*, the following encoding scheme* (2.45) *is a specific instance of* (2.44) *and satisfies all the properties of Property 2.4.5. This scheme is termed the Posterior Matching scheme and is first introduced by Shayevitz and Feder in [16]:*

$$\tilde{W}_0 = W \sim \mathcal{U}([0, 1]), \qquad \tilde{W}_{i+1} = T_y(\tilde{W}_i) = F_{\tilde{W}_i|Y_i=y}(\tilde{W}_i) \tag{2.45a}$$

$$X_{i+1} = F_X^{-1}(\tilde{W}_{i+1}) \tag{2.45b}$$

where $\mathcal{U}([0, 1])$ corresponds to the uniform distribution on $[0, 1]$ line and $F_{\tilde{W}_i|Y_i}$ is the CDF of the distribution on $\tilde{W}_i$ conditioned upon observation Y_i.

<u>Posterior Matching and its connection to Arithmetic Coding:</u>

Arithmetic coding is a form of variable-length entropy encoding used in lossless data compression. When a string is converted to arithmetic encoding, frequently used characters will be stored with fewer bits and not-so-frequently occurring characters will be stored with more bits, resulting in fewer bits used in total. Arithmetic coding differs from other forms of entropy encoding such as Huffman coding in that rather than separating the input into component symbols and replacing each with a code, arithmetic coding encodes the entire message into a single number, a fraction W where $0.0 \leq W \leq 1.0$.

The implementation of the encoding step in arithmetic coding is equivalent to the implementation of the posterior matching encoding scheme for a ***noiseless*** communication problem. To understand this, first let us look at an example implementation of arithmetic coding:

Consider the process for encoding a message with a four-symbol model and let the message that needs to be transmitted be $W = 0.538$. The encoding contains 3 steps. Fix $\tilde{W}_0 = W = 0.538$. $X_i \in \{$'$\mathrm{NEUTRAL}'$, '$\mathrm{POSITIVE}'$, '$\mathrm{NEGATIVE}'$, '$\mathrm{END\text{-}OF\text{-}DATA}'\}$.

- **Step 1:** Divide the [0,1] line into subparts depending on the optimal probabilities the model assigned to each symbol. For example, the system might ideally like 60% of its symbols to be NEUTRAL, 10% of its symbols to be POSITIVE, 10% of its symbols to be NEGATIVE and 10% of its symbols to indicate END-OF-DATA. These probabilities can even be adaptive, and hence re-evaluated and provided by the system at each step.

- **Step 2:** Determine the sub-interval in which $\tilde{W}_i$ falls. Consequently determine the next symbol X_i to be transmitted.

- **Step 3:** Rescale the sub-interval in which $\tilde{W}_i$ to be $[0, 1]$-interval. Compute rescaled $\tilde{W}_{i+1}$ accordingly.

See Figure 2.6.

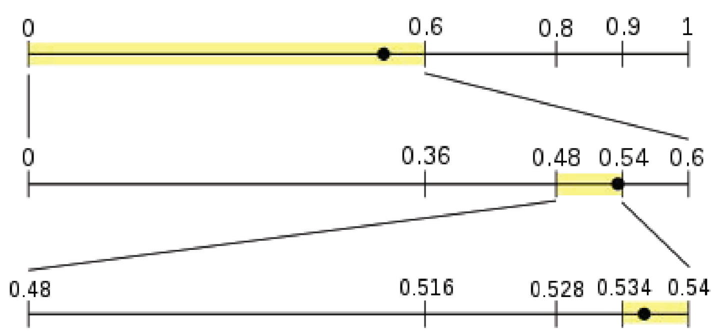

Figure 2.6: A diagram showing encoding of 0.538 (the circular point) in the example model. The region is divided into subregions proportional to symbol frequencies; then the subregion containing the point is successively subdivided in the same way.

When all symbols have been encoded, the resulting interval unambiguously identifies the sequence of symbols that produced it. Anyone who has the same final interval and model that is being used can reconstruct the symbol sequence that must have entered the encoder to result in that final interval.

In fact, the encoding process of arithmetic coding corresponds to the posterior matching scheme when the channel $P_{Y|X}$ is noiseless. Let $Y_i = X_i$ and the arithmetic coding can be rewritten as:

$$\tilde{W}_0 = W \sim \mathcal{U}([0,1]) \tag{2.46a}$$

$$\tilde{W}_{i+1} = T_y(\tilde{W}_i) = F_{\tilde{W}_i|Y_i=y}(\tilde{W}_i) \tag{2.46b}$$

$$X_{i+1} = F_X^{-1}(\tilde{W}_{i+1}) \tag{2.46c}$$

where F_X corresponds to the CDF of the optimal probabilities of the symbols determined in Step-1. Equation (2.46c) is equivalent to determining the sub-interval where $\tilde{W}_{i+1}$ lies in Step-2. Equation (2.46b) is equivalent to re-scaling the subinterval as described in Step-3 and computing $\tilde{W}_{i+1}$ from $\tilde{W}_i$ given the sub-interval $Y_i = X_i = y$.

The posterior matching scheme can be considered as an extension of arithmetic coding and works even when the decoder sees a noisy version of the symbols generated.

2.5 Problem Formulation

Throughout this discussion, we consider 4 random processes W, X, Y, Z associated with Borel metric spaces $\mathsf{W}, \mathsf{X}, \mathsf{Y}, \mathsf{Z}$ that are coupled according to Figure 1.1. The natural time ordering for the causal construction of the four random objects through time is given by:

$$\ldots, Z_{i-1}, \underbrace{W_i, X_i, Y_i, Z_i}_{i\text{th epoch}}, W_{i+1}, \ldots$$

The input process:

W is a time-homogenous Markov process such that for any $A \in \mathcal{F}_{\mathsf{W}}$:

$$P_{W_{i+1}|W^i=w^i,X^i=x^i,Y^i=y^i}(A) = P_{W_{i+1}|W_i=w_i}(A) \tag{2.47}$$

$$\equiv Q_W(A|w_i). \tag{2.48}$$

The causal encoder:

The *causal encoder* at time i has causal information about the source, W^i, and causal feedback about the channel outputs, Y^{i-1}, to specify the next channel input, X_i,

$$x_i = e_i(w^i, y^{i-1}). \tag{2.49}$$

We define the aspect of $e_i \in \mathsf{E}_i$ that maps W_i to X_i as $\tilde{e}_i \in \tilde{\mathsf{E}}$ where $\tilde{\mathsf{E}}$ is a space of Borel-measurable functions $f : \mathsf{W} \to \mathsf{X}$:

$$\tilde{e}_i(w^{i-1}, y^{i-1})(\cdot) = e_i\left(\begin{bmatrix} \cdot \\ w^{i-1} \end{bmatrix}, y^{i-1}\right) \equiv \tilde{e}_i(\cdot) \tag{2.50}$$

and we define E_i to be the space of Borel-measurable functions $f : \mathsf{W}^i \times \mathsf{Y}^{i-1} \to \mathsf{X}$ such that $\tilde{e}_i \in \tilde{\mathsf{E}}$ for all w^{i-1} and y^{i-1}.

The memoryless non-anticipative channel:

$X_i \in \mathsf{X}$ is passed through a time-homogenous, non-anticipative, memoryless channel to produce $Y_i \in \mathsf{Y}$; for any $A \in \mathcal{F}_{\mathsf{Y}}$:

$$P_{Y_i|Y^{i-1}=y^{i-1}, X^n=x^n, W^n=w^n}(A) = P_{Y|X}(A|x_i). \tag{2.51}$$

The causal decoder:

Lastly, the causal decoder at time i uses causal channel outputs, Y^i to specify $Z_i \in \mathsf{Z}$. Define D_i as a space of Borel-measurable functions $f : \mathsf{Y}^i \to \mathsf{Z}$ and $\mathsf{D} = \mathsf{D}_1 \times \ldots \times \mathsf{D}_n$. Then the causal decoder $d \in \mathsf{D}$ is a sequence of functions $d = (d_i : \ \ 1 \leq i \leq n)$:

$$z_i = d_i(y^i). \tag{2.52}$$

Belief update:

In the above discussion on the causal decoder, we deliberately consider Z to be general, not necessarily equal to W. Indeed, as we shall see, in some cases we set $\mathsf{Z} = \mathcal{P}(\mathsf{W})$ so that the outputs of the causal decoder represent *beliefs* about the source at time i. Define the beliefs $B_{i|j} \in \mathcal{P}(\mathsf{W})$ about the source at time i given the decoder's observations up until time $j \leq i$ as, for any $A \in \mathcal{B}(\mathsf{W})$:

$$\begin{aligned} B_{i|j}(A) &\triangleq \mathbb{P}\left(W_i \in A|Y^j\right) &\quad (2.53a)\\ b_{i|j}(A) &\triangleq P_{W_i|Y^j=y^j}(A). &\quad (2.53b) \end{aligned}$$

The beliefs can be interpreted as state variables that can be updated sequentially given new observations. The *nonlinear filter* $\Lambda : \mathcal{P}(\mathsf{W}) \times \mathsf{Y} \times \tilde{\mathsf{E}} \to \mathcal{P}(\mathsf{W})$ and *one step prediction update* $\Phi : \mathcal{P}(\mathsf{W}) \to \mathcal{P}(\mathsf{W})$ rules are given by [22]:

$$\begin{aligned} \Lambda(b, y, \tilde{e})(dw) &= \frac{dP_{Y|X}\left(\cdot|\tilde{e}(w)\right)}{dP_{\Lambda}(\cdot|b,\tilde{e})}(y) \times \Phi(b)(dw) &\quad (2.54)\\ \Phi(b)(dw) &\triangleq \int_{w' \in \mathsf{W}} Q_W\left(dw|w'\right) b(dw') &\quad (2.55)\\ P_{\Lambda}(dy|b,\tilde{e}) &\triangleq \int_{w' \in \mathsf{W}} P_{Y|X}(dy|\tilde{e}(w'))\Phi(b)(dw'). &\quad (2.56) \end{aligned}$$

Equation (2.54) can be interpreted as a standard manifestation of Bayes' rule: the numerator is simply the *likelihood*, the denominator is a *normalization constant*, and the coefficient $\Phi(b)$ is simply the *prior*. The aforementioned two equations specify how the beliefs are sequentially updated:

Lemma 2.5.1 ([20],[22])**.** *For any i and encoder policy e_i with associated $\tilde{e}_i$ given by* (2.50)*, the following holds:*

$$\begin{aligned} b_{i|i-1} &= \Phi\left(b_{i-1|i-1}\right) &\quad (2.57a)\\ b_{i|i} &= \Lambda\left(b_{i-1|i-1}, y_i, \tilde{e}_i\right). &\quad (2.57b) \end{aligned}$$

In Section 3.1, we demonstrate using a structural result how the beliefs arise as sufficient statistics in our main problem. In Section 3.2, we demonstrate how they additionally serve as optimal decision variables with information gain cost (3.1).

Additive cost function:

Denote a coordination strategy, also termed a *policy*, as $\gamma = (e_1, \ldots, e_n, d_1, \ldots, d_n)$ and the set of all feasible policies as $\Gamma = \{\gamma : e_i \in \mathsf{E}_i,\ d_i \in \mathsf{D}_i\}$. The causal encoder and decoder e and d are cooperating to achieve a common goal. The performance of their cooperation is measured in terms of an expected sum of costs over time horizon n with the following structure:

$$J_{n,\gamma}^{\alpha} = \mathbb{E}_{\gamma}\left[\sum_{i=1}^{n} \rho(W_i, Z_{i-1}, Z_i) + \alpha\eta(X_i)\right]. \quad (2.58)$$

The above expectation is taken with respect to an initial distribution P_{W_0,Z_0} where Z_0 is assumed known to the encoder and decoder. We assume that the functions ρ and η along with constant α have the following structure:

- $\rho : \mathsf{W} \times \mathsf{Z} \times \mathsf{Z} \to \mathbb{R}_+$ is a 'distortion-like' source cost, that relates the distortion between the source at time i and the outputs in the vicinity of time i.
- $\eta : \mathsf{X} \to \mathbb{R}_+$ is a 'power-like' channel input cost that penalizes channel inputs that deviate significantly from nominal desired values.
- $\alpha \in \mathbb{R}^+$ balances the relative importance of the two costs.

Definition 2.5.2. *We say that a sequential encoder-decoder pair $\gamma^* \in \Gamma$ is (globally) optimal if*

$$J^{\alpha}_{n,\gamma^*} \leq J^{\alpha}_{n,\gamma} \quad \textit{for all } \gamma \in \Gamma. \tag{2.59}$$

PART I

INFORMATION THEORETIC VIEWPOINTS ON OPTIMAL CAUSAL CODING-DECODING PROBLEMS

CHAPTER 3

OPTIMAL POLICY DESIGN IN AN INTERACTIVE DECISION MAKING PROBLEM

In this chapter we consider a causal coding/decoding problem where W is a Markov source process. We consider additive cost functions operating of the form $g(w_i, x_i, z_{i-1}, z_i)$. We do not impose assumptions (e.g. finiteness) on alphabets of the variables. Our motivation for this more general framework is an example (Section 3.0.1) motivated by feedback communication where the source alphabet is continuous, the decoder alphabet lies in a space of *beliefs* on the source alphabet, and the additive cost function is a log likelihood ratio pertaining to *sequential information gain*. Using dynamic programming, we provide a structural result whereby an optimal scheme exists that operates on appropriate sufficient statistics.

3.0.1 Example: Communication over a Noisy Channel with Feedback and the Sequential Information Gain Cost

We now consider the traditional feedback communication model and how its assumptions - along with traditional 'real-time' problem assumptions - can be modified so that fundamental limits are unchanged but the frameworks align. Consider the traditional information-theoretic communication model with feedback, consisting of an encoder, a decoder, and a fixed block length n. The encoder has a message $W \in \mathsf{W} = \{1, \ldots, 2^{nR}\}$. It specifies n inputs to the channel, $X_1, \ldots, X_n$. The channel is memoryless and non-anticipative where $P_{Y|X}(y|x)$ is the statistics of the output given the input. At each time step i, the encoder selects the message W and the previous channel outputs $Y_1, \ldots, Y_{i-1}$ at time i, to specify the next channel input X_i. The decoder, at time n, having acquired channel outputs $Y_1, \ldots, Y_n$, specifies a single decision, $\hat{W}_n \in \mathsf{W}$. The question asked in information theory is, how large can R be such that for sufficiently large n, there exist encoders and decoders for which $\mathbb{P}(\hat{W}_n \neq W) \to 0$? To demonstrate the existence of such encoders and decoders, a *random coding* argument [3] and the laws of large numbers are typically invoked.

Recently, a development by Shayevitz and Feder [14, 15, 16], has re-visited a philosophically different way to frame the feedback communication model - dating back to the 1960s [17, 18, 19]- that has a more dynamical systems and control theoretic flavor. Consider the following changes to the standard information theo-

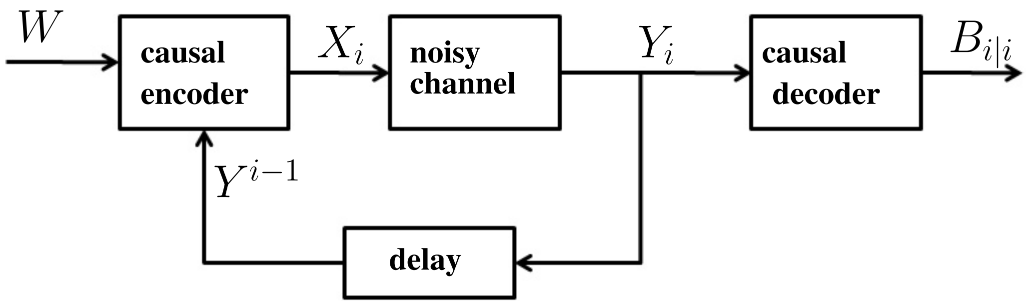

Figure 3.1: Communication of a message point W with causal feedback over a memoryless channel.

retic formulation that more closely resembles a causal coding/decoding problem, shown in Figure 3.1:

- *Message Point:* $W_i = W$ is equally likely over interval $\mathsf{W} = [0, 1]$.
- *Decoder:* At each i (not only at time n), the decoder specifies $Z_i = B_{i|i}$, the posterior belief about W given $Y_1, \ldots, Y_i$: $B_{i|i}(A) \triangleq \mathbb{P}\left(W_i \in A | Y^i\right)$.
- *Achievability:* As shown in Figure 3.2, with a set of uniform quantizers $\left(q_{iR} : [0, 1] \to 2^{iR},\ i \geq 1\right)$, a rate R is achievable if $B_{i|i}\left(\{w : q_{iR}(w) = q_{iR}(W)\}\right) \to 1$.

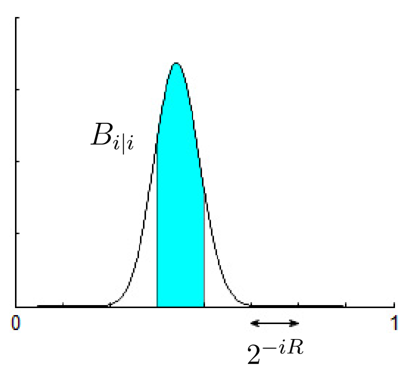

Figure 3.2: Representation of the posterior belief B_i in terms of its density.

Note the importance of W being a *continuous interval* and Z being the space of beliefs on W, $\mathcal{P}(\mathsf{W})$, in order for this 'real-time' flavored problem to relate to traditional information-theoretic notions of achievability. The fundamental limits under both formulations are equivalent [16], where achieving capacity subject to channel input cost $\eta(x)$ constraints pertains to maximizing the mutual information $I(W;Y^n)$ [10]. A time-invariant 'posterior matching' encoding scheme in Figure 3.1's framework achieves capacity on general memoryless channels [16]. Moreover, it is an optimal solution to a stochastic control problem [33] whose cost function at each time step is related to the *sequential information gain* $I(W;Y_i|Y^{i-1})$:

$$I(W;Y^n) = \sum_{i=1}^{n} I(W;Y_i|Y^{i-1}) = \sum_{i=1}^{n} \mathbb{E}\left[\log \frac{dB_{i|i}}{dB_{i-1|i-1}}(W)\right].$$

Note that the sequential information gain term represents the reduction in W's uncertainty from the previous posterior belief $B_{i-1|i-1}$ to the current, and so each term in the sum operates on W, $B_{i|i}$, **and** $B_{i-1|i-1}$. This alludes to a generalization of causal coding/decoding problems with a cost function $g(w_i, x_i, z_{i-1}, z_i)$, which in this case could plausibly be

$$g(w_i, x_i, z_{i-1}, z_i) = -\log \frac{dz_i}{dz_{i-1}}(w_i) + \alpha\eta(x_i) \tag{3.1}$$

where $Z_i \in \mathsf{Z} = \mathcal{P}(\mathsf{W})$ is a *decision variable* that can be any belief about the message. In this dissertation, we plan to build on this example and formulate general problems that capture this generalization and further elucidate an interplay between information theory and control theory within the context of both designing optimal strategies and performing inverse optimal control to characterize cost functions for which fixed strategies are optimal.

3.0.2 Chapter Outline and Main Results

We now outline the chapter, where in each section we provide bullet points about how it differs from other formulations and its main results.

Section 2.5 provided the problem setup. We emphasize the following properties that make it differ from traditional approaches:

- the Markov process source has a general alphabet W

- the traditional cost function $g(w_i, z_i)$ is replaced by

$$g(w_i, x_i, z_{i-1}, z_i) = \rho(w_i, z_{i-1}, z_i) + \alpha\eta(x_i) \tag{3.2}$$

- decision variables lie in arbitrary spaces X and Z

Section 3.1 considers a fixed cost function (3.2) and finding optimal coordination strategies (e, d). Results include:

- a structural result demonstrating the existence of optimal coordination strategies operating on sufficient statistics, capturing traditional results [20] as a special case.

Section 3.2 considers the *sequential information gain* cost function (3.1) with $\mathsf{Z} = \mathcal{P}(\mathsf{W})$ and finding optimal coordination strategies. Results include:

- an optimal coordination strategy always specifies $Z_i = B_{i|i}$
- a characterization of the problem as cost-penalized maximization of mutual information $I(W^n; Y^n)$

The first result uses dynamic programming and the second law of thermodynamics for Markov chains [52]. It synergizes with work in [54] but differs in how this is cast in the causal coding/decoding framework and the information gain cost (3.1).

Section 3.3 demonstrates how message point communication with feedback could be posed as a specific instantiation of our problem framework with the information gain cost (3.1). We show that the posterior matching scheme [16] is an optimal coordination strategy for the information gain cost (3.1) and source model $W_i = W_{i-1}$ with $\mathsf{W} = [0, 1]$. This example generalizes the result of [33] because here Z_i is a decision variable.

Section 3.4 provides example problems for which the aforementioned results apply, and shows how:

- under a particular constraint, the hidden Markov model and nonlinear filter [22] are an optimal coordination strategy for the information gain cost (3.1) with $\mathsf{W} = \mathsf{X}$
- the structural results aid the design of optimal and 'user-friendly' coordination strategies for brain-machine interfaces [21]

The first example is related to the variational characterization of the optimality of the nonlinear filter [54], but is different due to the information gain cost (3.1).

3.1 Main Structural Results

In this section, we prove that - under mild technical assumptions - for a general class of cost functions (ρ, η, α) inducing an average cost specified in (2.58), an optimal belief-based policy-estimator pair exists with the structure as shown in Figure 3.3.

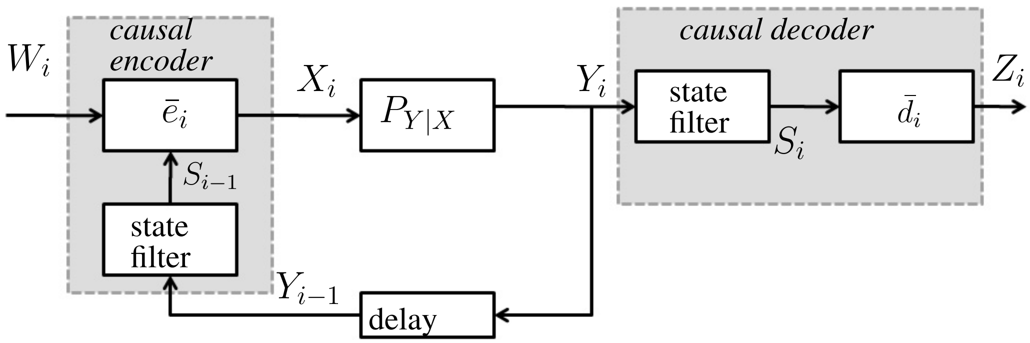

Figure 3.3: Structural result and sufficient statistics.

We first consider the basic solution approach to the problem by first demonstrating an example with two time steps. The essence of the idea is as follows:

$$\begin{aligned}
&\min_{e_1,d_1,e_2,d_2} \sum_{i=1}^{2} \mathbb{E}\left[\rho(W_i, Z_{i-1}, Z_i) + \alpha\eta(X_i)\right] \\
&= \mathbb{E}\left[\underbrace{\min_{e_1} \mathbb{E}\left[\alpha\eta(X_1)|\tilde{E}_1\right]}_{\text{controller at stage 0}}\right] \\
&+ \mathbb{E}\left[\underbrace{\min_{e_2,d_1} \mathbb{E}\left[\rho(W_1, Z_0, Z_1) + \alpha\eta(X_2)\middle| Z_0, Y^1, \tilde{E}_2, Z_1\right]}_{\text{controller at stage 1}}\right] \\
&+ \mathbb{E}\left[\underbrace{\min_{d_2} \mathbb{E}\left[\rho(W_2, Z_1, Z_2)\middle| Z_1, Y^2, Z_2\right]}_{\text{controller at stage 2}}\right]. \qquad (3.3a)
\end{aligned}$$

Note that by grouping (d_i, e_{i+1}) in this manner, in stage i, $d_i : \mathsf{Y}^i \to \mathsf{Z}$ and $e_{i+1} : \mathsf{W}^{i+1} \times \mathsf{Y}^i \to \mathsf{X}$ have access to a common piece of information, y^i (and thus also z_i). Note that $\tilde{e}_{i+1}(w^i, y^i)(\cdot) \equiv \tilde{e}_{i+1}(\cdot) : \mathsf{W} \to \mathsf{X}$ is a mapping as given by (2.50), whose alphabet, $\tilde{\mathsf{E}}$, does not grow with i. We secondly consider the belief $b_{i|i}$, which is a function of $(y^i, e_1, \ldots, e_i)$ and whose alphabet, $\mathcal{P}(\mathsf{W})$, does not grow with i. The 'control' action taken by the encoder (decoder) are given by $\tilde{E}_{i+1}$ (Z_i) respectively.

We next demonstrate that the conditional expectations in (3.3) can be described in terms of these and other variables whose alphabets do not grow with i:

Lemma 3.1.1. *For a fixed policy* $\gamma = (e, d)$*, define* $b' = b_{i|i} \in \mathcal{P}(\mathsf{W})$ *as in* (2.53)*, and* $\tilde{e}_{i+1}(w^i, y^i)(\cdot) \triangleq \tilde{e}_{i+1}(\cdot) \in \tilde{\mathsf{E}}$ *as in* (2.50)*. Define the state space* $\mathsf{S} = \mathsf{Z} \times \mathcal{P}(\mathsf{W})$ *and control space* $\mathsf{U} = \tilde{\mathsf{E}} \times \mathsf{Z}$ *with* $s_i \in \mathsf{S}, u_i \in \mathsf{U}$ *given by*

$$s_i = (z_{i-1}, b_{i|i}), \qquad u = (\tilde{e}_{i+1}, z_i). \tag{3.4}$$

Then

$$\mathbb{E}\left[\rho(W_i, Z_{i-1}, Z_i)|Z_{i-1}, Y^i, Z_i\right] = \bar{\rho}(S_i, Z_i) \tag{3.5a}$$

$$\mathbb{E}\left[\eta(X_{i+1})|Z_{i-1}, Y^i, \tilde{E}_{i+1}\right] = \bar{\eta}(S_i, \tilde{E}_{i+1}). \tag{3.5b}$$

To emphasize, this demonstrates state (s) and control (u) variables whose alphabets do not grow with i, for which 'distortion' and 'cost' like functions solely operate on. The definitions of $\bar{\rho}$ and $\bar{\eta}$ along with the lemma's proof can be found in Appendix A. We now demonstrate that these state and control variables comprise a controlled Markov chain:

Lemma 3.1.2. *The state* $s_i = (z_{i-1}, b_{i|i})$ *and control* $u = (\tilde{e}_{i+1}, z_i)$ *variables comprise a controlled Markov chain:*

$$\text{(a)} \quad J^{\alpha}_{n,\gamma} = \mathbb{E}_{\gamma}\left[\sum_{i=0}^{n} \bar{g}_i(S_i, U_i)\right] \tag{3.6}$$

$$\bar{g}_i(s, u) \triangleq \begin{cases} \alpha\bar{\eta}(s_i, \tilde{e}_{i+1}) & i = 0 \\ \bar{\rho}(s_i, z_i) + \alpha\bar{\eta}(s_i, \tilde{e}_{i+1}) & 1 \le i \le n-1 \\ \bar{\rho}(s_i, z_i) & i = n \end{cases} \tag{3.7}$$

$$\begin{aligned} \text{(b)} \quad P_{S_{i+1}|S^i, U^i}(ds_{i+1}|s^i, u^i) &= P_{S_{i+1}|S_i, U_i}(ds_{i+1}|s_i, u_i) \\ &\triangleq Q_S(ds_{i+1}|s_i, u_i). \end{aligned}$$

The proof of (a) follows directly from the law of iterated expectation and the definition (3.5). The proof of (b) can be found in Appendix B. Now define the cost-to-go function at stage $n-k$ as $V_{n-k} : \mathsf{S} \to \mathbb{R}$. Then for $V_{n+1}(s) \equiv 0$ and $k = 1, \ldots, n$ define:

$$V_{n-k}(s) = \inf_{u \in \mathsf{U}} \left[\bar{g}_{n-k}(s,u) + \int_{s'} V_{n-k+1}(s') Q_S(ds'|s,u) \right]. \tag{3.8}$$

This allows for us to state our main theorem of this section:

Theorem 3.1.3. *If for each $s \in \mathsf{S}$, the infimum in* (3.8) *is attained and the functions $(V_k : k = 0, \ldots, n)$ are universally measurable, then there exists an optimal encoder/decoder policy (e^*, d^*) pair of the form*

$$\begin{aligned} e^*_{i+1}(w^{i+1}, y^i) &\equiv \bar{e}^*_{i+1}(w_{i+1}, z_{i-1}, b_{i|i}) & \text{(3.9a)} \\ d^*_i(y^i) &\equiv \bar{d}^*_i(z_{i-1}, b_{i|i}). & \text{(3.9b)} \end{aligned}$$

Proof. Using standard dynamic programming arguments [4, Chapter 8], we have that $J^{\alpha}_{n,\gamma^*} \geq \mathbb{E}[V_0(S_0)]$. Next, $J^{\alpha}_{n,\gamma^*} = \mathbb{E}[V_0(S_0)]$ and it can be implemented by a policy of the form (3.9) by a policy that attains the infimum of (3.8) for each s [4, Prop 8.6]. □

The structural result in graphical form is shown in Figure 3.3. Note that within the causal encoder, the first process is a filter that computes sufficient statistics. From here, these sufficient statistics are given to another encoder, $\bar{e}_i$, that uses them, along with the current source value, W_i, to specify the next channel input X_i. Analogously, the causal decoder is comprised of first the same recursive filter that computes sufficient statistics, followed by another decoder, $\bar{d}_i$, that computes Z_i.

We now note that 'universal measurability' [4] is usually satisfied:

Remark 5. *Standard technical assumptions guarantee universal measurability and that the infimum is attained; one example is as follows: (a)* W, X, Y, *and* Z *are compact Borel metric spaces, (b) ρ and η are lower semi-continuous, (c) $P_{Y|X}(dy|x)$ and $Q_W(dw|w')$ are continuous stochastic kernels, and (d)* $\tilde{\mathsf{E}}$ *is an equicontinuous space of functions.*

We also note our result generalizes the classical result of Walrand and Varaiya [20]:

Remark 6. *This result instantiates the result in [20] which assumes all alphabets are finite, $\eta \equiv 0$, and $\rho(w_i, z_{i-1}, z_i) \equiv \rho(w_i, z_i)$: (i) because of the finite alphabets and costs, the infimum is attained in* (3.8)*; (ii)* (3.4) *can be collapsed to $s_i = b_{i|i}$ because of the absence of z_{i-1} in the function ρ. Secondly, our proof technique differs from [20, Sec. IV] in that we replace the three-step proof technique of ([20, Thm 1, Lemma 1, Thm 2]) - which includes two DP arguments ([20, Thm 1,Thm 2]) - with a single DP argument.*

However, our emphasis is not solely on allowing general alphabets or using the cost function of a particular form; both of these have in essence been accomplished using state augmentation and dynamic programming over general spaces. Rather, our emphasis is to carefully augment standard formulations to uncover an interplay information theory and control theory problems, as we shall see in the next section.

3.2 The Sequential Information Gain Cost

In this section, we specifically consider a class of problems that are not covered in traditional causal coding/decoding frameworks [20, 55],[35],[34, Ch. 6].

Traditional problems consider cost functions of the form $\rho(w_i, z_i)$ and assume that either all alphabets are finite [20, 55], or $\mathsf{W} = \mathsf{Z} = \mathbb{R}$ [35],[34, Ch. 6]. Motivated by the feedback communication example in Section 3.0.1, we now assume that $\mathsf{Z} = \mathcal{P}(\mathsf{W})$, the space of possible beliefs on the source. Secondly, we construct $\rho(w_i, z_{i-1}, z_i)$ to be a log-likelihood ratio that is suggestive of an 'information gain'-like quantity.

The following Lemma describes the relationship between $I(W^n; Y^n)$ and $I(W^n \to Y^n)$ for our problem setup (2.48)-(2.52). Because there is no feedback loop from Y to the generative process of W, these two quantities are equivalent:

Lemma 3.2.1. *For any 'sufficient statistic operating' encoder $\gamma \in \Gamma$ satisfying* (3.9a)*, i.e. $x_{i+1} = \bar{e}_{i+1}(w_{i+1}, z_{i-1}, b_{i|i})$, the following holds:*

$$I(W^n; Y^n) = I(W^n \to Y^n) = \sum_{i=1}^{n} I(W_i; Y_i | Y^{i-1}).$$

The proof is in Appendix C. Note that from the structural result in Theorem 3.1.3, there is no loss in performance for restricting attention to encoders of the form (3.9a). Under such encoders, note that the mutual information can be expressed as an accumulation of *sequential information gains*,

$$
\begin{aligned}
I(W^n;Y^n) &= \sum_{i=1}^{n} I(W_i;Y_i|Y^{i-1}) && (3.10a)\\
&= \sum_{i=1}^{n} \mathbb{E}\left[D\left(B_{i|i}\|B_{i|i-1}\right)\right] && (3.10b)\\
&= \sum_{i=1}^{n} \mathbb{E}\left[\log \frac{dB_{i|i}}{d\Phi(B_{i-1|i-1})}(W_i)\right] && (3.10c)
\end{aligned}
$$

where (3.10a) follows from Lemma 3.2.1; (3.10b) follows from (2.7) and (2.53); and (3.10c) follows from (2.4) and (2.55).

One may consider finding encoder policies e in order to maximize $I(W^n;Y^n)$, using a state space approach over the space of beliefs. [33] formulated a stochastic control problem where $B_{i-1|i-1}$ is a state variable and the only decision variable is the causal encoder's strategy - the decoder did not specify a decision variable Z_i. There, it was shown that when W is uniformly distributed on $\mathsf{W} = [0,1]$ and $(W_i = W : \ i \geq 1)$, the causal encoder given by the posterior matching scheme by Shayevitz and Feder [16] is an optimal solution to a control problem where costs are related to conditional mutual informations (3.10b). Anand and Kumar [56] have recently considered a related problem where $(W_i, \ i \geq 1)$ is a general Markov process over a *finite* alphabet, and the cost function is a conditional mutual information. There, also, however, the decoder did not specify a decision variable Z_i.

In this setting, we do not treat $B_{i|i}$ as a state variable; rather, we first consider a problem in the framework of causal coding/decoding, where the decoder's decision variable Z_i can be *any* possible belief: $\mathsf{Z} = \mathcal{P}(\mathsf{W})$. In order to reward larger information gains, we define an appropriate cost pertaining to the negative logarithm of the Radon-Nikodym derivative evaluated at w_i that is inspired by the expansion of mutual information given in (3.10c):

$$
\rho(w_i, z_{i-1}, z_i) = \begin{cases} -\log \frac{dz_i}{d\Phi(z_{i-1})}(w_i) & \text{if } z_i \ll \Phi(z_{i-1}) \\ \infty, & \text{otherwise} \end{cases} \tag{3.11}
$$

The reason we assign $\rho = \infty$ when $z_i \ll \Phi(z_{i-1})$ is because under any reasonable belief-setting strategy, if the belief about W_i given Y^{i-1} - the one-step prediction update (2.55) given by $\Phi(z_{i-1})$ - assigns zero probability mass to $A \in \mathcal{B}(\mathsf{W})$, then so should the belief about W_i given Y^i - which is given by Z_i.

We emphasize here that the beliefs on the source are themselves decision variables, which are what the causal decoder must specify. This viewpoint has been used within the sequential prediction literature [57] and statistical signal processing [58] but appears to not have been used as frequently in the literature that attempts to draw synergies between information theory and control.

Define $Z_0(A) = \mathbb{P}(W_0 \in A)$, the distribution on W_0. We now state the following useful Lemma that decomposes the cost into the state and distortion parts, that act on different aspects of the control input:

Lemma 3.2.2. *Under the information gain criterion* (3.11)*, for a state variable* $s_i = (z, b)$ *and control variable* $u_i = (\tilde{e}, z')$*,*

$$\bar{\eta}(s_i, \tilde{e}) = \int_{w \in \mathsf{W}} \eta\left(\tilde{e}\left(w\right)\right) \Phi(b)(dw) \tag{3.12}$$

$$\bar{\rho}(s_i, z') = \begin{cases} D\left(b \| z'\right) - D\left(b \| \Phi(z)\right) & b \ll z' \ll \Phi(z) \\ \infty & \textit{otherwise.} \end{cases} \tag{3.13}$$

The proof can be found in Appendix D.

With this, we state the main theorem of our section. It says that when treating beliefs as decision variables, under the information gain criterion (3.11), the optimal decision rule for the decoder is to select its belief about W_i to be $z_i = b_{i|i}$, and the optimal decision rule for the encoder is to maximize mutual information subject to a cost on channel inputs:

Theorem 3.2.3. *Under cost criterion* (3.13)*, there exists an optimal encoder/decoder policy* (e^*, d^*) *pair of the form*

$$e^*_{i+1}(w^{i+1}, y^i) \equiv \bar{e}^*_{i+1}(w_{i+1}, b_{i|i}) \tag{3.14}$$

$$d^*_i(y^i) \equiv \bar{d}^*_i(b_{i|i}) = b_{i|i} \tag{3.15}$$

where $b_{i|i} = \Lambda(b_{i-1|i-1}, y_i, \bar{e}^*_i(\cdot, b_{i-1|i-1}))$ *and the optimal cost is given by*

$$J^{\alpha}_{n,\gamma^*} = \min_{e \in \mathsf{E}} -I(W^n; Y^n) + \alpha \mathbb{E}_e\left[\sum_{i=1}^{n} \eta(X_i)\right]. \tag{3.16}$$

The proof can be found in Appendix E. The structural result corresponding to

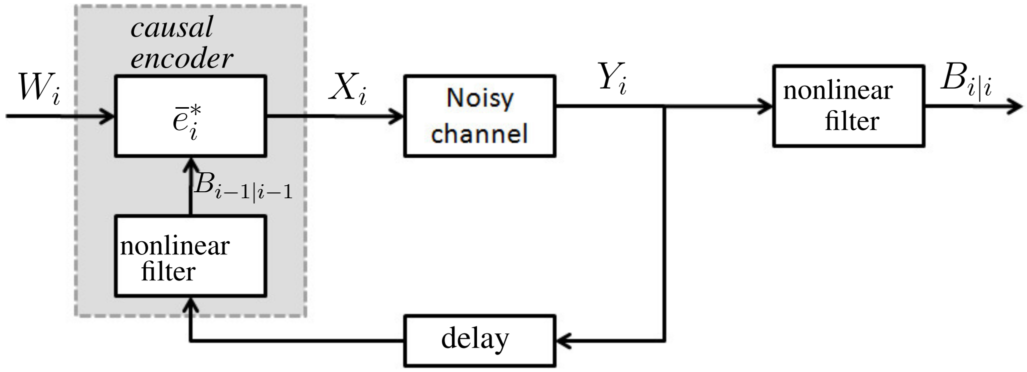

Figure 3.4: Simplified structural result with $\mathsf{Z} = \mathcal{P}(\mathsf{W})$ and sequential information gain cost (3.13).

equations (3.12)-(3.13) can be seen in Figure 3.4.

Remark 7. *The proof of Theorem 3.2.3 (Appendix E) uses dynamic programming, the second law of thermodynamics for Markov chains, and exploits how the divergence acts as a Lyapunov function for the stability of the nonlinear filter. This further demonstrates an interesting relationship between information theory and thermodynamics [54, 28]. This idea of using beliefs as decision variables where the posterior belief is optimal has been used in sequential prediction [57] and in variational approaches to nonlinear estimation [54], but within the constext of causal coding decoding problems, this is to the best of our knowledge, new.*

We will demonstrate in the examples section how this relates to the hidden Markov model and the nonlinear filter as well as the posterior matching scheme [16] for communication of a message point over a noisy channel with feedback.

3.3 Message Point Communication with Feedback and Posterior Matching Scheme

3.3.1 Likelihood Ratio Cost and Information Gain: Feedback Communication of a Message Point

Given that the natural mathematical framework to handle feedback is control theory, we consider the problem of communication over noisy channels with feedback from the dynamical systems perspective, and make use of recent sequential approaches to communication. This viewpoint has been made largely possible by a recent development in the information theory literature - the posterior matching (PM) scheme [16] - which generalizes other 'message-point' style feedback communication schemes [18, 19, 17]: rather than nR bits, a message point on the interval $[0, 1]$ is considered. The notion of "decoding nR bits" now becomes equivalent to determining the message point within an interval of length 2^{-nR} at the receiver (see Section 3.0.1).

The implementational details and fundamental limits are completely in line with traditional communication paradigms (see [16]) but there are subtle, yet striking differences. Because the message point is a point on the $[0, 1]$ line, there is no pre-specified block length; the system operates to sequentially give the user the information that is "still missing" at the receiver. Moreover, at each time step, the decoder specifies an output $Z_i \in \mathcal{P}(\mathsf{W})$, which is a belief about the message point. We now demonstrate how this notion of communication, and the problem of finding the optimal encoder with feedback, can be captured with our framework. Moreover, we will demonstrate that the PM scheme is an optimal solution to the problem.

Let $\mathsf{W} = [0, 1]$ and $\mathsf{Z} = \mathcal{P}(\mathsf{W})$. Further, let the source process be the 'repetition' Markov process $(W_i = W : i \geq 1)$ with W uniformly distributed over $[0, 1]$. If we assume that there is an expected cost constraint $\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\eta(X_i)] \leq L$, then we may formulate a communication problem of communicating a message point over a memoryless channel with causal feedback. First note that the mutual information between the message point and observations is given by

$$
\begin{aligned}
\frac{1}{n} I(W; Y^n) &= \frac{1}{n} \sum_{i=1}^{n} I(W; Y_i | Y^{i-1}) \\
&= \sum_{i=1}^{n} \mathbb{E}\left[\log \frac{dB_{i|i}}{dB_{i|i-1}}(W)\right].
\end{aligned}
$$

Shannon's converse to the channel coding theorem with feedback tells us that in order to achieve capacity, this aforementioned quantity must asymptotically be maximized. This allows for us to consider the following maximization problem:

$$
\max_{\gamma \in \Gamma} I(W; Y^n) + \alpha \mathbb{E}_{\gamma}\left[\sum_{i=1}^{n} \eta(X_i)\right]
$$

where α serves as a Lagrange multiplier such that under an optimal policy, the average state cost is upper bounded by L. We note that this can be captured in a causal coding/decoding framework by considering the sequential information gain distortion function (3.11). From Lemma 4.3.1, we note that a sufficient condition for optimality to this control problem is for

- $I(Y_i; Y^{i-1}) = 0$ for all i
- $X_i \sim P_X^*(\eta, P_{Y|X}, L)$, given in (2.9), for all i

Let $\mathsf{X} = \mathbb{R}$ and denote $F_X(\cdot)$ as the cumulative distribution function of the optimal input distribution $P_X^*(\eta, P_{Y|X}, L)$. The posterior matching (PM) scheme [16] simultaneously enables the two properties to hold for each i and is given by:

$$\begin{aligned} X_i &= F_X^{-1}\left(F_{W|Y^{i-1}}(W|Y^{i-1})\right) && \text{(3.17a)} \\ &= F_X^{-1}\left(B_{i-1|i-1}([0,W])\right) && \text{(3.17b)} \\ &= \bar{e}\,(W, Z_{i-1}) && \text{(3.17c)} \end{aligned}$$

where (3.17c) follows from Theorem 3.2.3 and because $W_i = W$. Note that the $F_{W|Y^{i-1}}$ operation constructs a uniform-$[0,1]$ random variable that is independent of the past channel outputs, and the F_X^{-1} shaping operation enables each input to be drawn according to the optimal channel input distribution $P_X^*(\eta, P_{Y|X}, L)$. Note that from (3.17c), the PM scheme can be interpreted as 'minimal' from our structural result in the causal coding/decoding framework in Section 3.2. Moreover, the causal encoder is *time-invariant*, and so likewise for the decoder acting as the nonlinear filter; thus, this means the PM scheme also can be interpreted as an instance of the inverse optimal control framework which will be discussed later via Lemma 4.3.1. See Figure 3.5 and its relationship with Figure 3.4. Also note

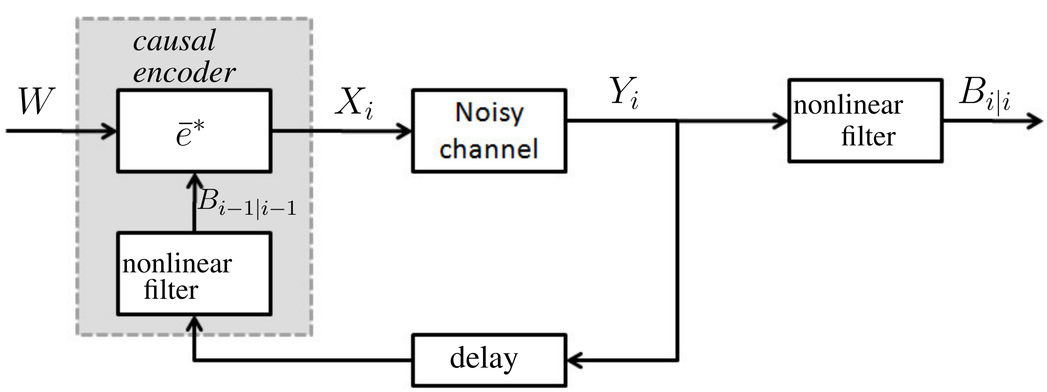

Figure 3.5: Posterior matching scheme by Shayevitz and Feder, interpreted as a time-invariant manifestation of the simplified structural result in Figure 3.4.

that in some cases, the time-reversibility condition for inverse optimal control in Section 4.2 is applicable: from Example2, the PM scheme (3.17) elicits reversibly feasible dynamics. Consider the additive Gaussian noise channel. Under the PM scheme, $(X_i, Y_i : i \geq 1)$ are jointly Gaussian (see (4.14e) and [14, Example 1]). Note that since $X_i = \tilde{X}_{i-1}$ in (4.14d), the joint reversibility sufficient condition is equivalent to reversibility of the Markov chain X. Since all stationary Gaussian processes are time-reversible, we see that in this scenario, the time-reversibility framework for our inverse optimal control framework is linked to the PM scheme. Although our control problem only addresses the maximization of mutual information - which is a necessary condition for reliable communication by the converse to the channel coding theorem - it can be shown that reliable communication, as defined in Section 3.0.1, results as a consequence of the mutual information maximization control problem under mild technical conditions [16].

3.4 Applications

In this section, we provide examples of the theorems and lemmas from previous sections.

3.4.1 Likelihood Ratio Cost and Information Gain: HMMs and the Nonlinear Filter

We now demonstrate that the information gain cost framework of Section 3.2 demonstrates the causal coding/decoding and information-theoretic optimality of the nonlinear filter in a specific sense. Related work on using variational principles to characterize the nonlinear filter was reported in [54]. However, demonstrating that the nonlinear filter is acting as an optimal controller with respect to this information gain cost function, is - to the best of our knowledge - new. We start by considering the following assumptions:

(i) the source and channel inputs have the same alphabets: $\mathsf{W} = \mathsf{X}$

(ii) the causal encoder alphabet $\mathsf{E}_i = \{e_i : \mathsf{W}^i \times \mathsf{Y}^{i-1} \to \mathsf{X}\} = \{=\}$ where $=$ is the identity function: $x_i = w_i$.

Under these conditions, the only feasible encoder simply specifies w_i as the channel inputs, and thus this becomes a hidden Markov model.

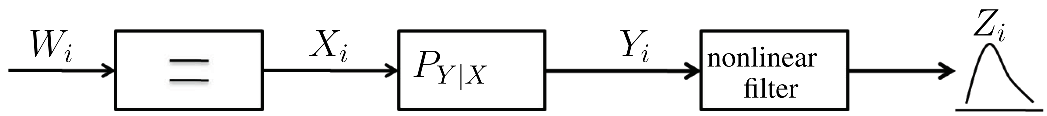

Figure 3.6: The information gain cost when the encoder set consists of only the identity function. This becomes a hidden Markov model where the nonlinear filter is an optimal solution.

As such, we can consider maximizing the mutual information from W^n to Y^n over all possible causal decoder policies. As such, the optimal design of e_i disappears and the focus becomes optimal design of $\{d_i\}$. We now show that, assuming $Z_0(A) = \mathbb{P}(W_0 \in A)$, the optimal policy for the decoder is given by the true posterior - which can be computed recursively using the nonlinear filter:

Lemma 3.4.1. *Under assumptions (i) and (ii) above, and cost functions $\eta \equiv 0$ and ρ given by* (3.11)

$$\rho(w_i, z_{i-1}, z_i) = \begin{cases} -\log \frac{dz_i}{d\Phi(z_{i-1})}(w_i) & \textit{if } z_i \ll \Phi(z_{i-1}) \\ \infty, & \textit{otherwise} \end{cases}$$

the policy $\bar{\gamma}$ consisting of the identity function encoder and nonlinear filter decoder $Z_i = B_{i|i} = \Lambda(Z_{i-1}, Y_i, =)$, is globally optimal where $J^{\alpha}_{n,\gamma^*} = -I(W^n; Y^n)$.

Proof. See Figure 3.6. Because E_i is a singleton consisting of the identity function, and because $\eta(x) = 0$, this follows directly from Theorem 3.2.3. $\square$

3.4.2 Structural Result: Brain-Machine Interfaces

A brain-machine interface (BMI) is a system that elicits a direct communication pathway between a human and an external device. In many cases, it is the objective of the human to control an external device merely by imagination, and the external device acquires neural signals, actuates some physical system, and perceptual feedback is given to the user to complete the loop. We now demonstrate how our structural result can be applied to the design of brain-machine interfaces that have a ‘user-friendly’ structure: displaying the minimal amount of useful perceptual feedback to the user, and designing an interaction strategy between the user and the external device.

Consider a brain-machine interface where a human has a desired high-level intent represented by the Markov process $(W_i : i \geq 1)$. At each time step, the human imagines a control signal X_i which is statistically linked to neural activity Y_i that is observed by the external device. For example, the statistics of Y_i are different when imagining a left-oriented movement $X_i = 0$ as compared to imagining a right-oriented movement $X_i = 1$ [59]. At each time step, the external device maps all its recorded observations Y^i to actuate some system, whose state is given by Z_i. Equally as important, the user gets perceptual feedback from the external device and allows this, along with causal information about the high-level intent, W^i, to specify the subsequent imagined control signal X_i.

Without loss of generality, because we do not know yet what perceptual feedback is the most relevant, we could consider a scenario where all information available to the decoder at any time i is fed back to the subject. Secondly, we may assume that we are planning to design the coordination strategy between the user and the interface: not only how the interface should take its observations and actuate the plant, but also what perceptual feedback should be specified back to the user *and* how the user should react to the perceptual feedback to specify the subsequent control signal X_i. In such a case, this problem boils down to our problem formulation in Section 2.5. Note that because of the causal nature of the problem, real-time constraints with a human in the loop obviate the possibility of using 'block-coding' like paradigms. Secondly, such settings are more complicated than simply optimally representing intent with an arithmetic coding procedure as in [60] - because of the inherent uncertainty also due to the noisy channel mapping intent to neural signals.

Almost all previous approaches to design BMIs failed to consider how the desired control signals change in response to sensory feedback. For example, many previous schemes simply attempt to recursively estimate X_i from Y^i under the assumption that $(X_i : i \geq 1)$ is a Markov process. However, as we know from our structural result, for an arbitrary objective with additive cost function, it is crucially important for the system to keep a running estimate, or belief, on W_i given Y^i. Moreover, it is critically important that the user and the system agree on an interaction protocol that specifies both *what* sensory feedback is provided to the user (e.g. the sufficient statistics) and *how* the user should react to this feedback in pursuit of high-level intent (e.g. the function $\bar{e}_i$).

Our structural result says that first a state filter can construct sufficient statistics $S_i = (Z_{i-1}, B_{i|i})$, and then the external device can actuate the plant using S_i and the user only needs to be fed back S_{i-1} as perceptual feedback. This information, along with the current high-level goal Z_i, is all that is needed to specify an optimal

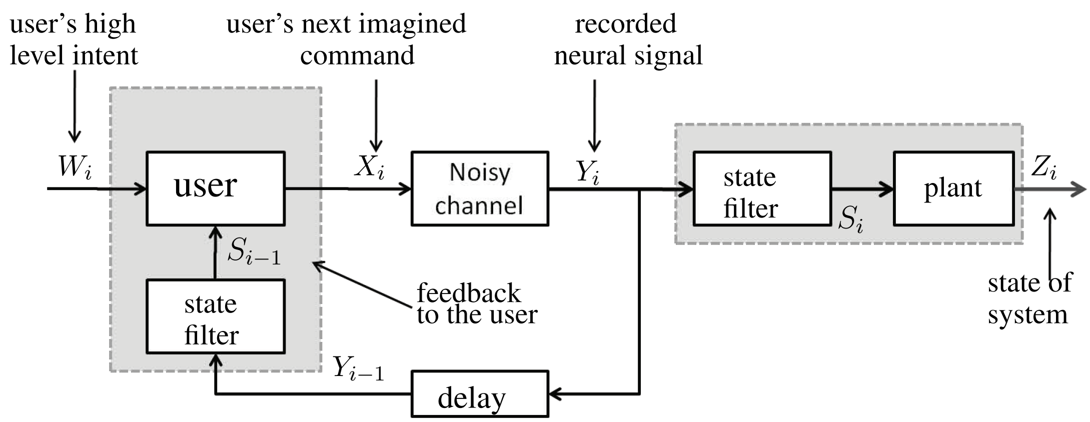

Figure 3.7: Structural result within the context of a brain-machine interface: in an optimal system, the user acts as part of the causal encoder. The other part accumulates all causal observations and summarizes them into sufficient statistics acting as perceptual feedback to the user.

causal encoder $\bar{e}_i$. See Figure 3.7.

In [21], we instantiate this idea in an EEG-based BMI in two steps. We assume the high-level intent can be mathematically represented as a Markov process $(W_i = W : i \geq 1)$ on W $= [0, 1]$ for which W is uniformly distributed over the $[0, 1]$ line. As such, this means we are assuming that the whole high-level intent is known to the user at all times. To relate this to a variety of practical applications, the user interprets the message point as a countably infinite sequence of symbols $D = (D_1, D_2, \ldots)$ in an *ordered* countable set $\mathcal{D}$ with a known statistical model (typically modeled as a fixed-order Markov process). Examples of the sequence D include an infinite sequence of text characters or an infinite sequence of small path arcs pertaining to a smooth path of bounded curvature. We use arithmetic coding [3] to develop a one-to-one mapping between any such sequence D and a point $W = \tau(D)$ uniformly distributed on the $[0, 1]$ line. We subsequently use an EEG system and specify a binary-input (left/right motor imagery) noisy channel with a spatial filter to extract beliefs $B_{i|i}$ sequentially [59]. With this, we implement the Posterior Matching scheme for the binary symmetric channel [17, 16]. Here, what is nice for a human in the loop is that $\bar{e}_i = \bar{e}$, and secondly, for the BSC, it only requires a *functional* of the posterior $B_{i-1|i-1}$ to be given to the encoder at time i: the median (denoted as $m(B_{i-1|i-1})$) [17, 14]:

$$X_i = \begin{cases} 0, & W < m(B_{i-1|i-1}) \\ 1, & W \geq m(B_{i-1|i-1}) \end{cases}. \tag{3.18}$$

Because of the one-to-one mapping τ, at time i, this can be implemented by visually displaying the median path $\tau^{-1}(m(B_{i-1|i-1}))$ on the screen and instructing the user to obey the time-invariant PM scheme (3.18) within the context of the median path. This simply means performing a lexicographically comparison to D (i.e. identify the first location where the sequences differ and perform a symbol-based comparison). We have successfully implemented this to demonstrate reliable text spelling and two-dimensional smooth path specification. Secondly, wedding arithmetic coding with the PM scheme has the added benefit that a natural 'propagation' of uncertainty ensues: the locations where D and $\tau^{-1}(m(B_{i-1|i-1}))$ differ increase to later and later parts of their sequences; this leads to a natural real-time implementation plausibility. Remote-control of an unmanned aerial vehicle using this paradigm has recently been shown in [61].

We also comment how the PM scheme by Shayevitz and Feder [16] is particularly relevant here: formulating this problem as one where the encoder has one of 2^{nR} hypotheses would mean that the human agent attempting to elicit neural control of an external device would have to implement an a strategy that differentiates possible inputs based upon one of 2^{nR} hypotheses. Even with visualization, this could be cumbersome. Moreover, it is unclear how the design specification would change when $n = 100$ as compared to when $n = 101$. Remarkably, using the posterior matching framework makes this problem truly solvable both theoretically and practically - by simply changing the starting point to be $\mathsf{W} = [0, 1]$ and $\mathsf{Z} = \mathcal{P}(\mathsf{W})$ and defining an appropriate information gain cost criterion. These observations speak to the fragility at which information theoretic problems with the same fundamental limits can be formulated.

The structural result demonstrated in this paper now enables the opportunity to design many brain-machine interface paradigms for a variety of cost functions beyond the the information gain paradigm and with assumption that $W_i = W_{i-1}$. The structural result has the potential more generally to enable an interesting intersection of desires on one platform: (i) guaranteed optimality from a decision-theoretic viewpoint; (ii) elucidation of the minimal amount of perceptual feedback information required to optimally display to the user; and (iii) potential ease-of-use when (e.g. when $\bar{e}_i = \bar{e}$ and it has a simple operational interpretation).

CHAPTER 4

INVERSE OPTIMAL CONTROL IN INTERACTIVE DECISION MAKING PROBLEMS

In the last chapter, we demonstrated a structural result for the two-agent team decision problem described in Section 2.5 and showed that dynamic programming (DP) provides a general methodology to solve this team decision problem. But this involves performing dynamic programming over the space of probability measures (Theorem 3.1.3), which is a hard problem.

In this chapter, we focus on an alternate approach - the 'inverse optimal control' approach, that can help in bypassing the dynamic programming step in certain cases. In this approach, we identify a fixed strategy of the agents and verify if it is optimal. Verification is done by identifying a set of "easy-to-describe" cost functions for which this fixed strategy is optimal using our inverse optimal control result. If the actual cost function falls in this set, then we know the policies we started with are in fact optimal, and there is no further need to perform the dynamic programming step. One downside in this approach is the guesswork involved in identifying the policies at the start.

In Section 4.1, we provide our inverse optimal control result for a two-agent team decision system based on an information-theoretic approach. We identify a set of "easy-to-describe" cost functions - through the variational equations for rate-distortion and capacity-cost functions - for which a fixed policy is optimal. As a consequence of this result, we were able to make an interesting connection of inverse optimal control with time reversibility as discussed below.

In Section 4.2, we extend our inverse optimal control result, to show that if a fixed coordination policy elicits reversibly feasible dynamics and a condition on time-reversibility, then it is a sufficient condition for the policy to be optimal. We then look at the following examples in Section 4.4, show that they are inverse-control optimal and deduce the cost functions for which the schemes are optimal.

- Gauss-Markov source, AGN channel pair - pertains to the decentralized control problems in [34, Ch 6],[35] with quadratic state cost and squared error distortion,

- Markov counting-function source, Z channel pair - pertains to the $\cdot/M/1$ queue for timing channels [30, 31],
- Markov counting-function source, 'inverted E' channel pair - pertains to Blackwell's trapdoor communication channel [36, 37, 38].

4.1 Inverse Optimal Control with Stationary Markov Coordination Strategies

In the last chapter, we demonstrated that for a specific "information-gain" related cost function (3.11), there existed an optimal encoder policy of the form $X_i = \bar{e}_i^*(W_i, Z_{i-1})$ and decoder policy of the form

$$Z_i = B_{i|i} = \Lambda\left(B_{i-1|i-1}, Y_i, \bar{e}_i^*(\cdot, B_{i-1|i-1})\right) = \bar{d}(Z_{i-1}, Y_i)$$

In light of this, we now consider a general Markov process $W \in \mathsf{W}$ and $Z \in \mathsf{Z}$ where Z need not be $\mathcal{P}(\mathsf{W})$ and fix the coordination strategy $\bar{\pi}$ to be *stationary Markov* (SM), meaning that for fixed functions $\bar{e} : \mathsf{W} \times \mathsf{Z} \to \mathsf{X}$ and $\bar{d} : \mathsf{Z} \times \mathsf{Y} \to \mathsf{Z}$, the following holds:

$$\begin{aligned} x_i &= \bar{e}(w_i, z_{i-1}) && \text{(4.1a)} \\ z_i &= \bar{d}(z_{i-1}, y_i). && \text{(4.1b)} \end{aligned}$$

See Figure 4.1. Equation (4.1b) is sometimes termed a decoder of 'finite-memory' [20],[62]. Since the encoder and decoder both utilize z_{i-1}, this can be also interpreted as a collection of 'equi-memory' encoders and decoders [34, Definition 6.3.2]. For a fixed SM coordination strategy $\bar{\gamma}$, we compare it against arbitrary policies of the form $\gamma = (e_1, \ldots, e_n, d_1, \ldots, d_n)$ where $e_i : \mathsf{W}^i \times \mathsf{Y}^{i-1} \to \mathsf{X}$ is given in (2.49), and $d_i : \mathsf{Y}^i \to \mathsf{Z}$ is given in (2.52). Here we identify the structure of cost functions $\rho(w_i, z_{i-1}, z_i)$ under which $\bar{\gamma}$ is globally optimal.

Definition 4.1.1. *A coordination strategy γ is inverse-control optimal for a source-channel pair $(P_{W^n}, P_{Y|X})$ if $J^{\alpha}_{n,\gamma} \leq J^{\alpha}_{n,\gamma'}$ for all $\gamma' \in \Gamma$ for some $\alpha \geq 0$, $\rho : \mathsf{W} \times \mathsf{Z} \times \mathsf{Z} \to \mathbb{R}$ and $\eta : \mathsf{X} \to \mathbb{R}_+$ in* (2.58).

To develop high-level conditions under which γ is indeed inverse-control optimal, we first develop some preliminary machinery. Fix a specific ρ and η function.

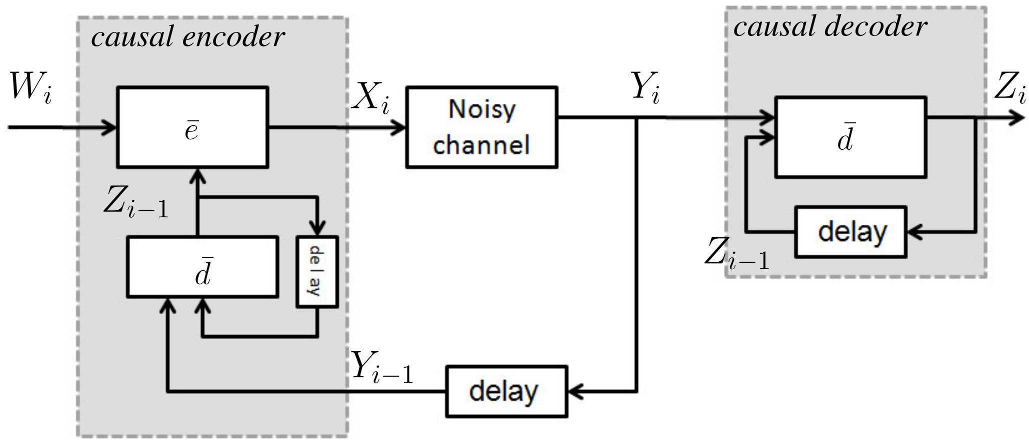

Figure 4.1: A stationary Markov coordination strategy.

For any coordination strategy γ, define $P^{\bar{\gamma}}_{Z^n|W^n}$ as the conditional distribution induced statistical law under γ and also define:

$$D_\gamma \triangleq \frac{1}{n}\mathbb{E}_\gamma\left[\sum_{i=1}^{n}\rho(W_i, Z_{i-1}, Z_i)\right] \tag{4.2}$$

$$L_\gamma \triangleq \frac{1}{n}\mathbb{E}_\gamma\left[\sum_{i=1}^{n}\eta(X_i)\right]. \tag{4.3}$$

Define the rate-distortion function for P_{W^n} and ρ as [3]

$$R_n\left(\rho, P_{W^n}, D\right) \triangleq \min_{P_{Z^n|W^n}:\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\rho(W_i,Z_{i-1},Z_i)\right]\le D} \frac{1}{n}I(P_{W^n}, P_{Z^n|W^n}) \tag{4.4}$$

and denote $P^*_{Z^n|W^n}\left(\rho, P_{W^n}, D\right)$ as the minimizer in (4.4). We now state the following standard lemma:

Lemma 4.1.2. *Fix a γ,P_{W^n}, $P_{Y|X}$, ρ, and η. Then*

$$\begin{aligned}
R_n\left(\rho, P_{W^n}, D_\gamma\right) &\le \frac{1}{n}I\left(P_{W^n}, P^{\bar{\gamma}}_{Z^n|W^n}\right) && \text{(4.5a)}\\
&\le \frac{1}{n}I\left(P_{W^n}, P^{\gamma}_{Y^n|W^n}\right) && \text{(4.5b)}\\
&\le \frac{1}{n}\sum_{i=1}^{n}I(P^{\gamma}_{X_i}, P_{Y|X}) && \text{(4.5c)}\\
&\le C\left(\eta, P_{Y|X}, L_\gamma\right) && \text{(4.5d)}
\end{aligned}$$

where equality holds if and only if:

- *(a)* $P^{\bar{\gamma}}_{Z^n|W^n} = P^*_{Z^n|W^n}(\rho, P_{W^n}, D_\pi)$
- *(b)* $I\left(P_{W^n}, P^{\bar{\gamma}}_{Z^n|W^n}\right) = I\left(P_{W^n}, P^{\gamma}_{Y^n|W^n}\right)$
- *(c)* $I(Y_i; Y^{i-1}) = 0$ *for each* i
- *(d)* $P^{\gamma}_{X_i} = P^*_X(\eta, P_{Y|X}, L_\gamma)$ *for each* i

The proof is standard [3] but for the sake of completeness, we include it in Appendix F. This leads to an intermediate sufficient condition for inverse control optimality that applies for *any* $\gamma \in \gamma$ (e.g. γ need not be stationary-Markov):

Lemma 4.1.3. *If a policy* $\gamma \in \Gamma$ *results in* (4.5) *holding with equality, then it is inverse control optimal.*

Proof. Note that $J^{\alpha}_{n,\gamma} = D_\gamma + \alpha L_\gamma = \langle (1, \alpha), (D_\gamma, L_\gamma) \rangle$. Define $\mathcal{R} = \{(D_{\gamma'}, L_{\gamma'} : \gamma') : \gamma' \in \Gamma\}$. Define $\widetilde{\Gamma}$ to be the set of randomized policies in Γ. Note that any $\gamma' \in \widetilde{\Gamma}$ still induces a conditional distribution $P^{\bar{\gamma}'}_{Z^n|W^n}$ and thus an induced $D_{\gamma'}$ and $L_{\gamma'}$ so that we may define $\widetilde{\mathcal{R}} = \{(D_{\gamma'}, L_{\gamma'} : \gamma') : \gamma' \in \widetilde{\Gamma}\}$. Clearly, $\mathcal{R} \subset \widetilde{\mathcal{R}}$, and secondly, $\widetilde{\mathcal{R}}$ is convex. Next, note that if (4.5) holds with equality for some $\gamma \in \Gamma$, then $D_\gamma \le D_{\gamma'}$ for for any $\gamma' \in \widetilde{\Gamma}$ for which $L_{\gamma'} \le L_\gamma$ from the definition of $R_n(\rho, P_{W^n}, D_\gamma)$ in (4.4) and $C\left(\eta, P_{Y|X}, L_\gamma\right)$ in (2.9) (See also [5, Lemma 1]). Thus (D_γ, L_γ) is a boundary point of $\widetilde{\mathcal{R}}$. Therefore there exists a supporting hyperplane parametrized by $\alpha \ge 0$ that intersects (D_γ, L_γ):

$$J^{\alpha}_{n,\gamma} = \langle (1, \alpha), (D_\gamma, L_\gamma) \rangle \le \langle (1, \alpha), (D_{\gamma'}, L_{\gamma'}) \rangle = J^{\alpha}_{n,\gamma'}.$$

for all $\gamma' \in \widetilde{\Gamma} \supset \Gamma$ [63], where $\langle \cdot, \cdot \rangle$ denotes inner product. □

We now consider SM policies for which condition (c) in Lemma 4.1.2 holds, and demonstrate a stationary Markov relationships between (W^n, Z^n) random variables:

Lemma 4.1.4. *If an SM coordination strategy* $\bar{\gamma}$ (4.1) *induces the channel outputs* $(Y_i : 1 \le i \le n)$ *being i.i.d., then*

$$P^{\bar{\pi}}_{Z_i|Z^{i-1}=z^{i-1}, W^n=w^n}(dz_i) \equiv Q^{\bar{\pi}}_{Z'|Z,W'}(dz_i|z_{i-1}, w_i) \tag{4.6a}$$

$$P^{\bar{\pi}}_{Z_i|Z^{i-1}=z^{i-1}}(dz_i) \equiv Q^{\bar{\pi}}_{Z'|Z}(dz_i|z_{i-1}). \tag{4.6b}$$

The proof of this can be found in Appendix G and exploits the equivalence between a random process being a time-homogeneous Markov chain and it being represented as an iterated function system [64]. With this, we can now state the main theorem of this section:

Theorem 4.1.5. *If under a SM policy $\bar{\gamma}$, $(Y_i : 1 \leq i \leq n)$ are i.i.d. and $I\left(P_{W^n}, P^{\bar{\gamma}}_{Z^n|W^n}\right) = I\left(P_{W^n}, P^{\bar{\gamma}}_{Y^n|W^n}\right)$, then $\bar{\pi}$ is inverse control optimal with ρ and η given by:*

$$\eta(x) \propto_+ D\left(P_{Y|X=x} \| P^{\bar{\pi}}_Y\right) \tag{4.7a}$$

$$\rho(w_i, z_{i-1}, z_i) \propto_+ -\log \frac{dQ^{\bar{\pi}}_{Z'|Z,W'}(\cdot|z_{i-1}, w_i)}{dQ^{\bar{\pi}}_{Z'|Z}(\cdot|z_{i-1})}(z_i) \tag{4.7b}$$

where $\propto_+$ denotes proportional to with a positive constant.

Proof. Note that it suffices to show that (4.5) holds with equality and then invoke Lemma 4.1.3. First note that from the theorem definition, clearly conditions (b) in Lemma 4.1.2 holds with equality. Since $(Y_i : i = 1, \ldots, n)$ are i.i.d. and since the channel is memoryless, it follows that the $(X_i : i = 1, \ldots, n)$ are identically distributed and so condition (c) in Lemma 4.1.2 holds with equality. Thus the two remaining conditions are to show that conditions (a) and (d) in Lemma 4.1.2 hold with equality.

The variational equations for an optimal solution to (4.4) state that a necessary and sufficient condition for $P^{\bar{\gamma}}_{Z^n|W^n} = P^*_{Z^n|W^n}(\rho, P_{W^n}, D_\pi)$ is the following relationship [65]:

$$\frac{dP_{Z^n|W^n=w^n}}{dP_{Z^n}}(z^n) = \zeta(w^n)e^{-\alpha_2\left(\sum_{i=1}^n \rho(w_i,z_{i-1},z_i)\right)}. \tag{4.8}$$

For our case, note that

$$\begin{aligned}\log\frac{dP^{\bar{\gamma}}_{Z^n|W^n=w^n}}{dP^{\bar{\gamma}}_{Z^n}}(z^n) &= \sum_{i=1}^n \log \frac{dP^{\bar{\gamma}}_{Z_i|Z^{i-1}=z^{i-1},W^n=w^n}}{dP^{\bar{\gamma}}_{Z_i|Z^{i-1}=z^{i-1}}}(z_i) \\ &= \sum_{i=1}^n \log \frac{dQ^{\bar{\gamma}}_{Z'|Z,W'}(\cdot|z_{i-1}, w_i)}{dQ^{\bar{\gamma}}_{Z'|Z}(\cdot|z_{i-1})}(z_i) \end{aligned} \tag{4.9}$$

where (4.9) follows from Lemma 4.1.4. Thus we see that with ρ given by (4.7b), from (4.8) we see that condition (a) of Lemma 4.1.2 holds with equality.

Lastly, condition (d) of Lemma 4.1.2 holds with equality if and only if each

$P_{X_i}^{\bar{\gamma}} \sim P_X^*(\eta, P_{Y|X}, L_\gamma)$. Variational arguments [5, Lemma 3],[66, p. 147] demonstrate that this criterion is equivalent to (4.7a). □

Corollary 4.1.6. *If the function* $\bar{d}(z_{i-1}, \cdot) \triangleq \bar{d}_{z_{i-1}}(\cdot)$ *in* (4.1b) *is invertible, then condition* (4.7b) *in Theorem 4.1.5 becomes*

$$\rho(w_i, z_{i-1}, z_i) \propto_+ \log \frac{dP_{Y|X=\bar{e}(w_i,z_{i-1})}}{dP_Y^{\bar{\pi}}} \left(\bar{d}_{z_{i-1}}^{-1}(z_i)\right).$$

We now first relate this to 'source-channel' matching and how it is in some sense it is also 'natural' within the inverse control framework to have a distortion function of the form $\rho(w_i, z_{i-1}, z_i)$:

Remark 8. *The problem setup leading up to Theorem 4.1.5 is philosophically inspired by the 'source channel matching' work [5, 67], but here, we are relating this to a causal coding-decoding problem with* **causal encoder feedback***, and* **time-invariant additive costs***. These two properties appear to make the distortion function* $\rho(w_i, z_{i-1}, z_i)$ *- as compared to* $\rho(w_i, z_i)$ *- crucially important: note the time-invariant statistical relationships in Lemma 4.1.4 and how they relate to* $\rho(w_i, z_{i-1}, z_i)$ *in* (4.9) *and Corollary 4.1.6 pertaining to condition (a) in Lemma 4.1.2. With this more general* $\rho(w_i, z_{i-1}, z_i)$ *framework, we can characterize time-invariant cost functions for problems where neither* W *nor* Z *are stationary (see the linear quadratic Gaussian decentralized control and M/M/1 queue examples in Section 4.4).*

Next, we demonstrate how time-reversibility of an appropriately defined Markov chain can serve for Theorem 4.1.5 - and thus inverse control optimality - to hold for SM coordination strategies.

4.2 Time-Reversibility of Markov Chains and Inverse Optimal Control for Stationary Markov Policies

Time reversibility plays an important role in disciplines concerning dynamical systems, e.g. in physics (conservation laws); statistical mechanics (in terms of equilibrium states); stochastic processes (e.g. queuing networks [23, 24] and convergence rates of Markov chains [25, ch 20]); and biology (e.g. trans paths in ion channels [26]). However, its use in acting as a sufficient condition to saturate fundamental information-theoretic limits appears to be somewhat limited. One noteworthy exception is how Mitter and colleagues have related Markov chain reversibility to rate of entropy production in non-equilibrium thermodynamics [27],[28, Remark 2.1].

In queuing systems, the celebrated Burke's theorem [24, 29] uses Markov chain time reversibility to show that, in a certain stochastic dynamical system - an M/M/1 queue in steady-state - *the state of the system (queue) at time i is statistically independent of all outputs (departures) before time i.* This observation has been used in proving achievability theorems using for queuing timing channels [30, 31, 32], and for implementing recursive schemes that maximize mutual information according to the converse to the channel coding theorem with feedback [17, 18, 19, 16]. Here we demonstrate how time reversibility of Markov chains provides a sufficient condition for inverse optimal control with SM coordination strategies.

We first note that from (2.48), W is a time-homogenous Markov chain and so it can be represented as an iterated function system [64]:

$$W_i = \psi(\tilde{W}_i, W_{i-1}) \equiv \psi_{\tilde{W}_i}(W_{i-1}), \quad i \geq 1 \tag{4.10}$$

where $\tilde{W}_i$ are i.i.d. To ensure, (2.47), we assume

$$I(\tilde{W}_i; X^{i-1}, Y^{i-1}) = 0. \tag{4.11}$$

We next suppose the structure of the SM coordination strategy is such that the following assumption holds

Definition 4.2.1. *We say that the SM coordination strategy $\bar{\gamma} = (\bar{e}, \bar{d})$ elicits ' reversibly feasible dynamics ' if $I\left(P_{W^n}, P^{\bar{\gamma}}_{Z^n|W^n}\right) = I\left(P_{W^n}, P^{\gamma}_{Y^n|W^n}\right)$ and the statistical dynamics can be described as*

$$\begin{aligned} X_i &= f(\tilde{X}_{i-1}, \tilde{W}_i) \equiv f_{\tilde{X}_{i-1}}(\tilde{W}_i) && (4.12) \\ \tilde{X}_i &= g(X_i, Y_i) \equiv g_{X_i}(Y_i) && (4.13) \end{aligned}$$

where $f_{\tilde{X}_{i-1}} : \mathsf{W} \rightarrow \mathsf{X}$ and $g_{X_i} : \mathsf{Y} \rightarrow \mathsf{X}$ are $\mathbb{P}$-a.s. invertible functions for $i = 1, \ldots, n$.

Note that $\tilde{x}$ in condition (4.13) is the update to the state *after* the output of the channel is taken into consideration and *before* the source w is updated to the state.

We now show an example that is related to feedback communication with posterior matching [16].

Example 2. *Let* $\mathsf{W} = \mathsf{X} = [0, 1]$ *and* $\mathsf{Z} = \mathcal{P}(\mathsf{W})$. *Then the 'posterior matching' scheme [16] given by*

$$
\begin{aligned}
\tilde{W}_0 &\sim \textit{unif}[0,1], \quad \tilde{W}_i \equiv 1, i \geq 1 && (4.14a)\\
W_0 &= \tilde{W}_0 \quad W_i = W_{i-1},\ i \geq 1 && (4.14b)\\
Z_i &= B_{i|i} && (4.14c)\\
X_0 &= 0,\ X_i = \tilde{X}_{i-1}, i \geq 1 && (4.14d)\\
\tilde{X}_i &= Z_{i-1}([0, W_i]) = F_{X|Y}(X_i|Y_i) && (4.14e)
\end{aligned}
$$

elicits reversibly feasible dynamics . Note that this clearly is a SM coordination policy because for the decoder $B_{i|i}$ *is given by the nonlinear filter, and for the encoder, this follows from the first equality in* (4.14e)*. To verify that the last equality in* (4.14e) *holds, see [16, Corollary 6].*

Next, we consider a SM coordination strategy that results in a *birth-death* Markov chain [23, 24] where X can increase or decrease by at most 1 from time i to time $i + 1$ (see Figure 4.2):

Example 3. *Let* $\tilde{\mathsf{W}} = \mathsf{W} = \mathsf{X} = \mathsf{Y} = \mathsf{Z} = \mathbb{F}$ *for some field. Then the following SM coordination strategy*

$$
\begin{aligned}
W_i &= W_{i-1} + \tilde{W}_i && (4.15a)\\
Z_i &= Z_{i-1} + Y_i && (4.15b)\\
\tilde{X}_i &= X_i - Y_i && (4.15c)\\
X_i &= W_i - Z_{i-1} = \tilde{X}_{i-1} + \tilde{W}_i && (4.15d)
\end{aligned}
$$

elicits reversibly feasible dynamics . This follows from inspection.

See Figure 4.2.

Lemma 4.2.2. *Consider an SM coordination strategy with dynamics given by* (4.15) *where* $\mathbb{P}\left(\tilde{W}_i \in \{0, 1\}\right) = \mathbb{P}(Y_i \in \{0, 1\})$. *If* X *is a time-reversible Markov chain, then* $(X, \tilde{X})$ *is jointly a time-reversible Markov chain,* Y_i *are i.i.d., and* $\bar{\gamma}$ *is inverse-control optimal.*

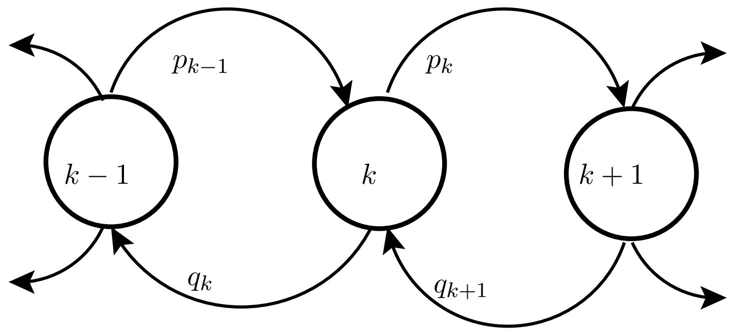

Figure 4.2: A birth-death Markov chain X.

The proof that $(X, \tilde{X})$ is jointly stationary and that Y_i are i.i.d. is a generalization [68] of the discrete-time Burke's theorem [29] from queuing theory. From there, the Lemma follows by simply invoking Definition 4.1.1 and Theorem 4.1.5. As of now, time-reversibility is only wed to inverse control optimality in the algebraic setup of Example 3. This alludes to there being a more general statement under which time-reversibility implies inverse optimality:

Lemma 4.2.3. *If an SM coordination strategy $\bar{\gamma}$ elicits reversibly feasible dynamics and $(X, \tilde{X})$ is jointly a time-reversible Markov chain, then $\bar{\gamma}$ is inverse-control optimal.*

Proof. We now develop a generalization to the discrete-time proof of Burke's theorem [29, 68] from queuing theory. Note that from Assumption 4.2.1 that

$$\begin{aligned} \tilde{W}_i &= f^{-1}_{\tilde{X}_{i-1}}(X_i) && (4.16) \\ Y_i &= g^{-1}_{X_i}(\tilde{X}_i). && (4.17) \end{aligned}$$

Now note that, from the time-reversibility assumption, we have

$$\left(X_1, \tilde{X}_1, \ldots, \tilde{X}_{i-1}, X_i\right) \stackrel{d}{=} \left(X_{2i-1}, \tilde{X}_{2i-2}, \ldots, \tilde{X}_i, X_i\right). \quad (4.18)$$

Re-arranging terms, we have

$$\begin{aligned} \left(X_i, \tilde{X}^{i-1}, X^{i-1}\right) &\stackrel{d}{=} \left(X_i, \tilde{X}_i^{2i-2}, X_{i+1}^{2i-1}\right) && (4.19) \\ \Rightarrow I(X_i; \tilde{X}^{i-1}, X^{i-1}) &= I(X_i; \tilde{X}_i^{2i-2}, X_{i+1}^{2i-1}) && (4.20) \\ \Leftrightarrow I(X_i; Y^{i-1}, X^{i-1}) &= I(X_i; \tilde{X}_i^{2i-2}, \tilde{W}_{i+1}^{2i-1}) && (4.21) \end{aligned}$$

where (4.21) follows from invariance of mutual information to bijective transformations and (4.12)-(4.13), (4.16)-(4.17). Analogously, from (4.19),

$$I(X_i; X^{i-1}|\tilde{X}^{i-1}) = I(X_i; X_{i+1}^{2i-1}|\tilde{X}_i^{2i-2}) \quad (4.22)$$

$$\Leftrightarrow I(X_i; X^{i-1}|Y^{i-1}) = I(X_i; \tilde{W}_{i+1}^{2i-1}|\tilde{X}_i^{2i-2}). \quad (4.23)$$

Therefore

$$I(X_i; Y^{i-1}) = I(X_i; \tilde{W}_{i+1}^{2i-1}) \quad (4.24)$$

$$= 0 \quad (4.25)$$

where (4.24) follows from the chain rule of mutual information (2.8) and subtracting (4.23) from (4.21); and (4.25) follows from (4.11). Because of the nature of the memoryless channel $P_{Y|X}$ in (2.51), it follows that $I(Y_i; Y^{i-1}) = 0$ for all i. Moreover, because Markov chain reversibility implies stationarity, it follows that $(Y_i : 1 \leq i \leq n)$ are i.i.d. Thus we can invoke Theorem 4.1.5. □

4.3 Information Gain Cost and Inverse Optimal Control

In the beginning of this section, we motivated the definition of stationary Markov coordination strategies, of the type $x_i = \bar{e}_i(w_i, z_{i-1})$ and $z_i = \bar{d}_i(z_{i-1}, y_i)$ by noting from Section 3.2 that such an optimal decoder exists when $\mathsf{Z} = \mathcal{P}(\mathsf{W})$ and the cost function is of the "information-gain" related structure (3.11):

$$X_i = \bar{e}_i^*(W_i, B_{i-1|i-1}) = \bar{e}_i(W_i, Z_{i-1})$$

$$Z_i = B_{i|i} = \Lambda\left(B_{i-1|i-1}, Y_i, \bar{e}_i^*(\cdot, B_{i-1|i-1})\right) = \bar{d}_i(Z_{i-1}, Y_i).$$

We now demonstrate that the information gain cost function in Section 3.2 can be seen to be a consequence of our inverse optimal control framework for any coordination strategy for $(Y_i : 1 \leq i \leq n)$ are i.i.d. and $\bar{d}$ is the nonlinear filter:

Lemma 4.3.1. *Let $\mathsf{Z} = \mathcal{P}(\mathsf{W})$. If a SM coordination strategy $\bar{\gamma}$ contains a nonlinear filter decoder $z_i = \bar{d}(z_{i-1}, y_i) = \Lambda(z_{i-1}, y_i, \bar{e}(\cdot, y_{i-1}))$ and $(Y_i : 1 \leq n)$ are i.i.d., then $\bar{\gamma}$ is inverse control optimal with information gain distortion $\rho(w_i, z_{i-1}, z_i) = -\log \frac{dz_i}{d\Phi(z_{i-1})}(w_i)$ and state cost function $\eta(x) = D\left(P_{Y|X=x} \| P_Y^{\bar{\gamma}}\right)$.*

The optimal cost is given by

$$J^{\alpha}_{n,\bar{\gamma}} = (\alpha-1)nC\left(\eta, P_{Y|X}, L_{\gamma}\right). \tag{4.26}$$

Proof. First note that under this policy $\bar{\gamma}$, $Z_i = B_{i|i}$. Now note that clearly

$$\mathbb{P}\left(W_i \in A|Y^i\right) = Z_i(A) = \mathbb{P}\left(W_i \in A|Z_{i-1}, Z_i\right), \tag{4.27a}$$

$$\mathbb{P}\left(W_i \in A|Y^{i-1}\right) = \Phi(Z_{i-1})(A) = \mathbb{P}\left(W_i \in A|Z_{i-1}\right). \tag{4.27b}$$

As such, we have that

$$\begin{aligned}
\frac{dz_i}{d\Phi(z_{i-1})}(w_i) &= \frac{dP^{\bar{\gamma}}_{W_i|Y^i}}{dP^{\bar{\gamma}}_{W_i|Y^{i-1}}}(w_i) \\
&= \frac{dP^{\bar{\gamma}}_{W_i|Z_{i-1}=z_{i-1},Z_i=z_i}}{dP^{\bar{\gamma}}_{W_i|Z_{i-1}=z_{i-1}}}(w_i) &(4.28)\\
&= \frac{dP^{\bar{\gamma}}_{Z_i|Z_{i-1}=z_{i-1},W_i=w_i}}{dP^{\bar{\gamma}}_{Z_i|Z_{i-1}=z_{i-1}}}(z_i) &(4.29)\\
&= \frac{dQ^{\bar{\gamma}}_{Z'|Z,W}(\cdot|z_{i-1},w_i)}{dQ^{\bar{\gamma}}_{Z'|Z}(\cdot|z_{i-1})}(z_i)\,(z_i) &(4.30)
\end{aligned}$$

where (4.28) follows from (4.27); (4.29) follows from a simple application of Bayes' rule: $\frac{\mathbb{P}(A|B,C)}{\mathbb{P}(A|B)} = \frac{\mathbb{P}(C|A,B)}{\mathbb{P}(C|B)}$; and (4.30) follows from Lemma 4.1.4. Also, since $Z_i = B_{i|i}$, it follows that $I(W^n;Y^n) = I(W^n;Z^n)$. Thus Theorem 4.1.5 applies and so $\bar{\gamma}$ is inverse control optimal. To characterize the final cost, note that for the associated α,

$$\begin{aligned}
J^{\alpha}_{n,\bar{\gamma}} &= -I(W^n;Y^n) + \alpha\mathbb{E}_{\bar{e}}\left[\sum_{i=1}^{n}\eta(X_i)\right] &(4.31)\\
&= -nC\left(\eta, P_{Y|X}, L\right) + \alpha\mathbb{E}_{\bar{e}}\left[\sum_{i=1}^{n}\eta(X_i)\right] &(4.32)\\
&= -nC\left(\eta, P_{Y|X}, L\right) + \alpha\left(\sum_{i=1}^{n}I(X_i;Y_i)\right) &(4.33)\\
&= (\alpha-1)nC\left(\eta, P_{Y|X}, L\right) &(4.34)
\end{aligned}$$

where (4.31) follows from Theorem 3.2.3; (4.32) follows from the fact that Theorem 4.1.5 applies which means that (4.5) holds with equality; (4.33) follows from the definition of mutual information and that $\eta(x) = D\left(P_{Y|X=x} \| P_Y^{\bar{\gamma}}\right)$; and (4.34) follows from the fact that (4.5) holds with equality. $\square$

Traditionally, inverse optimal control is performed through finding a control-Lyapunov function (4.1), which involves performing a sequential decomposition of the problem and finding a consistent value function (3.8). When $\mathsf{Z} = \mathcal{P}(\mathsf{W})$, this can be done using only the decision variables and stationary-Markov coordination strategies as in Section 3.2: $X_i = \bar{e}_i(W_i, Z_{i-1})$ and $Z_i = B_{i|i} = \Lambda\left(B_{i-1|i-1}, Y_i, \bar{e}_i(\cdot, B_{i-1|i-1})\right) = \bar{d}(Z_{i-1}, Y_i)$. This means that using a control-Lyapunov approach, first a sequential decomposition resting upon the structural result work in Section 3.1 would be needed, with the additional effort of showing that coordination strategies of the structural result form (3.9) can be reduced to stationary Markov strategies of the form (4.1). However, our inverse control optimality sufficient conditions apply for a general Z (which need not be $\mathcal{P}(\mathsf{Z})$) and do not involve a sequential decomposition. As such, the approach developed in this section - when applicable - appears to require'less effort' than typically required in arriving at an inverse optimal control result.

4.4 Examples

In this section, we provide examples of the theorems and lemmas from previous sections.

4.4.1 Inverse Optimal Control: Gauss-Markov Source and AGN Channel

Here we show that a stationary Markov coordination strategy consisting of a linear 'estimation error' encoder and MMSE decoder is inverse-control optimal for a Gauss-Markov Q_W and a power-constrained additive Gaussian channel. A variant of this problem for $\rho(w_i, z_{i-1}, z_i) \equiv \rho(w_i, z_i) = (w_i - z_i)^2$ has been studied by [35],[34].

See Figure 4.3. Let $\mathsf{W} = \mathsf{X} = \mathsf{Y} = \mathsf{Z} = \mathbb{R}$. The source is a Gauss-Markov process with i.i.d. $\tilde{W}_i \sim \mathcal{N}(0, \sigma_m^2)$,

$$W_0 \sim \mathcal{N}\left(0, \frac{\sigma_m^2 \sigma_v^2}{L + \sigma_v^2(1-\rho^2)}\right) \tag{4.35a}$$

$$W_i = \rho W_{i-1} + \tilde{W}_i \quad i \geq 1, \tag{4.35b}$$

$$I(\tilde{W}_i; X^{i-1}, Y^{i-1}) = 0, \quad i \geq 1. \tag{4.35c}$$

Note that we are not assuming that W is stationary. As such, this problem can be connected to problems in 'control over noisy channels. In such problems with quadratic cost and linear Gaussian dynamics, the essence of optimally solving the control over noisy channels problem is optimally solving this causal coding/decoding 'active tracking' problem [35],[34].

The channel additive with Gaussian noise (AGN):

$$Y_i \;\; = \;\; X_i + V_i \quad V_i \sim \mathcal{N}\left(0, \sigma_v^2\right). \tag{4.36}$$

A typical objective in practice is to design an encoder and decoder than can minimize the mean-squared error in estimating the source process, i.e., minimize $J(e^n, d^n) = \mathbb{E}\left[\sum_{i=1}^n (Z_i - W_i)^2 + \alpha X_i^2\right]$. It is known [35],[34] that an optimal linear coordination strategy exists, pertaining to "error" encoding and MMSE estimation decoding:

$$X_i \;\; = \;\; \beta_i \left(W_i - \mathbb{E}\left[W_i | Y^{i-1}\right]\right) \tag{4.37a}$$

$$Z_i \;\; = \;\; \mathbb{E}\left[W_i | Y^i\right] \tag{4.37b}$$

where β_i are time-varying normalizing constants that result in $X_i \sim \mathcal{N}(0, L)$ for all i, and the power-constraint L depends on the value of α.

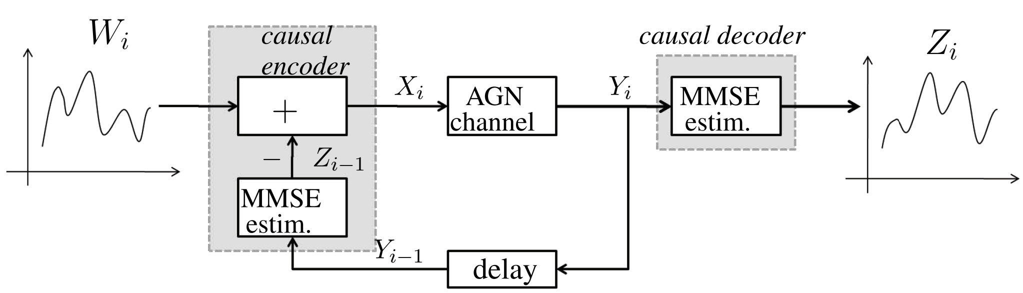

Figure 4.3: With Q_W Gauss-Markov and $P_{Y|X}$ an AGN channel, "error" encoding and MMSE estimation decoding is inverse control optimal. The induced cost function is squared error.

We now consider observing this problem from the lens of inverse optimal control for a distortion function of the form $\rho(w_i, z_{i-1}, z_i)$:

Lemma 4.4.1. *For the problem setup in* (4.35)*, define the following stationary Markov coordination policy:*

$$X_i = \beta \left(W_i - \rho Z_{i-1}\right) \tag{4.38a}$$

$$Z_i = \rho Z_{i-1} + \varrho Y_i \tag{4.38b}$$

where $\beta = \sqrt{\frac{L}{C}}, \gamma = \frac{\sqrt{LC}}{L+\sigma_v^2}$, *and* $C = \frac{\sigma_m^2}{1-\rho^2 \frac{\sigma_v^2}{L+\sigma_v^2}}$.

- *(a) The policy pair in* (4.38) *is inverse control optimal*

$$\eta(x_i) \propto_+ x_i^2 \tag{4.39a}$$

$$\rho(w_i, z_{i-1}, z_i) \propto_+ (w_i - z_i)^2 - \frac{\sigma_v^2}{L+\sigma_v^2}(w_i - z_{i-1})^2 \tag{4.39b}$$

- *(b) The total cost can be represented as a weighted MMSE cost given by* $J^{\alpha}_{n,\pi}$*:*

$$\begin{aligned} & \mathbb{E}\left[\sum_{i=1}^{n} \rho(W_i, Z_{i-1}, Z_i)\right] \\ = \ & \mathbb{E}\left[\sum_{i=1}^{n} (Z_i - W_i)^2 + \left(\frac{1}{1 - \frac{\sigma_v^2 \rho^2}{L+\sigma_v^2}}\right)(Z_n - W_n)^2\right] \end{aligned}$$

The proof is provided in Appendix H.

Remark 9. *The policy-pair in* (4.37) *is optimal for a mean-square distortion cost (MMSE) problem for Gauss-Markov sources except that the last reconstruction has higher penalty. For* $n \to \infty$*, the cost for which* (4.37) *is optimal is exactly equivalent to a MMSE cost problem* $\lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^{n} (Z_i - W_i)^2 + \alpha X_i^2\right]$. *Thus, asymptotically, we can recover the results of [35],[34, Ch. 6] using inverse-optimal control and time-invariant cost functions.*

4.4.2 Inverse Optimal Control: the M/M/1 Queue

Here we show that the $\cdot/M/1$ queue's dynamics can be interpreted as a stationary Markov coordination strategy that is inverse control optimal for Q_W being a

Poisson process. It is well-known from Burke's theorem [24, 29] that for a Poisson process of rate λ entering a $\cdot/M/1$ queue, in steady state the queue state at time t is independent of the output before time t. We now demonstrate that this statement has implications not only for the capacity of queuing timing channels [30, 31, 32, 69], but also for inverse optimal control.

See Figure 4.4. Divide time into units of interval Δ where $\Delta \ll 1$. The input to the queue W_i represents the number of arrivals to the queue till time i. For a Poisson source, $(W_i : i \geq 1)$ is the discrete-time equivalent of the counting function representation of a Poisson process.

$$Q_W\left(w_i | w_{i-1}\right) = \begin{cases} \lambda\Delta, & \text{if } w_i = w_{i-1} + 1 \\ 1 - \lambda\Delta, & \text{if } w_i = w_{i-1} \\ 0, & \text{otherwise.} \end{cases} \tag{4.40}$$

In other words, $W_i = W_{i-1} + \tilde{W}_i$ where $\tilde{W}_i$ are i.i.d. with $\mathbb{P}\left(\tilde{W}_i = 1\right) = \lambda\Delta$. Assume the following model for the channel:

$$P_{Y|X}(1|x) = \begin{cases} 0 & x = 0 \\ \mu\Delta & x > 0 \end{cases}. \tag{4.41}$$

For a queuing system, note that this means that a departure $(Y_i = 1)$ can only occur when the number of customers in the queue is positive, and the likelihood of a departure in that scenario for a bin of length Δ is $\mu\Delta$. Continuing on with the queuing analogy, note that we represent Z as the counting function representation of the departure process as $Z_i = \sum_{k \leq i} Y_k$ where $Y_k \in \{0, 1\}$. X_i is the queue size representing the number of customers in the queue at the i-th time instant: $X_i = W_i - Z_{i-1}$. Thus, the update equations for the state X_i and output of the queue Z_i are linear stationary Markov policies given by

$$X_i = W_i - Z_{i-1} \tag{4.42a}$$

$$Z_i = Z_{i-1} + Y_i \tag{4.42b}$$

The departure at i-th time instant Y_i depends on the state by the following discrete memoryless 'Z' channel model: That is, there will be no departure if the queue is empty, and there will be departure with probability $\mu\Delta$ if the queue is not empty. See Figure 4.5. The initial number of arrivals W_0 is drawn according to $\mathbb{P}(W_0 = k) = \left(1-\frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^k$, $k \geq 0$ and the initial number of departures $Z_0 = 0$. Note that the aggregate statistical dynamics of $P^{\bar{\gamma}}_{Z^n|W^n}$ in Figure 4.4 are

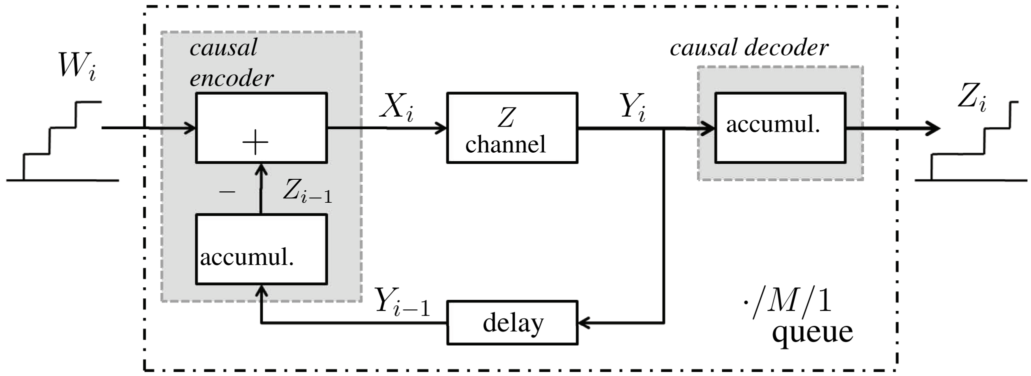

Figure 4.4: With Q_W a Poisson process and $P_{Y|X}$ a Z channel, the $\cdot/M/1$ queue is inverse control optimal.

precisely that of the discrete-time exponential server timing channel, also termed a $\cdot/M/1$ queue of rate μ, which is a first-come, first-serve queuing system with i.i.d. service times geometrically distributed of rate μ [24]. As $\Delta \to 0$, this becomes the continuous-time $\cdot/M/1$ queue. From standard queuing theory it follows that

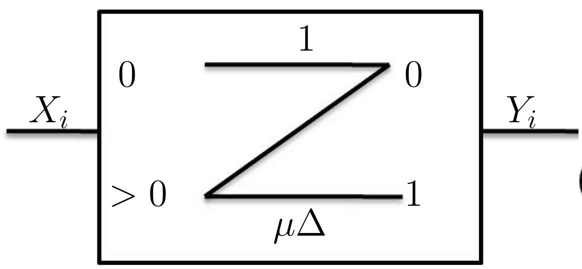

Figure 4.5: $P_{Y|X}$ for the $\cdot/M/1$ queue sampled at length-Δ intervals.

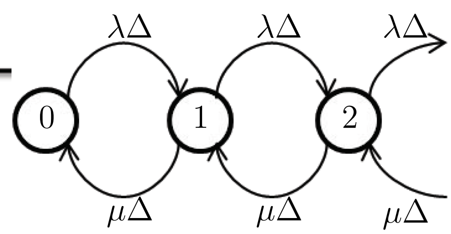

Figure 4.6: Birth-death chain for X in the $M/M/1$ queue.

X is a birth-death Markov chain (see Figure 4.6) in steady-state with distribution

$$\pi_k = \left(1-\frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^k, k \geq 0. \tag{4.43}$$

Therefore Lemma 4.2.2 holds and note that the fixed coordination strategy given by (4.42) is inverse-control optimal for

$$\rho(w_i, z_{i-1}, z_i) = \begin{cases} -\log(1-\lambda\Delta), & x_i = 0, y_i = 0; \\ \log\frac{1-\mu\Delta}{1-\lambda\Delta}, & x_i > 0, y_i = 0; \\ \log\frac{\mu}{\lambda}, & x_i > 0, y_i = 1; \\ +\infty, & \text{otherwise.} \end{cases}$$

$$\Rightarrow \lim_{\Delta\to 0} \rho(w_i, z_{i-1}, z_i) = \begin{cases} 0, & y_i = 0 \\ \log\frac{\mu}{\lambda}, & x_i > 0, y_i = 1 \\ \infty & \text{otherwise.} \end{cases} \tag{4.44}$$

Figure 4.4 is akin to [30, Fig 4], where it is shown that this insight and (4.44) leads to the derivation of the capacity of the exponential server timing channel.

Remark 10. *Though the ESTC is time-varying, non-memoryless, and has non-linear dynamics from a inter-arrival time viewpoint [30], when viewed appropriately, its internal structure consists of a time-invariant memoryless 'Z' channel and a feedback loop comprising a linear SM coordination strategy $\bar{\gamma}$. Moreover, for a Poisson process input, $\bar{\gamma}$ is inverse control optimal. As such, the internal structure of the $\cdot/M/1$ queue can be interpreted as an optimal decentralized controller. Also, note how the internal structure is exactly synonymous to the Gaussian case* (4.35) *([35],[34]) in that the encoder and decoder are both linear dynamical systems.*

The result differs from the source-channel matching results in [67, Sec 3] for two reasons: i) the problem is approached through an inter-arrival viewpoint in [67], while we use counting function representation (inputs and outputs to the queue). ii) The dynamics of the ESTC are fixed and [67] considers a possible encoder between the poisson process and the ESTC input, and a decoder between ESTC output and the reconstruction and shows that the encoder and decoder should be identity mappings. In our case, the linear encoder and decoder policies are fixed and internal to the structure of the queue dynamics. As a consequence, the source-channel matching results has to be performed over a less complicated memoryless 'Z' channel.

Other extensions to queuing timing channels fit within this framework as well: see for example the variety of queuing systems in [68] for which joint reversibility holds. Similar results hold for other queuing timing channels, such as:

- $\cdot/M/c$ queue: There are c servers each with an i.i.d exponential service time. In this case, the queue dynamics-the linear encoder and the decoder will be the same (Fig 4.4). The structure of memoryless channel ($P_{Y|X}$) will depend on c.

- 'The queue with feedback' [24, p 204-205]. Here, with probability $1 - p_0$ departures from the queue instantaneously return to the input of the queue (independent of all other processes). The 'effective' Z channel changes $\mu\Delta$ to $p_0\mu\Delta$ and all other arguments hold.

4.4.3 Inverse Optimal Control: Blackwell's Trapdoor Channel

Here we show that the internal structure of Blackwell's trapdoor channel can be interpreted as a stationary Markov coordination strategy that is inverse control optimal.

Consider 'the chemical (trapdoor) channel' [36, 37, 38] as shown in Fig 4.7. Initially (Figure 4.7a), a ball labeled either 0 (red) or 1 (blue) is present in one of the two slots. Then (Figure 4.7b) a ball, either a 0 or 1, is placed in the empty slot, after which (Figure 4.7c) one of the trapdoors opens at random with probability $(\frac{1}{2}, \frac{1}{2})$. The ball lying above the open door then falls through. The door closes (as in Figure 4.7a) and the process is repeated.

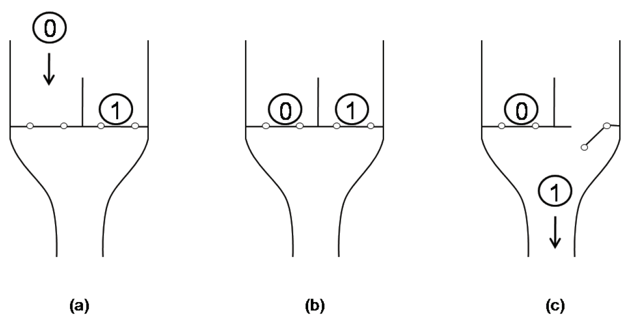

Figure 4.7: Blackwell's trapdoor channel.

Let $\tilde{W}_i \in \{0, 1\}$ and $Y_i \in \{0, 1\}$ represent the color of the ball that is input and output of the trapdoor respectively. Define the channel input X_i to pertain to the composition of balls before one of the doors is opened (Fig 4.7b). That is, $X_i \in \{0, 1, 2\}$ where $X_i = 0$ represents two red balls (0,0), $X_i = 1$ represents a blue ball and a red ball (0,1) and $X_i = 2$ represents two blue balls (1,1). Thus, the dynamics are given by

$$X_i = X_{i-1} + \widetilde{W}_i - Y_{i-1}. \tag{4.45}$$

From a counting function viewpoint, let $\{W_i\}$ and $\{Z_i\}$ be the counting processes representing the number of blue balls that were input and output from the system. See Figure 4.8. Hence X_i, as defined above tells about the composition of balls, or equivalently the number of blue balls that are 'in' the system at time i.

$$\begin{aligned} W_i &= W_{i-1} + \widetilde{W}_i && (4.46a)\\ Z_i &= Z_{i-1} + Y_i && (4.46b)\\ X_i &= X_{i-1} + \widetilde{W}_i - Y_{i-1} \\ &= W_i - Z_{i-1}. && (4.46c) \end{aligned}$$

Note that the state-update equation and the decoding policy (4.46b)-(4.46c) are reversibly feasible dynamics by Example 3. The output depends on the state according to the channel law $P_{Y|X}(Y|X)$ as (the inverse erasure channel) as shown in Figure 4.9:

$$P_{Y|X}(1|x) = \begin{cases} 0 & x = 0 \\ \frac{1}{2} & x = 1 \\ 1 & x = 2 \end{cases}. \tag{4.47}$$

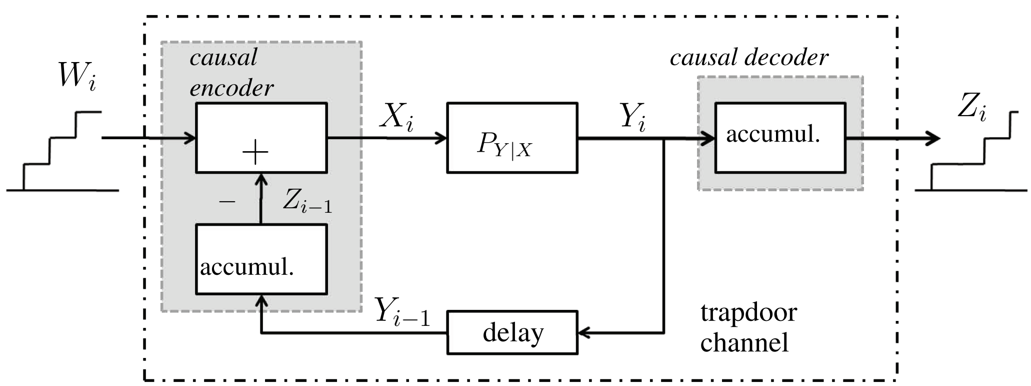

Figure 4.8: With Q_W a Markov counting process (i.i.d. $\tilde{W}_i$ inputs) and an 'inverted E' channel, Blackwell's trapdoor channel is inverse control optimal.

Fixing $\tilde{W}$ to be an i.i.d process, with $\mathbb{P}\left(\tilde{W}_i = 0\right) = p.$, and $Z_0 = 0$. The

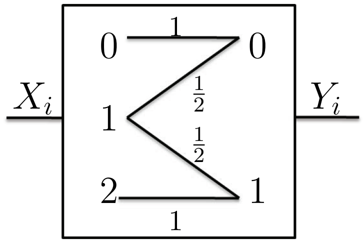

Figure 4.9: $P_{Y|X}$ for the trapdoor channel.

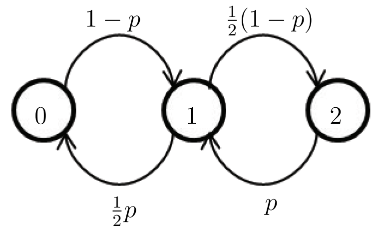

Figure 4.10: Birth-death chain for X in the trapdoor channel with Q_W a Markov counting process.

transition probabilities of the Markov chain X_i are given by Fig 4.10.

$$P = \begin{bmatrix} p & 1-p & 0 \\ \frac{1}{2}p & \frac{1}{2} & \frac{1}{2}(1-p) \\ 0 & p & 1-p \end{bmatrix}$$

And if W_0 is drawn according to the initial distribution given by

$$\mathbb{P}\left(W_0 = k\right) = \begin{cases} p^2, & \text{k=0;} \\ 2p(1-p), & \text{k=1;} \\ (1-p)^2, & \text{k=2;} \\ 0, & \text{otherwise} \end{cases},$$

then it follows that we have a birth-death chain initially in steady-state with distribution $\pi(\cdot) = \mathbb{P}\left(W_0 = \cdot\right)$. Thus from Lemma 4.2.2, we have that $\bar{\pi}$ is inverse-control optimal. Moreover, the from Corollary 4.1.6, the trapdoor policy (4.45) is optimal for the cost function of the form

$$\rho(w_i, z_{i-1}, z_i) = \begin{cases} \log p, & x_i = 0, y_i = 0; \\ \log 2p, & x_i = 1, y_i = 0; \\ \log 2(1-p), & x_i = 1, y_i = 1; \\ \log(1-p), & x_i = 2, y_i = 1; \\ +\infty, & \text{otherwise.} \end{cases}$$

Note that when $p = \frac{1}{2}$, $\mathbb{E}\left[\rho(W, Z_{i-1}, Z_i)\right] = -I(\pi, P_{Y|X}) = -\frac{1}{2}$, which coincides with the achievable rate coding scheme developed for the trapdoor channel in [37].

PART II

RELIABILITY OF COMMUNICATION POLICIES

CHAPTER 5

THE RELATIONSHIP BETWEEN RELIABLE FEEDBACK COMMUNICATION AND NONLINEAR FILTER STABILITY

In this chapter, we further demonstrate an interplay between information theory and control theory, at the level of achievability of message-point communication schemes. We establish necessary and sufficient conditions for when a message point feedback communication - the posterior matching scheme - achieves reliable communication in terms of the stability of the posterior belief's nonlinear filter. By making this connection to hidden Markov models, we can provide easily testable sufficient conditions (e.g. ergodicity of the Markov process and non-degeneracy of the noisy channel) on when reliable communication occurs.

5.1 Introduction

We look at a specific instance of the causal coding-decoding problem introduced in Sec 2.5 corresponding to sequential information gain cost and when $W_i \equiv W$ (See example 3.3.1). This formulation was motivated from Shannon's converse and is the message-point communication system with feedback introduced in Sec 2.4. See Figure 5.1. An explicit optimal solution for when $\mathsf{W} = [0, 1]$ line was shown in Chapter 3.3.1 to be the posterior matching scheme (Section 2.4.3). However, it is unclear if this will indeed guarantee 'reliability', or if there are concise sufficient conditions when one can do so. It is the purpose of this chapter to delve into this further. In the process, because the message point W is continuous, we provide a very different mathematical framework to give necessary and sufficient conditions. We show a fundamental relationship with the stability of the nonlinear filter - thus further showing an intimate relationship between information theory and control theory - particularly when discretization is not fundamental to the problem formulation.

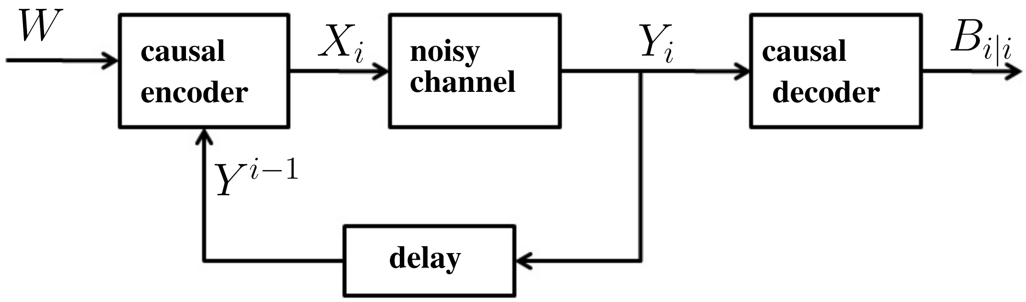

Figure 5.1: Communication of a message point W with causal feedback over a memoryless channel.

5.1.1 Posterior Matching Scheme - Relevant Properties

The PM scheme in many cases achieves capacity, and its 'simple' encoding structure allows it to be used beyond standard communications paradigms [21]. In other cases, however, the PM scheme cannot achieve any positive rate [16, Example 11]. In this section, we focus on the 'PM-style' encoding scheme described in Section 2.4.3 and derive the following results:

Lemma 5.1.1. *Under the PM-style encoding scheme:*

(i) Y_i*'s are i.i.d.*

(ii) For a non-degenerate channel, $I(W; Y_{n+1}|\mathcal{F}^Y_{1,n}) = I$*. And* $I(W; Y^n) = nI$*.*

(iii) For all $n \geq 1$*:* $\sigma(\tilde{W}_{n+1}) \vee \mathcal{F}^Y_{1,n} \supset \sigma(W)$*.*

(iv) $\{\tilde{W}_i\}_{i\geq 1}$ *is a time-homogenous Markov chain with the transition probability kernel such that for any* $A \in \mathcal{B}(\mathsf{W})$*,*

$$\varpi_W(\tilde{W}_i, A) \triangleq \mathbb{P}\left(\tilde{W}_{i+1} \in A|\sigma(\tilde{W}_i)\right) = \int_{y\in\mathsf{Y}} P_{Y|X}(y|\phi(\tilde{W}_i))1_{\{T_y(\tilde{W}_i)\in A\}}dy. \tag{5.1}$$

(v) Define

$$S_n = (\tilde{W}_n, Y_n) \qquad \textit{such that} \quad S_n[1] = \tilde{W}_n \textit{ and } S_n[2] = Y_n. \tag{5.2}$$

$\{S_n, Y_n\}_{n\geq 0}$ *forms a Hidden Markov Model* (2.16c) *with transition law and* *observation law given by*

$$\xi_S(S_n,(A,B)) = 1_{\left\{T_{S_n[2]}(S_n[1])\in A\right\}} \times P_{Y|\tilde{W}}(B|S(S_n[1],S_n[2]) \tag{5.3a}$$

$$P_{Y|S}(B|S_n) = 1_{\{S_n[2]\in B\}}. \tag{5.3b}$$

(vi) Define $S_n = (\tilde{W}_n, Y_n)$ *as in* (5.2)*.* $\{S_n\}_{n\geq 0}$ *is ergodic iff* $\{\tilde{W}_n\}_{n\geq 0}$ *is ergodic.*

Proof. For (i), from Property 2.4.5-(a) and (2.44b), $X_i \perp\!\!\!\perp Y^{i-1}$. Further, from the memoryless property of the channel $P_{Y|X}$ (2.51), $Y_i \perp\!\!\!\perp Y^{i-1}$.

For (ii), using Property 2.4.5-(b) and because Y_i are i.i.d:

$$\begin{aligned} I(W;Y_{n+1}|\mathcal{F}^Y_{1,n}) &= H(Y_{n+1}|\mathcal{F}^Y_{1,n}) - H(Y_{n+1}|\mathcal{F}^Y_{1,n}\vee\sigma(W)) &(5.4a)\\ &= H(Y_{n+1}) - H(Y_{n+1}|X_{n+1}) &(5.4b)\\ &= I(X_{n+1};Y_{n+1}) = I. &(5.4c)\end{aligned}$$

Hence,

$$I(W;Y^n) = \sum_{i=1}^{n} I(W;Y_i|\mathcal{F}^Y_{1,i-1}) = nI. \tag{5.4d}$$

For (iii), because the mapping S_y is invertible from Property 2.4.5-(c), we can recover W_0 from $\tilde{W}_{n+1}$ and Y^n.

$$\tilde{W}_{n+1} = S_{Y_n}(\tilde{W}_n) = S_{Y_n}\circ S_{Y_{n-1}}\circ\cdots\circ S_{Y_1}(W) \tag{5.5}$$

Hence, $\sigma(\tilde{W}_{n+1})\vee\mathcal{F}^Y_{1,n}\supset\sigma(W)$.

For (iv), because Y's are i.i.d., the above equation (5.5) corresponds to an iterated function system, and $(\tilde{W}_i)_{i\geq 1}$ is a time-homogenous Markov Chain [64]. For (v), $\{S_n, Y_n\}_{n\geq 0}$ is in fact a HMM because

- the observation process Y_n is a component of S_n. $Y_n = S_n[2]$ thus satisfying (5.3b).
- $\tilde{W}_{n+1}$ is a deterministic function of $S_n = (\tilde{W}_n, Y_n)$ for the 'PM-style' scheme and Y_{n+1} depends only on $\tilde{W}_{n+1}$ (and hence $S_n = (\tilde{W}_n, Y_n)$) and the channel law $P_{Y|\tilde{W}}$- thus satisfying (5.3a).

For (vi), because Y's are i.i.d., the second component of S_n is always ergodic. Hence the process $\{S_n\}_{n\geq 0}$ is ergodic if and only if its first component $\{\tilde{W}_n\}_{n\geq 0}$ is ergodic.

$\square$

Assumption 1. *(*ERGODICITY*) A process $\{S_n\}_{n\geq 0}$ starting at $S_0 = z$ is said to be ergodic, if there exists a steady state probability measure ϖ_S such that the following holds:*

$$||\mathbb{P}^z(S_n \in \cdot) - \varpi_S||_{TV} \overset{n\to\infty}{\longrightarrow} 0 \qquad \textit{for } \varpi - a.e. z \in \mathsf{U}.$$

Assumption 2. *(*NON-DEGENERACY*) There exists a probability measure φ on $\mathcal{B}(\mathsf{Y})$ and a strictly positive measurable function $g : \mathsf{W} \times \mathsf{Y} \to (0, \infty)$ such that*

$$P_{Y|\tilde{W}}(A|z) = \int 1_{\{u\in A\}} g(z, u)\varphi(du) \qquad \forall A \in \mathcal{B}(\mathsf{Y}), z \in \mathsf{W}.$$

We know from Lemma 2.3.5 that there exists an equivalent conditional Markov process $(P^S, \mu, \mathbf{P}^Y)$ for every hidden Markov process $(\xi_S, P_{Y|S})$ with the signal process $\{S_n\}$ and the observation process $\{Y_n\}$. We will now define a class of HMMs for which a certain continuity condition exists.

Definition 5.1.2. *(*CONDITIONAL MUTUAL ABSOLUTE CONTINUITY*) We say that a hidden Markov model (S, Y) with the generative law $(\xi_S, P_{Y|S})$ is 'conditional mutual absolute continuous' if*

- *the signal process S_n is ergodic, and*
- *there is a strictly positive measurable function $h : \mathsf{U} \times \Omega^Y \times \mathsf{U} \to (0, \infty)$ such that for $\mu\mathbf{P}^Y$ -a.e. (z, y),*

$$P^S(z, y, A) = \int 1_{\{\tilde{z}\in A\}} h(z, y, \tilde{z})\xi_S(z, d\tilde{z}) \qquad \forall A \in \mathcal{B}(\mathsf{U}) \qquad (5.6)$$

where (P^S, μ, P^Y) corresponds to the generative law of a conditional Markov process constructed according to (2.23).

5.2 Main Results

5.2.1 Necessary and Sufficient Conditions on Reliability

We know from the definition of reliability, Lemma 2.4.3, that a feedback encoding scheme is reliable iff (2.35) holds. We now derive an equivalent necessary and sufficient condition that begins to resemble the type of conditions required for nonlinear filter stability [52, eqn 1.10]:

Theorem 5.2.1. *A feedback communication scheme is reliable iff*

$$\overline{\mathbb{E}}\left[\frac{d\nu^{k,u}}{d\overline{\nu}}(W)|\mathcal{F}^{Y}_{1,\infty}\right]\Big|_{u=W} = 2^k, \quad \forall k \geq 1 \tag{5.7}$$

where $\pi_0^{k,W} = \nu^{k,W}$ *and* $\bar{\pi}_0 = \overline{\nu}$ *are defined in* (2.31) *and* (2.32)*. Since the L.H.S in* (5.7) *is* $\sigma(W) \vee \mathcal{F}^{Y}_{1,\infty}$*-measurable, the following condition will suffice for* (5.7) *to hold*

$$\mathcal{F}^{Y}_{1,\infty} \vee \sigma(W) = \mathcal{F}^{Y}_{1,\infty} \quad \overline{\mathbb{P}} - a.s. \tag{5.8}$$

The proof can be found in Section 5.3. Equation (5.8) implies that communication is reliable if the message point W can be inferred from the infinite observation sequence $\{Y_n\}_{n\geq 1}$. The condition (5.7) relates to the condition on nonlinear filter stability given in [52, eqn 1.10] but not exactly - note the absence of tail sigma fields in (5.7).

The PM-style feedback encoding scheme developed in (2.44) is of interest because it is optimal, $I(W;Y^n) = nI$. As mentioned before, the PM scheme in many cases achieves capacity, but in others does not achieve any positive rate [16, Ex. 11]. Sufficient conditions based upon the kernel being 'fixed-point-free' are given in [16, Thm. 4]. It is our objective to develop general **necessary and sufficient** conditions - illustrated through the lens of hidden Markov model filter stability - under which all rates $R < C$ are achievable for the PM scheme. This leads to the following:

Lemma 5.2.2. *The PM scheme is reliable iff*

$$\overline{\mathbb{E}}\left[\frac{d\nu^{k,u}}{d\overline{\nu}}(W)\Big|\bigcap_{n\geq 1}\mathcal{F}^{Y}_{1,\infty}\vee\mathcal{F}^{\tilde{W}}_{n,\infty}\right]\Bigg|_{u=W} = \overline{\mathbb{E}}\left[\frac{d\nu^{k,u}}{d\overline{\nu}}(W)\Big|\mathcal{F}^{Y}_{1,\infty}\right]\Bigg|_{u=W}. \tag{5.9}$$

The proof can be found in Section 5.3. Lemma 5.2.2 can be proved in two different ways. The first approach combines the result from Lemma 5.2.1 and the property Lemma 5.1.1-(iii) of PM scheme.

Remark 11. *The condition* (5.9) *on PM scheme is intriguing because this has exactly the same mathematical structure as the condition required for nonlinear filter stability of HMMs recently shown in [52, eqn 1.10].*

Directly from this connection to filter stability in HMMs, we can now develop sufficient conditions under which the PM scheme is reliable:

Theorem 5.2.3. *If $\{S_n\}_{n\geq 0}$ is ergodic such that $\tilde{W}_n$ is ergodic and the channel law $P_{Y|\tilde{W}}$ is non-degenerate, then the PM scheme is reliable.*

Proof. See Section 5.4. □

Remark 12. *These conditions relate to the example Shayevitz and Feder gave in [16, Ex. 11], where no positive rate was achievable: there, the process $(\tilde{W}_i)_{i\geq 1}$ was non-ergodic. The aforementioned Lemma uses the concept of filter stability to provide sufficient conditions on reliability. Some conditions relate to ergodicity (provided non-degeneracy of the channel), thus synergizing with [16, Ex. 11]. Eliminating non-degeneracy cannot in general be done to ensure* (5.9) *in HMMs (see [52, Ex 1.1]).*

5.2.2 Necessary and Sufficient Conditions for Achieving All Rates $R < C$

Lastly, we state our main result in this paper:

Theorem 5.2.4. *For a discrete and non-degenerate memoryless channel, the PM scheme achieves all rates $R < C$ if and only if it is reliable.*

The proof can be found in Section 5.5.

Corollary 5.2.5. *For a discrete and non-degenerate memoryless channel, if the channel inputs $\tilde{W}$ are ergodic and the channel law $P_{Y|X}$ is non-degenerate, then the PM scheme achieves all rates $R < C$.*

Proof. Using Theorem 5.2.3 and 5.2.4. □

5.3 Proofs of Reliability Theorems

Proof of Theorem 5.2.1. By the definition of KL-divergence, $D\left(\pi_n^{k,W} \| \bar{\pi}_n\right) = \mathbb{E}\left[\log \frac{d\pi_n^{k,W}}{d\bar{\pi}_n}(W) \Big| \mathcal{F}_{1,n}^Y\right]$. Since $\pi_0^{k,W} = \nu^{k,W}$ is a random measure (of W), we first compute the KL-divergence for fixed priors.

From (2.29), for any *fixed* measures ν and $\bar{\nu}$ s.t. $\nu \ll \bar{\nu}$, $\pi_0 = \nu$ and $\bar{\pi}_0 = \bar{\nu}$, $\frac{d\mathbb{P}}{d\bar{\mathbb{P}}} = \frac{d\nu}{d\bar{\nu}}(W)$. From Bayes' rule,

$$
\begin{aligned}
\mathbb{E}\left[g(W) | \mathcal{F}_{1,n}^Y\right] &= \frac{\bar{\mathbb{E}}\left[g(W)\frac{d\mathbb{P}}{d\bar{\mathbb{P}}} \Big| \mathcal{F}_{1,n}^Y\right]}{\bar{\mathbb{E}}\left[\frac{d\mathbb{P}}{d\bar{\mathbb{P}}} \Big| \mathcal{F}_{1,n}^Y\right]} \\
&= \int_{\mathsf{W}} g(w) \frac{\frac{d\nu}{d\bar{\nu}}(w)}{\bar{\mathbb{E}}\left[\frac{d\nu}{d\bar{\nu}}(W) \Big| \mathcal{F}_{1,n}^Y\right]} \bar{\pi}_n(dw).
\end{aligned}
$$

Therefore,

$$
\frac{d\pi_n}{d\bar{\pi}_n}(W) = \frac{\frac{d\nu}{d\bar{\nu}}(W)}{\bar{\mathbb{E}}\left[\frac{d\nu}{d\bar{\nu}}(W) \Big| \mathcal{F}_{1,n}^Y\right]}. \tag{5.10}
$$

Hence, for fixed measures $\nu^{k,u}$ and $\bar{\nu}$, from (5.10)

$$
\begin{aligned}
D\left(\pi_n^{k,u} \| \bar{\pi}_n\right) &= \mathbb{E}\left[\log \frac{d\pi_n^{k,u}}{d\bar{\pi}_n}(W) \Big| \mathcal{F}_{1,n}^Y\right] \\
&= k - \log \bar{\mathbb{E}}\left[\frac{d\nu^{k,u}}{d\bar{\nu}}(W) \Big| \mathcal{F}_{1,n}^Y\right]. \qquad (5.11)
\end{aligned}
$$

where (5.11) is because $\frac{d\nu^{k,u}}{d\bar{\nu}}(W) = 2^k 1_{\{Q_k(u) = Q_k(W)\}}$. Moreover,

$$
\begin{aligned}
k - \log \bar{\mathbb{E}}\left[\frac{d\nu^{k,u}}{d\bar{\nu}}(W) \Big| \mathcal{F}_{1,n}^Y\right] &= k - \log \int_{w \in \mathsf{W}} 2^k 1_{\{Q_k(w) = Q_k(u)\}} \bar{\pi}_n(dw) \qquad (5.12) \\
&= -\log \bar{\pi}_n(\{l : Q_k(l) = Q_k(u)\}) \qquad (5.13) \\
&= D\left(\pi_{n|\mathcal{G}_k}^{k,u} \| \bar{\pi}_{n|\mathcal{G}_k}\right) \qquad (5.14)
\end{aligned}
$$

where (5.12) is because $\frac{d\nu^{k,u}}{d\overline{\nu}}(W) = 2^k 1_{\{Q_k(u)=Q_k(W)\}}$. (5.14) follows from (2.42). Hence,

$$\lim_{n\to\infty}\left[k - \log \overline{\mathbb{E}}\left[\frac{d\nu^{k,u}}{d\overline{\nu}}(W)\Big|\mathcal{F}^Y_{1,n}\right]\right]\Big|_{u=W} = \lim_{n\to\infty} D\left(\pi^{k,u}_{n|\mathcal{G}_k}\|\bar{\pi}_{n|\mathcal{G}_k}\right)\Big|_{u=W}. \tag{5.15}$$

Using Lemma 2.4.3, if the communication is reliable then the R.H.S in (5.15) equals 0 and hence (5.7) of Theorem 5.2.1 holds. □

Proof of Lemma 5.2.2. In the first approach, we re-visit (5.10)-(5.13). From Definition 2.4.1,

$$\begin{aligned} 2^k 1_{\{Q_k(u)=Q_k(W)\}} &= \frac{d\nu^{k,u}}{d\overline{\nu}}(W) \\ &= \overline{\mathbb{E}}\left[\frac{d\nu^{k,u}}{d\overline{\nu}}(W)\Big|\mathcal{F}^Y_{1,\infty}\vee\sigma(W),\right] \quad (5.16)\\ &= \overline{\mathbb{E}}\left[\frac{d\nu^{k,u}}{d\overline{\nu}}(W)\Big|\bigcap_{n\geq 1}\mathcal{F}^Y_{1,\infty}\vee\mathcal{F}^{\tilde{W}}_{n,\infty}\right]. \quad (5.17)\end{aligned}$$

(5.16) is because $\frac{d\nu^{k,u}}{d\overline{\nu}}(W)$ is $\sigma(W)$-measurable; (5.17) from Lemma 5.1.1(iii). From here, simply invoke Lemma 5.2.1. □

5.4 Sufficiency Condition for Reliability of PM Scheme - Proof of Theorem 5.2.3

Define $S_n = (\tilde{W}_n, Y_n)$ as in (5.2). Denote $S_n[1] = \tilde{W}_n$ and $S_n[2] = Y_n$. Before we prove Theorem 5.2.3, we prove the following sub-theorem:

Lemma 5.4.1. *If a HMM $(\xi_S, P_{Y|S})$ with signal process $S_n = (\tilde{W}_n, Y_n)$ is such that S_n is ergodic, and $P_{Y|\tilde{W}}$ is non-degenerate, then it is 'conditional mutual absolute continuous' as defined in Defn 5.1.2.*

Proof. It thus suffices to show the existence of a strictly positive measurable function $h : \mathsf{U}\times\Omega^Y\times\mathsf{U}\to(0,\infty)$ such that for $\mu\mathbf{P}^Y$ -a.e. (z,y),

$$P^S(z,y,A) = \int 1_{\{\tilde{z}\in A\}} h(z,y,\tilde{z})\xi_S(z,d\tilde{z}) \qquad \forall A\in\mathcal{B}(\mathsf{U}) \tag{5.18}$$

where (P^S, μ, P^Y) corresponds to the generative law of a conditional Markov process constructed according to (2.23).

Existence of strictly positive measurable function h: Invoke Lemma 5.4.2 and Lemma 5.4.3 to show the existence of strictly positive measurable function h.

Lemma 5.4.2. *Let ν and $\overline{\nu}$ be two priors on $\tilde{W}_0$ such that $||\mathbf{P}^{\nu}(\tilde{W}_n \in \cdot) - \mathbf{P}^{\overline{\nu}}(\tilde{W}_n \in \cdot)||_{TV} \stackrel{n\to\infty}{\longrightarrow} 0$ and the kernel $P_{Y|\tilde{W}}$ is non-degenerate (Defn 2). Then $\mathbf{P}^{\nu}\Big|_{\mathcal{F}_+^Y} \sim \mathbf{P}^{\overline{\nu}}\Big|_{\mathcal{F}_+^Y}$.*

Proof. See Appendix I. $\square$

Lemma 5.4.3. *Suppose S_n is ergodic and $\mathbf{P}^{\nu}\Big|_{\mathcal{F}_+^Y} \sim \mathbf{P}^{\overline{\nu}}\Big|_{\mathcal{F}_+^Y}$, then there exists a strictly positive measurable function h satisfying*

$$P^S(z, y, A) = \int 1_{\{\tilde{z}\in A\}} h(z, y, \tilde{z}) \xi_S(z, d\tilde{z}) \qquad \forall A \in \mathcal{B}(\mathsf{U}). \tag{5.19}$$

Proof. The proof follows directly from [1, Lemma 3.8] except for slight differences. We present the complete proof in Appendix J. $\square$

$\square$

5.4.1 Proof of Theorem 5.2.3

Proof. See Figure 5.2. We first start by showing that for a posterior matching scheme $\{S_n, Y_n\}_{n\in\mathbb{Z}}$ forms a hidden Markov model. We next point out that this hidden Markov process is in fact a disguised Markov chain in a random environment (or equivalently $\{S_n\}$ is a conditional signal process). Further, under ergodicity and non-degeneracy assumptions (Assumption 1 and 2), the HMM is in fact *conditional mutual absolute continuous*. Finally, if the HMM is 'conditional mutual absolute continuous', from [1], it can be shown that the condition for PM scheme reliability (5.9) holds.

- When using PM Scheme, from Lemma 5.1.1(v) $\{S_n, Y_n\}_{n\in\mathbb{Z}}$ forms a hidden Markov model with $(\xi_S, P_{Y|X})_{PM}$ given by (5.3).
- For the PM Scheme if $\{\tilde{W}_n\}$ is ergodic, from Lemma 5.1.1(vi), $\{S_n\}$ is ergodic.

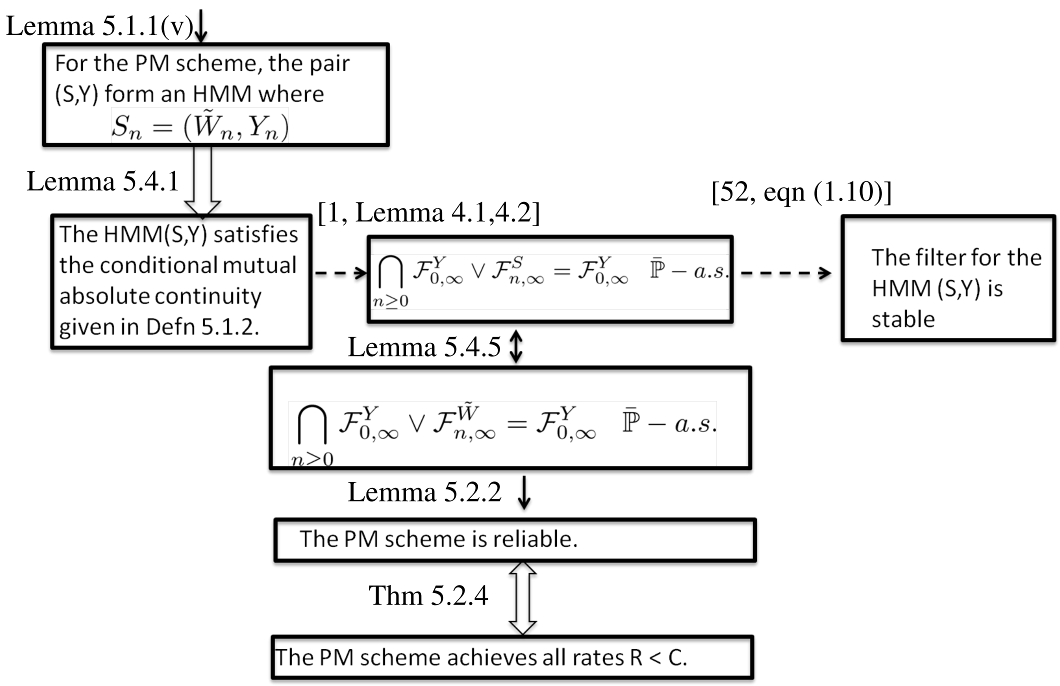

Figure 5.2: Proof outline of results corresponding to posterior matching scheme (Lemma 5.2.2, Theorem 5.2.3, Theorem 5.2.4). Arrows: Filled in arrows represent no assumptions needed. Hollow arrows represent ergodicity of $\tilde{W}$ and nondegeneracy of $P_{Y|\tilde{W}}$ needed. Solid arrow outlines represent results that we proved. Dashed arrow outline represent results from [1].

- From the above steps, consider the HMM $\{S_n, Y_n\}$ such that $\{S_n\}$ is ergodic. Invoke Lemma 5.4.1 to see that the HMM $\{S_n, Y_n\}$ is *conditional mutual absolute continuous*.
- If the HMM $\{S_n, Y_n\}$ is *conditional mutual absolute continuous*, using Van Handel's results [1], it can be shown that a certain stability condition holds.

Lemma 5.4.4. *If the HMM $\{S_n, Y_n\}$ is conditional mutual absolute continuous, then*

$$\bigcap_{n\geq 0} \mathcal{F}^S_{n,\infty} \vee \mathcal{F}^Y_{0,\infty} = \mathcal{F}^Y_{0,\infty} \qquad \bar{\mathbb{P}} - a.s. \tag{5.20}$$

Proof. From [1, Theorem 3.5, 4.1 and 4.2],

$$\bigcap_{n\geq 0} \mathcal{F}^S_{n,\infty} \vee \mathcal{F}^Y_{0,\infty} = \mathcal{F}^Y_{0,\infty} \qquad \overline{\mathbb{P}} - a.s.$$

□

- Invoking Lemma 5.4.5, we can show that infact $\mathcal{F}^{\tilde{W}}_{n,\infty} \vee \mathcal{F}^Y_{0,\infty} = \mathcal{F}^Y_{0,\infty}$, thus satisfying the reliability condition (5.9). This completes the proof.

Lemma 5.4.5. $\mathcal{F}^S_{n,\infty} \vee \mathcal{F}^Y_{0,\infty} = \mathcal{F}^{\tilde{W}}_{n,\infty} \vee \mathcal{F}^Y_{0,\infty} \qquad \overline{\mathbb{P}} - a.s.$

Proof.

$$\begin{aligned}
\bigcap_{n\geq 0} \mathcal{F}^{\tilde{W}}_{n,\infty} \vee \mathcal{F}^Y_{0,\infty} &= \bigcap_{n\geq 0} \sigma(\tilde{W}^\infty_n, Y^\infty_0) && (5.21)\\
&= \bigcap_{n\geq 0} \sigma((\tilde{W}^\infty_n, Y^\infty_n), Y^\infty_0) && (5.22)\\
&= \bigcap_{n\geq 0} \sigma(S^\infty_n, Y^\infty_0) && (5.23)\\
&= \mathcal{F}^Y_{0,\infty} \qquad \overline{\mathbb{P}} - a.s. && (5.24)
\end{aligned}$$

□

□

5.5 Necessary and Sufficient Condition for Achieving Any Rate $R < C$ - Proof of Theorem 5.2.4

We now prove our main result, Theorem 5.2.4: the PM scheme can achieve any rate $R < C$ iff it is reliable. Since (⇒) is trivial, we now focus on (⇐). We first develop some machinery. Define the random variables Z_n, Z'_n and G_n as

$$Z_n \triangleq \frac{d\bar{\pi}_{n+1}}{d\bar{\pi}_n}(W), \qquad Z'_n \triangleq \frac{d\bar{\pi}_{n+1|\mathcal{G}_{nR}}}{d\bar{\pi}_{n|\mathcal{G}_{nR}}}(W). \tag{5.25}$$

$$\mathcal{B}_n \triangleq \mathcal{G}_\infty \vee \mathcal{F}^Y_{1,n+1}, \qquad \mathcal{B}'_n \triangleq \mathcal{G}_{nR} \vee \mathcal{F}^Y_{1,n+1} \tag{5.26}$$

$$G_n \triangleq \overline{\mathbb{E}}\left[\log Z_n - \log Z'_n | \mathcal{F}^Y_{1,n}\right]. \tag{5.27}$$

Proof of Thm 5.2.4. The proof is outlined as follows:

1. Define the candidate Lyapunov function $V_n(b) : \mathcal{P}([0,1]) \to \mathbb{R}^+$ as

$$\begin{aligned} V_n(b) \triangleq& D\left(\pi^{nR,W}_{n|\mathcal{G}_{nR}} \| b_{|\mathcal{G}_{nR}}\right) \\ =& -\log b\left(\{l : Q_{nR}(l) = Q_{nR}(W)\}\right). \end{aligned} \tag{5.28}$$

2. Define $A_{n,\epsilon} \triangleq \{\omega : G_n(\omega) \le \epsilon\}$. Invoke Lemma 5.5.1.

   Lemma 5.5.1. *If $\omega \in A_{n,\epsilon}$ and $V_{n+1}(\bar{\pi}_n) > 0$, then the drift of the Lyapunov function is negative:*

$$\overline{\mathbb{E}}\left[V_{n+1}(\bar{\pi}_{n+1}) - V_n(\bar{\pi}_n) | \mathcal{F}^Y_{1,n}\right](\omega) \le \begin{cases} -(C - R - \epsilon) 1_{\{V_{n+1}(\bar{\pi}_n) > 0\}} & \omega \in A_{n,\epsilon} \\ R 1_{\{V_{n+1}(\bar{\pi}_n) > 0\}} & \textit{otherwise.} \end{cases} \tag{5.29}$$

3. Define $B_{n,\epsilon} \triangleq \bigcap_{k \ge n} A_{k,\epsilon}$. Invoke Lemma 5.5.2.

   Lemma 5.5.2. *If $\mathcal{F}^Y_{1,\infty} = \mathcal{F}^Y_{1,\infty} \vee \sigma(W)$ $\overline{\mathbb{P}} - a.s.$, then for any $\epsilon > 0$ there exists $N(\epsilon)$ s.t.*

$$\overline{\mathbb{P}}(B_{n,\epsilon}) = 1, \qquad n \ge N(\epsilon).$$

4. Invoke Lemma 5.5.3.

   Lemma 5.5.3. *If $\mathcal{F}^Y_{1,\infty} = \mathcal{F}^Y_{1,\infty} \vee \sigma(W)$ $\overline{\mathbb{P}} - a.s.$, then*

$$\lim_{M \to \infty} \overline{\mathbb{P}}(V_{n+M+1}(\bar{\pi}_{n+M}) > 0) = 0. \tag{5.30}$$

Thus, $V_{n+1}(\bar{\pi}_n) \to 0$ in $\overline{\mathbb{P}}$-probability. Since $V_n(\bar{\pi}_n) \leq V_{n+1}(\bar{\pi}_n)$, this implies that $V_n(\bar{\pi}_n) \to 0$ in $\overline{\mathbb{P}}$-probability. So from Lemma 2.4.3, any rate $R < C$ is achievable. $\square$

Before proving Lemma 5.5.1, 5.5.2 and 5.5.3, we develop the following machinery. We exploit the following properties of Z_n and G_n:

Lemma 5.5.4. *Z_n and Z'_n have the following properties*

(i) Z_n *is* $\mathcal{B}_n$*-measurable*

(ii) Z'_n *is* $\mathcal{B}'_n$*-measurable*

(iii) $Z'_n = \overline{\mathbb{E}}\left[Z_n|\mathcal{B}'_n\right]$

(iv) $D\left(\bar{\pi}_{n+1}\|\bar{\pi}_n\right) = \overline{\mathbb{E}}\left[\log Z_n|\mathcal{F}^Y_{1,n+1}\right]$, *and*
$D\left(\bar{\pi}_{n+1|\mathcal{G}_{nR}}\|\bar{\pi}_{n|\mathcal{G}_{nR}}\right) = \overline{\mathbb{E}}\left[\log Z'_n|\mathcal{F}^Y_{1,n+1}\right]$.

Proof. (i) follows directly from the definitions of Z_n and $\mathcal{B}_n$ (5.25) and (5.26) along with the fact that $\mathcal{G}_\infty = \sigma(W)$ (Definition 2.4.1) and the fact that $\bar{\pi}_{n+1}$ and $\bar{\pi}_n$ are functions of Y^{n+1}.(ii) follows for analogous reasons as (i) along with the fact that $\bar{\pi}_{n+1|\mathcal{G}_{nR}}$ and $\bar{\pi}_{n|\mathcal{G}_{nR}}$ are probability measures on the space $([0,1],\mathcal{G}_{nR})$.

As for (iii), we have already shown that Z'_n is $\mathcal{B}'_n$-measurable. Thus it remains to be shown that for any $A \in \mathcal{B}'_n$, $\overline{\mathbb{P}}(A|\mathcal{F}^Y_{1,n+1}) = \overline{\mathbb{E}}_{\bar{\pi}_n}\left[1_A\overline{\mathbb{E}}\left[Z_n|\mathcal{B}'_n\right]\right]$, where $1_A(\omega) = 1$ if $\omega \in A$ and is 0 otherwise. Note that

$$
\begin{aligned}
\overline{\mathbb{P}}(A|\mathcal{F}^Y_{1,n+1}) &= \int_{\omega\in\Omega} 1_A(\omega)Z_n(\omega)\bar{\pi}_n(d\omega) &(5.31)\\
&= \int_{\omega\in\Omega} \overline{\mathbb{E}}[1_A Z_n|\mathcal{B}'_n](\omega)\bar{\pi}_{n|\mathcal{G}_{nR}}(d\omega)\\
&= \int_{\omega\in\Omega} 1_A(\omega)\overline{\mathbb{E}}[Z_n|\mathcal{B}'_n](\omega)\bar{\pi}_{n|\mathcal{G}_{nR}}(d\omega) &(5.32)\\
&= \overline{\mathbb{E}}_{\bar{\pi}_n}\left[1_A\overline{\mathbb{E}}\left[Z_n|\mathcal{B}'_n\right]\right]
\end{aligned}
$$

where (5.31) follows from (5.25) and (5.32) follows because $A \in \mathcal{B}'_n$. (iii) follows from the definition of KL-divergence and definitions of Z_n and Z'_n in (5.25). $\square$

We next state a lemma about special properties of Z_n unique to the PM scheme.

Lemma 5.5.5. *For the PM scheme, over a discrete and non-degenerate memoryless channel:*

(i) $Z_n = \frac{dP_{Y|X=\bar{e}(\bar{\pi}_n,W)}}{dP_Y}(Y_{n+1})$.

(ii) $0 < (Z_n)_{n\geq 1} \subset L_1(\overline{\mathbb{P}})$.

(iii) $\overline{\mathbb{E}}\left[D\left(\bar{\pi}_{n+1}\|\bar{\pi}_n\right)\middle|\mathcal{F}^Y_{1,n}\right] = \overline{\mathbb{E}}\left[\log Z_n|\mathcal{F}^Y_{1,n}\right] = C$.

(iv) $0 \leq G_n \leq C$.

(v) If $\mathcal{F}^Y_{1,n} = \mathcal{F}^Y_{1,n} \vee \sigma(W)$ $\overline{\mathbb{P}}$*-a.s., then* $G_n = 0$ $\overline{\mathbb{P}}$*-a.s.*

Proof: (i) holds for the PM scheme because the Y_i's are i.i.d. (ii) holds because the channel is discrete and non-degenerate and so from (2.25):

$$0 < \min_{x\in\mathsf{X},y\in\mathsf{Y}} \frac{dP_{Y|X=x}}{dP_Y}(y) \leq Z_n \leq \max_{x\in\mathsf{X},y\in\mathsf{Y}} \frac{dP_{Y|X=x}}{dP_Y}(y) < \infty.$$

(iii) follows because:

$$\overline{\mathbb{E}}\left[D\left(\bar{\pi}_{n+1}\|\bar{\pi}_n\right)|\mathcal{F}^Y_{1,n}\right] = \overline{\mathbb{E}}\left[\overline{\mathbb{E}}\left[\log Z_n|\mathcal{F}^Y_{1,n+1}\right]|\mathcal{F}^Y_{1,n}\right] \tag{5.33a}$$

$$= \overline{\mathbb{E}}\left[\log Z_n\middle|\mathcal{F}^Y_{1,n}\right] \tag{5.33b}$$

$$= I(W;Y_{n+1}|\mathcal{F}^Y_{1,n}) = C. \tag{5.33c}$$

where (5.33a) follows from the definition of KL-divergence and Z_n (5.25);(5.33b) from the law of iterated expectation; and (5.33c) from Lemma 5.1.1(iii). (iv) follows because

$$G_n = \overline{\mathbb{E}}\left[\log\frac{d\bar{\pi}_{n+1}}{d\bar{\pi}_n}(W) - \log\frac{d\bar{\pi}_{n+1|\mathcal{G}_{nR}}}{d\bar{\pi}_{n|\mathcal{G}_{nR}}}(W)|\mathcal{F}^Y_{1,n}\right] \tag{5.34}$$

$$= \overline{\mathbb{E}}\left[\overline{\mathbb{E}}\left[\log\frac{d\bar{\pi}_{n+1}}{d\bar{\pi}_n}(W) - \log\frac{d\bar{\pi}_{n+1|\mathcal{G}_{nR}}}{d\bar{\pi}_{n|\mathcal{G}_{nR}}}(W)\middle|\mathcal{F}^Y_{1,n+1}\right]\middle|\mathcal{F}^Y_{1,n}\right]$$

$$= \overline{\mathbb{E}}\left[\underbrace{D\left(\bar{\pi}_{n+1}\|\bar{\pi}_n\right) - D\left(\bar{\pi}_{n+1|\mathcal{G}_{nR}}\|\bar{\pi}_{n|\mathcal{G}_{nR}}\right)}_{\geq 0}\middle|\mathcal{F}^Y_{1,n}\right] \tag{5.35}$$

$$\leq \overline{\mathbb{E}}\left[D\left(\bar{\pi}_{n+1}\|\bar{\pi}_n\right)\middle|\mathcal{F}^Y_{1,n}\right] = C \tag{5.36}$$

where (5.34) follows from (5.25) and (5.27); (5.35) follows from Jensen's inequality [52, eqn 4.1] proving $G_n \geq 0$; and (5.36) follows from part (iii). (v) follows because if $\mathcal{F}^Y_{1,n} = \mathcal{F}^Y_{1,n} \vee \sigma(W)$ $\overline{\mathbb{P}}$-a.s., then $\bar{\pi}_n$ and $\bar{\pi}_{n+1}$ are both Dirac measures centered at W. Thus $\log Z_n, \log Z'_n$, and D_n are all 0 $\overline{\mathbb{P}}$-a.s. □

With that, we now prove the main theorem of this section:

Proof of Lemma 5.5.1:

Proof. If $V_n(\bar{\pi}_n) = 0$, then $V_n(\bar{\pi}_{n+1}) = 0$ because $\bar{\pi}_{n+1} \ll \bar{\pi}_n$. If $V_n(\bar{\pi}_n) > 0$,

$$\begin{aligned}
&\overline{\mathbb{E}}\left[V_n(\bar{\pi}_{n+1}) - V_n(\bar{\pi}_n)|\mathcal{F}^Y_{1,n}\right](\omega) \\
=&\overline{\mathbb{E}}\left[-\log \frac{\bar{\pi}_{n+1}\left(\{l : Q_{nR}(l) = Q_{nR}(W)\}\right)}{\bar{\pi}_n\left(\{l : Q_{nR}(l) = Q_{nR}(W)\}\right)}\middle|\mathcal{F}^Y_{1,n}\right] \qquad (5.37a)\\
&= \overline{\mathbb{E}}\left[-\log \frac{\bar{\pi}_{n+1|\mathcal{G}_{nR}}}{\bar{\pi}_{n|\mathcal{G}_{nR}}}(W)\middle|\mathcal{F}^Y_{1,n}\right](\omega) \\
&= \overline{\mathbb{E}}\left[-\log Z'_n|\mathcal{F}^Y_{1,n}\right](\omega) \\
&= \overline{\mathbb{E}}\left[-\log Z_n|\mathcal{F}^Y_{1,n}\right](\omega) + G_n(\omega) \\
&\leq \begin{cases} -(C-\epsilon)1_{\{V_n(\bar{\pi}_n)>0\}}, & \omega \in A_{n,\epsilon} \\ 0, & \text{otherwise.} \end{cases} \qquad (5.37b)
\end{aligned}$$

where (5.37a) follows from (5.28); For $\omega \in A_{n,\epsilon}$ case in (5.37b), use Lemma 5.5.5(iii) and defn of $A_{n,\epsilon}$. For $\omega \notin A_{n,\epsilon}$, use Lemma 5.5.5(iv).

Further, for any $\omega \in \Omega$,

$$\begin{aligned}
&\overline{\mathbb{E}}\left[V_{n+1}(\bar{\pi}_{n+1}) - V_n(\bar{\pi}_{n+1})|\mathcal{F}^Y_{1,n}\right](\omega) \\
=&\overline{\mathbb{E}}\left[-\log \frac{\bar{\pi}_{n+1}\big(l : Q_{(n+1)R}(l) = Q_{(n+1)R}(W)\big)}{\bar{\pi}_{n+1}\left(l : Q_{nR}(l) = Q_{nR}(W)\right)}\middle|\mathcal{F}^Y_{1,n}\right] \qquad (5.38a)\\
&\leq R 1_{\{V_{n+1}(\bar{\pi}_n)>0\}}. \qquad (5.38b)
\end{aligned}$$

Equation (5.38b) is because if $V_{n+1}(\bar{\pi}_n) = 0$ then $V_{n+1}(\bar{\pi}_{n+1}) = 0$ and so (5.38a) is 0; otherwise, $\frac{\bar{\pi}_{n+1}\left(u:Q_{(n+1)R}(u)=Q_{(n+1)R}(W)\right)}{\bar{\pi}_{n+1}(u:Q_{nR}(u)=Q_{nR}(W))}$ corresponds to the probability mass function of a discrete random variable of cardinality 2^R, and thus (5.38a) is its entropy - which is atmost R.

Adding inequalities (5.37) and (5.38), we have:

$$\overline{\mathbb{E}}\left[V_{n+1}(\bar{\pi}_{n+1})-V_n(\bar{\pi}_n)|\mathcal{F}^Y_{1,n}\right] \leq \begin{cases} R1_{\{V_{n+1}(\bar{\pi}_n)>0\}} - (C-\epsilon)1_{\{V_n(\bar{\pi}_n)>0\}}, & \omega \in A_{n,\epsilon} \\ R1_{\{V_{n+1}(\bar{\pi}_n)>0\}}, & \text{otherwise.} \end{cases} \tag{5.39a}$$

$$\leq \begin{cases} -(C-R-\epsilon)1_{\{V_{n+1}(\bar{\pi}_n)>0\}}, & \omega \in A_{n,\epsilon} \\ R1_{\{V_{n+1}(\bar{\pi}_n)>0\}}, & \text{otherwise.} \end{cases} \tag{5.39b}$$

where (5.39b) holds true because $V_{n+1}(\bar{\pi}_n) > 0 \Rightarrow V_n(\bar{\pi}_n) > 0$. □

Proof of Lemma 5.5.2:

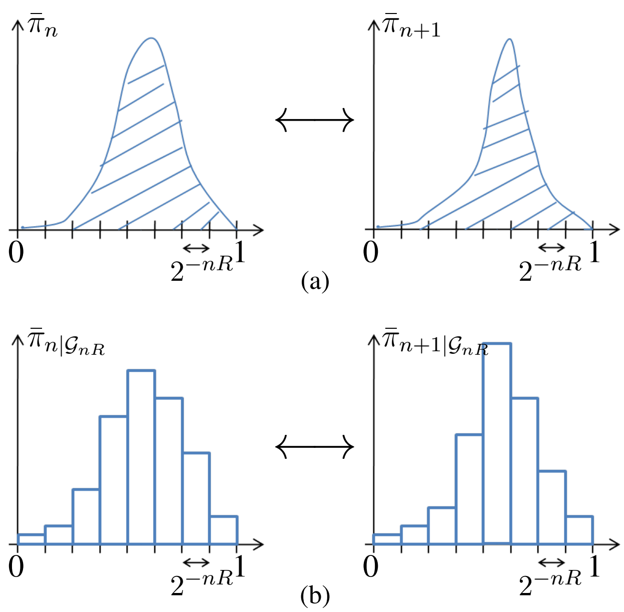

Figure 5.3: (a) Posterior beliefs $\bar{\pi}_n$ and $\bar{\pi}_{n+1}$. (b) Posterior beliefs restricted to field $\mathcal{G}_{nR}$. The difference in KL-divergence between $\bar{\pi}_n$ and $\bar{\pi}_{n+1}$ and KL-divergence between $\bar{\pi}_{n+1|\mathcal{G}_{nR}}$ and $\bar{\pi}_{n|\mathcal{G}_{nR}}$ converges to 0 uniformly in $\overline{\mathbb{P}}$.

Proof. See Figure 5.3.

It is sufficient to show that for every $\epsilon \geq 0$, there exists an $N(\epsilon)$ such that

$$\overline{\mathbb{P}}\left(\left\{\omega : \sup_{k\geq n} G_k(\omega) \leq \epsilon\right\}\right) = 1 \quad \text{for } n \geq N(\epsilon). \tag{5.40}$$

We prove this lemma in two steps

A. Define the set of functions $\Phi_1 = \{f_n : f_n = \log Z_n - \log Z'_n, n \geq 1\}$ and argue that Φ_1 is a subset of $(\mathcal{F}^Y_{1,\infty} \vee \sigma(W))-$measurable bounded functions.

B. Show that $\sup_{k\geq n} \left|\overline{\mathbb{E}}[f_k|\mathcal{F}^Y_{1,k}]\right|(\omega) \leq K'\delta_n$ if $\mathcal{F}^Y_{1,\infty} = \mathcal{F}^Y_{1,\infty} \vee \sigma(W)$, for some constant K' and monotonously decreasing sequence $\delta_n \downarrow 0$. And then argue how this implies (5.40) - what we require in this lemma.

[A] Define the set of functions $\Phi_1 = \{f_n : f_n = \log Z_n - \log Z'_n, n \geq 1\}$. Clearly $f \in \Phi_1$ is a $(\mathcal{F}^Y_{1,\infty} \vee \sigma(W))-$measurable function. It is also bounded because from Lemma 5.5.5, for the PM Scheme $0 < (Z_n)_{n\geq 1} \subset L_1(\overline{\mathbb{P}})$ and moreover,

$$0 < \min_{x\in\mathsf{X},y\in\mathsf{Y}} \frac{dP_{Y|X=x}}{dP_Y}(y) \leq Z_n \leq \max_{x\in\mathsf{X},y\in\mathsf{Y}} \frac{dP_{Y|X=x}}{dP_Y}(y) < \infty. \tag{5.41}$$

Therefore,
$|\log Z_n| \leq \max\left(\left|\log \min_{x\in\mathsf{X},y\in\mathsf{Y}} \frac{dP_{Y|X=x}}{dP_Y}(y)\right|, \left|\log \max_{x\in\mathsf{X},y\in\mathsf{Y}} \frac{dP_{Y|X=x}}{dP_Y}(y)\right|\right) \triangleq \frac{K}{2}$.

From Lemma 5.5.4(iii), $Z'_n = \overline{\mathbb{E}}[Z_n|\mathcal{B}'_n]$ and is thus bounded above and below by the same bounds of Z_n in (5.41). Therefore $|\log Z'_n| \leq \frac{K}{2}$ and for any $f \in \Phi_1$, $|f| \leq |\log Z_n| + |\log Z'_n| \leq K$. Thus, Φ_1 is a subset of $(\mathcal{F}^Y_{1,\infty} \vee \sigma(W))-$measurable bounded functions.

[B] From [70, 71], for every uniformly integrable bounded function (bounded above by K) and nested sub-σ-algebras $\left\{\mathcal{F}^Y_{1,n}\right\}_{n\geq 1}$ such that $\mathcal{F}^Y_{1,m} \subset \mathcal{F}^Y_{1,k} \subset \mathcal{F}^Y_{1,\infty}$ for every $k \geq m$,

$$\sup_{f\in\Phi_1} \left|\overline{\mathbb{E}}[f|\mathcal{F}^Y_{1,k}] - \overline{\mathbb{E}}[f|\mathcal{F}^Y_{1,\infty}](\omega)\right| \leq 4K\delta_k(1-\delta_k) \tag{5.42}$$

where $\delta_m \downarrow 0$ is a decreasing sequence converging to 0 for the nested sub-σ-algebras $\mathcal{F}^Y_{1,m} \uparrow \mathcal{F}^Y_{1,\infty}$. Hence,

$$\sup_{k\geq m} \sup_{f\in\Phi_1} \left|\overline{\mathbb{E}}[f|\mathcal{F}^Y_{1,k}] - \overline{\mathbb{E}}[f|\mathcal{F}^Y_{1,\infty}](\omega)\right| \leq \sup_{k\geq m} 4K\delta_k(1-\delta_k) \tag{5.43}$$

$$= 4K\delta_m(1-\delta_m), \qquad m \geq N_\delta(\frac{1}{2}) \tag{5.44}$$

where $N_\delta(\frac{1}{2})$ is chosen such that $\delta_m \leq \frac{1}{2}$ for every $m \geq N_\delta(\frac{1}{2})$. Any $f \in \Phi_1$ is of the form $f = f_n = \log Z_n - \log Z'_n$. If $\mathcal{F}^Y_{1,\infty} = \mathcal{F}^Y_{1,\infty} \vee \sigma(W)$, then $\overline{\mathbb{E}}[f_n|\mathcal{F}^Y_{1,\infty}] = \overline{\mathbb{E}}[f_n|\mathcal{F}^Y_{1,\infty} \vee \sigma(W)] = 0$ because both $\log Z_n = \log \frac{d\bar{\pi}_{n+1}}{d\bar{\pi}_n}(W)$ and $\log Z'_n =$

$\log \frac{d\bar{\pi}_{n+1|\mathcal{G}_{nR}}}{d\bar{\pi}_{n|\mathcal{G}_{nR}}}(W)$ equals 0 when W is known, for any $n = 1, 2, \cdots$. Hence,

$$\sup_{k \geq m} \sup_{f \in \Phi_1} \left| \overline{\mathbb{E}}[f | \mathcal{F}^Y_{1,k}](\omega) \right| \leq 4K\delta_m(1-\delta_m), \qquad m \geq N_\delta(\frac{1}{2}). \quad (5.45)$$

Note that, for any $f_k \in \Phi_1$

$$\begin{aligned} \left| \overline{\mathbb{E}}[f_k | \mathcal{F}^Y_{1,k}](\omega) \right| &\leq \sup_{f \in \Phi_1} \left| \overline{\mathbb{E}}[f | \mathcal{F}^Y_{1,k}] \right| & (5.46) \\ \Leftrightarrow \sup_{k \geq m} \left| \overline{\mathbb{E}}[f_k | \mathcal{F}^Y_{1,k}](\omega) \right| &\leq \sup_{k \geq m} \sup_{f \in \Phi_1} \left| \overline{\mathbb{E}}[f | \mathcal{F}^Y_{1,k}] \right| & (5.47) \\ &\leq 4K\delta_m(1-\delta_m), \qquad m \geq N_\delta(\frac{1}{2}) & (5.48) \end{aligned}$$

Note that $G_k = \overline{\mathbb{E}}[f_k | \mathcal{F}^Y_{1,k}]$. We can drop the absolute value because of non-negativity of G_k Lemma 5.5.5(iv).

$$\begin{aligned} &\sup_{k \geq m} G_k(\omega) \leq 4K\delta_m(1-\delta_m), \qquad m \geq N_\delta(\frac{1}{2}) \\ &\Leftrightarrow \overline{\mathbb{P}}\left(\left\{ \omega : \sup_{k \geq m} G_k(\omega) \leq \epsilon \right\} \right) = 0 \qquad m \geq N(\epsilon) \qquad (5.49) \end{aligned}$$

where $N(\epsilon)$ is chosen such that $\delta_m(1-\delta_m) \leq \frac{\epsilon}{4K}$ for $m \geq N(\epsilon)$. $N(\epsilon)$ exists because $\delta_m(1-\delta_m) \downarrow 0$ if $\delta_m \downarrow 0$. Moreover, $\delta_m(1-\delta_m) \leq \frac{\epsilon}{4K} \Rightarrow \delta_m \leq \frac{1-\sqrt{1-\frac{\epsilon}{K}}}{2}$. Hence, $N(\epsilon) = N_\delta\left(\frac{1-\sqrt{1-\frac{\epsilon}{K}}}{2}\right)$. □

Proof of Lemma 5.5.3:

Proof. For any $M > 0$ and $\omega \in B_{n,\epsilon}$:

$$\begin{aligned} &\overline{\mathbb{E}}\left[V_{n+M}(\bar{\pi}_{n+M}) - V_n(\bar{\pi}_n) | \mathcal{F}^Y_{1,n} \right](\omega) \\ &= \sum_{k=n}^{n+M-1} \overline{\mathbb{E}}\left[V_{k+1}(\bar{\pi}_{k+1}) - V_k(\bar{\pi}_k) | \mathcal{F}^Y_{1,n} \right](\omega) \\ &= \sum \overline{\mathbb{E}}\left[\overline{\mathbb{E}}\left[V_{k+1}(\bar{\pi}_{k+1}) - V_k(\bar{\pi}_k) | \mathcal{F}^Y_{1,k} \right] | \mathcal{F}^Y_{1,n} \right](\omega) \\ &\leq \sum \overline{\mathbb{E}}\left[-(C-R-\epsilon) 1_{\{V_{k+1}(\bar{\pi}_k) > 0\}} | \mathcal{F}^Y_{1,n} \right](\omega) \\ &= \sum_{k=n}^{n+M-1} -(C-R-\epsilon) \overline{\mathbb{P}}(V_{k+1}(\bar{\pi}_k) > 0 | \mathcal{F}^Y_{1,n})(\omega). \end{aligned}$$

The above sum is non-increasing and thus has a limit. Since

$\overline{\mathbb{E}}\left[V_{n+M+1}(\bar{\pi}_{n+M})|\mathcal{F}^Y_{1,n}\right](\omega) \geq 0, \forall\, \omega \in B_{n,\epsilon}$:

$$\lim_{M\to\infty} \overline{\mathbb{P}}(V_{n+M+1}(\bar{\pi}_{n+M}) > 0|\mathcal{F}^Y_{1,n})(\omega) = 0. \tag{5.50a}$$

Thus,

$$\begin{aligned}
&\lim_{M\to\infty} \overline{\mathbb{P}}(V_{n+M+1}(\bar{\pi}_{n+M}) > 0) \\
&\quad = \lim_{M\to\infty} \overline{\mathbb{E}}\left[\overline{\mathbb{P}}(V_{n+M+1}(\bar{\pi}_{n+M}) > 0|\mathcal{F}^Y_{1,n})\right] \\
&\quad = \overline{\mathbb{E}}\left[\lim_{M\to\infty} \overline{\mathbb{P}}(V_{n+M+1}(\bar{\pi}_{n+M}) > 0|\mathcal{F}^Y_{1,n})\right] && (5.50b) \\
&\quad \leq 0 \times \overline{\mathbb{P}}(B_{n,\epsilon}) + 1 \times \overline{\mathbb{P}}(B^c_{n,\epsilon}) && (5.50c) \\
&\quad = 0 && (5.50d)
\end{aligned}$$

where (5.50b) follows from the bounded convergence theorem; (5.50c) follows from (5.50a); and (5.50d) follows from Lemma 5.5.2. □

CHAPTER 6

CONCLUSIONS AND FUTURE DIRECTIONS

6.1 Conclusions

In this thesis, we focus using perspectives from both information and control theory in analyzing an interactive two-agent sequential decision-making problem. We consider an interacting two-agent decision-making problem consisting of a Markov source process, a causal encoder with feedback, and a causal decoder. We augment the standard formulation by considering general alphabets and a non-trivial cost function operating on current and previous symbols; this enables us to introduce the 'sequential information gain cost' function that can capture information gains accumulated at each time step. We emphasize how this problem formulation leads to a different style of coding schemes with a control-theoretic flavor. Further, we solve for structural results on these optimal policies using dynamic programming principles.

We then demonstrate another interplay between information theory and control theory, at the level of reliability of message-point communication schemes, by establishing a relationship between '*reliability*' in feedback communication to the '*stability*' of the posterior belief's nonlinear filter. We also consider the two-agent inverse optimal control (IOC) problem, where a fixed policy satisfying certain statistical conditions is shown to be optimal for some cost function, using probabilistic matching.

We provide examples of the applicability of this framework to communication with feedback, hidden Markov models and the nonlinear filter, decentralized control, brain-machine interfaces, and queuing theory.

From the context of message point communication schemes, our framework provides a meaningful approach in i) solving for optimal policies (maximizing communication rate) using control-theoretic principles; and ii) defining reliability and providing conditions for these control-theory based policies to be reliable/achievable, completely by using control-theoretic analysis.

6.2 Future Directions

- **Extension of Reliable Feedback based communication schemes to more network scenarios.**

  The posterior matching scheme was inspired by a converse to a known fundamental limit. In fact, the takeaway lesson from the posterior matching scheme described in Section 2.4.3 is that the update/decision policy at the encoder must always satisfy certain properties for it to be optimal, and these properties were motivated from the converse to the point-to-point communication channel. There are other multi-terminal problems with feedback with tight information-theoretic converses and interactions. This includes the degraded broadcast channel with feedback studied by El Gamal [72]. Further investigation is required to see how control-theory based communication policies (e.g., posterior matching scheme) can be used for simplifying encoder-decoder schemes for many multi-terminal problems with feedback.

- **Error Exponent Analysis for control-theory based communication policies.**

  In the context of message-point communication schemes, we have provided optimality conditions when a certain control-based coding policy maximizes mutual information (converse for the communication problem) in Part I. We have also provided conditions when a control-based coding policy is reliable (achievability result for the communication scheme) in Part II. The next step is to explore ways to perform error-exponent analysis for a given control-based policy. Further investigation is required to extend the Lyapunov analysis performed in Section 5.5 to develop fundamental limits of error exponents for feedback w/ fixed block length using Martingale condition implied by the Lyapunov function (equation (5.29)).

- **Sequential information gain as a metric**

  The sequential information gain cost can be used as a metric for suitable problems where a measure of information transfer has to be computed. Prof. Douglas Jones' group at the University of Illinois works on ultra-low-power energy devices, where energy becomes a critical resource. Their group is focusing on developing intelligent management strategies to use energy only when it is useful. Prof. Jones group focused on strategies where they use sequential information gain as a metric and provided key results on sensor scheduling problems. This approach to energy management is most effective in dynamic environments and finds meaningful applications in the detection and monitoring of physical dynamic phenomena (i.e. wildlife, healthcare, civil infrastructure, smart buildings, military). This might be a starting point for other practical applications.

- **Other directions**

  - **Delays $\geq$ 2**: The two-agent team decision problem formulation we consider involves a one-step delay in sharing the observations from the second agent to the first agent. It is interesting to see what happens when the delays are ≥ 2. In fact, the proof technique in the structural result (Section 3.1) is based on grouping decisions of two agents (e_{i+1}, d_i) who share a common piece of information $(Y_1, \cdots, Y_i)$. This technique might not work when delays are greater than 2. Further analysis is required to bring in ideas from control theory literature which deal with delayed information sharing patterns (see Nayyar, Mahajan and Teneketzis [73]).
  - **Noisy Feedback:** We consider a perfect feedback loop in our two agent problem setting. This might not be the case in some practical applications. Feedback could be noisy or limited, and solving for *reliable* communication policies that can make use of noisy feedback has been an open problem. There are reliable communication schemes provided by information theory literature for the case when there is no feedback (Shannon [10]) and for the case when there is perfect feedback (posterior matching [16]), but no explicit scheme is known which can make use of noisy or limited feedback. It would be interesting to see if control theory can provide intuition on using noisy/limited feedback and thus see if encoding policies could be designed in a simpler, low-complex fashion based on (noisy) feedback.
  - **Connection to Learning with Expert Advice:** We define reliability of a communication policy (Sec 2.4.2) in terms of performance of a

uniform(or non-genie aided) prior in comparison with a genie-aided prior who knows what the actual message is. We say a communication policy is reliable if we can decode the message by starting with a uniform prior. In other words, the performance of the non-genie aided prior should be same as the performance of the genie-aided prior. This perspective can provide a connection to Learning with expert advice where

* the genie-aided prior could be the expert, the uniform prior could be the user,
* the expected value of the regret could be the KL-divergence between the posterior beliefs of the expert and the user (uniform prior),
* the actual loss function at each time step could be the negative of sequential information gain (negative log RN-derivative of the posterior beliefs of the user between consecutive time-steps).

It remains to be seen whether the regret goes to zero, and if it does, how fast.

APPENDIX A

PROOF OF LEMMA 3.1.1

Proof. As described in the statement of the lemma, define the state space $\mathsf{S} = \mathsf{Z} \times \mathcal{P}(\mathsf{W})$ and control space $\mathsf{U} = \tilde{\mathsf{E}} \times \mathsf{Z}$ with $s_i \in \mathsf{S}, u_i \in \mathsf{U}$ given by (3.4):

$$s_i = (z_{i-1}, b_{i|i}), \qquad u = (\tilde{e}_{i+1}, z_i).$$

Then

$$\begin{aligned}
& \mathbb{E}\left[\rho(W_i, Z_{i-1}, Z_i) | Z_{i-1} = z_{i-1}, Y^i = y^i, Z_i = z_i\right] \\
= \; & \int_{w_i \in \mathsf{W}} \rho(w_i, z_{i-1}, z_i) b_{i|i}(dw_i) && \text{(A.1)} \\
\equiv \; & \bar{\rho}(s_i, z_i) \\
& \mathbb{E}\left[\eta(X_{i+1}) | Z_{i-1} = z_{i-1}, Y^i = y^i, \tilde{E}_{i+1} = \tilde{e}_{i+1}\right] \\
= \; & \int_{w_{i+1} \in \mathsf{W}} \alpha\eta\left(\tilde{e}_{i+1}\left(w_{i+1}\right)\right) b_{i+1|i}(dw_{i+1}) \\
= \; & \int_{w_{i+1} \in \mathsf{W}} \alpha\eta\left(\tilde{e}_{i+1}\left(w_{i+1}\right)\right) \Phi(b_{i|i})(dw_{i+1}) && \text{(A.2)} \\
\equiv \; & \bar{\eta}(s_i, \tilde{e}_{i+1})
\end{aligned}$$

where (A.1) follows from (2.53); (A.2) follows from (2.55). $\square$

APPENDIX B

PROOF OF LEMMA 3.1.2

Proof. Note that

$$
\begin{aligned}
& P_{S_{i+1}|S^i=s^i,U^i=u^i}(ds_{i+1}) \\
= {} & 1_{\{s_{i+1,1}=u_{i,2}\}} \int_{w_{i+1}\in \mathsf{W}} \int_{y_{i+1}\in \mathsf{Y}} 1_{\left\{b_{i+1|i+1}=\Lambda(b_{i|i},y_{i+1},\tilde{e}_{i+1})\right\}} \\
& \quad P_{Y|X}\left(dy_{i+1}|\tilde{e}_{i+1}(w_{i+1})\right) b_{i+1|i}(dw_{i+1}) \qquad \text{(B.1a)} \\
= {} & 1_{\{s_{i+1,1}=u_{i,2}\}} \int_{w_{i+1}\in \mathsf{W}} \int_{y_{i+1}\in \mathsf{Y}} 1_{\{s_{i+1,2}=\Lambda(s_{i,2},y_{i+1},u_{i,1})\}} \\
& \quad P_{Y|X}\left(dy_{i+1}|u_{i,1}(w_{i+1})\right) \Phi(s_{i,2})(dw_{i+1}) \qquad \text{(B.1b)} \\
= {} & P_{S_{i+1}|S_i=s_i,U_i=u_i}(ds_{i+1}) \\
\equiv {} & Q_S(ds_{i+1}|s_i,u_i) \qquad \text{(B.1c)}
\end{aligned}
$$

where (B.1a) follows from (2.51) and (2.48); (B.1b) follows from (3.4); and (B.1c) demonstrates that this is a controlled Markov chain with *time-invariant* statistical dynamics. □

APPENDIX C

PROOF OF LEMMA 3.2.1

Proof.

$$\begin{aligned}
I(W^n;Y^n) &= \sum_{i=1}^{n} I(W^n;Y_i|Y^{i-1}) && \text{(C.1)}\\
&= \sum_{i=1}^{n} I(W^i;Y_i|Y^{i-1}) \\
&\qquad + I(W_{i+1}^n;Y_i|Y^{i-1},W^i) && \text{(C.2)}\\
&= \sum_{i=1}^{n} I(W^i;Y_i|Y^{i-1}) \\
&\qquad + I(W_{i+1}^n;Y_i|Y^{i-1},W^i,X_i) && \text{(C.3)}\\
&= \sum_{i=1}^{n} I(W^i;Y_i|Y^{i-1}) && \text{(C.4)}\\
&= \sum_{i=1}^{n} I(W_i;Y_i|Y^{i-1}) \\
&\qquad + I(W^{i-1};Y_i|W_i,Y^{i-1}) && \text{(C.5)}\\
&= \sum_{i=1}^{n} I(W_i;Y_i|Y^{i-1}) \\
&\qquad + I(W^{i-1};Y_i|W_i,Y^{i-1},X_i) && \text{(C.6)}\\
&= \sum_{i=1}^{n} I(W_i;Y_i|Y^{i-1}) && \text{(C.7)}
\end{aligned}$$

where (C.1) follows from (2.8); (C.2) follows from from (2.8); (C.3) follows from (2.49); (C.4) follows follows from (2.51); (C.5) follows from (2.8); (C.6) follows from our assumption (3.9a) that the encoder operates on sufficient statistics; and (C.7) follows from (2.51). □

APPENDIX D

PROOF OF LEMMA 3.2.2

Proof. (3.12) follows directly from Lemma 3.1.1. Now, let us focus on (3.13). From Lemma 3.1.1, we have that

$$\begin{aligned}
\tilde{\rho}(s,z) &= \int_{w \in \mathsf{W}} \rho(w,b,z) b'(dw) \\
&= \int_{w \in \mathsf{W}} -\log \frac{dz}{d\Phi(b)}(w) b'(dw) \qquad \text{(D.1)}
\end{aligned}$$

where (D.1) follows from (3.11) for any z satisfying $z \ll \Phi(b)$ and is infinite otherwise. Now note that if it is not the case that $b' \ll z$, then there exists a set $A \in \mathcal{B}(\mathsf{W})$ for which $z(A) = 0$ and $b'(A) > 0$ and thus it follows that $\frac{dz}{d\Phi(b)}(w) = 0 \Rightarrow -\log \frac{dz}{d\Phi(b)}(w) = \infty$ for all $w \in A$. Thus $\tilde{\rho}(s,z) = \infty$. Now assume $b' \ll z \ll \Phi(b)$. Then since if $\beta \ll \nu \ll \mu$ then $\frac{d\beta}{d\mu} = \frac{d\beta}{d\nu}\frac{d\nu}{d\mu}$, μ-almost everywhere [74, Sec 5.5], it follows that

$$\begin{aligned}
\tilde{\rho}(s,z) &= \int_{w \in W} -\log \frac{db'}{d\Phi(b)}(w) b'(dw) + \log \frac{db'}{dz}(w) b'(dw) \\
&= -D\left(b' \| \Phi(b)\right) + D\left(b' \| z\right). \qquad \text{(D.2)}
\end{aligned}$$

$\square$

APPENDIX E

PROOF OF THEOREM 3.2.3

Proof. In order to find the optimal cost J_n^* given by $J_n^* = \mathbb{E}\left[V_0(S_0)\right]$, we use the standard dynamic programming approach and evaluate optimal cost-to-go functions $\{V_k : k = 0, \cdots, n\}$. Consider the final-stage problem of finding $V_n(s_n)$, where $s_n = (z_{n-1}, b_{n|n})$ and describe any control u_n as $u_n = (\tilde{e}_{n+1}, z_n)$. Then the one-stage problem is

$$\begin{aligned} V_n\left(s_n\right) &= \inf_{u_n=(\tilde{e}_{n+1},z_n)} \bar{g}_n\left(s_n, u_n\right) \\ &= \inf_{z_n\in\mathcal{P}(\mathsf{W})} \bar{\rho}(s_n, z_n) \qquad \text{(E.1)} \\ &= -D\left(b_{n|n} \| \Phi(z_{n-1})\right) + \inf_{z_n\in\mathcal{P}(\mathsf{W}),\ b_{n|n}\ll z_n\ll\Phi(z_{n-1})} D\left(b_{n|n}\|z_n\right) \qquad \text{(E.2)} \\ &= -D\left(b_{n|n}\|\Phi(z_{n-1})\right) \qquad \text{(E.3)} \end{aligned}$$

where (E.1) follows (3.7); (E.2) follows from (3.13); and (E.3) follows from the non-negativity of the KL divergence. The optimal choice of z_n is the one for which the equality in (E.3) holds true and hence under an optimal policy, $z_n = b_{n|n}$. This follows the same reasoning that elicits how for in the self-information loss sequential probability assignment, the best probability assignment is the true belief [57].

For the second step $k = n-1$, consider finding $V_{n-1}(s_{n-1})$, where $s_{n-1} = (z_{n-2}, b_{n-1|n-1})$ and describe any control u_{n-1} as $u_{n-1} = (\tilde{e}_n, z_{n-1})$. Then we have:

$$\begin{aligned}
&V_{n-1}(s_{n-1}) \\
&= \inf_{u_{n-1}=(\tilde{e}_n, z_{n-1})} \bar{g}_{n-1}(s_{n-1}, u_{n-1}) \\
&+ \mathbb{E}\left[V_n\left(z_{n-1}, B_{n|n}\right) | S_{n-1}=s_{n-1}, U_{n-1}=u_{n-1}\right] && \text{(E.4)} \\
&= \inf_{\tilde{e}_n, z_{n-1}} \alpha\bar{\eta}(s_{n-1}, \tilde{e}_n) + \bar{\rho}(s_{n-1}, z_{n-1}) \\
&+ \mathbb{E}\left[-D\left(B_{n|n} \| \Phi(z_{n-1})\right) | S_{n-1}=s_{n-1}, U_{n-1}=u_{n-1}\right] && \text{(E.5)} \\
&= -D\left(b_{n-1|n-1} \| \Phi(z_{n-2})\right) + \inf_{\tilde{e}} \alpha\bar{\eta}(s_{n-1}, \tilde{e}_n) \\
&+ \inf_{b_{n-1|n-1} \ll z_{n-1} \ll \Phi(z_{n-2})} D\left(b_{n-1|n-1} \| z_{n-1}\right) \\
&+ \mathbb{E}\left[-D\left(B_{n|n} \| \Phi(z_{n-1})\right) | S_{n-1}=s_{n-1}, \tilde{E}_n=\tilde{e}_n\right] && \text{(E.6)}
\end{aligned}$$

where (E.5) follows by substituting values of $\bar{g}_{n-1}$ and V_n from (3.7) and (E.3); (E.6) follows from (3.13).

For any fixed encoder policy $\tilde{e}_n$, the optimal choice for z_{n-1} is to pick $z_{n-1} = b_{n-1|n-1}$ as shown:

$$\begin{aligned}
&z^*_{n-1}(s_{n-1}) = \\
&\quad \operatorname*{arg\,inf}_{b_{n-1|n-1} \ll z_{n-1} \ll \Phi(z_{n-2})} D\left(b_{n-1|n-1} \| z_{n-1}\right) \\
&- \mathbb{E}\left[D\left(B_{n|n} \| \Phi(z_{n-1})\right) | S_{n-1}=s_{n-1}, \tilde{E}_n=\tilde{e}_n\right] && \text{(E.7)} \\
&= \operatorname*{arg\,inf}_{b_{n-1|n-1} \ll z_{n-1} \ll \Phi(z_{n-2})} D\left(P_\Lambda\left(\cdot | b_{n-1|n-1}, \tilde{e}_n\right) \| P_\Lambda\left(\cdot | z_{n-1}, \tilde{e}_n\right)\right) \\
&\underbrace{+ D\left(b_{n-1|n-1} \| z_{n-1}\right) - \mathbb{E}\left[D\left(\Lambda\left(b_{n-1|n-1}, Y_n, \tilde{e}_n\right) \| \Lambda(z_{n-1}, Y_n, \tilde{e}_n)\right) | S_{n-1}=s_{n-1}, \tilde{E}_n=\tilde{e}_n\right]}_{\geq 0} \\
& && \text{(E.8)} \\
&= b_{n-1|n-1}. && \text{(E.9)}
\end{aligned}$$

where (E.8) and the non-negativity of the difference follow because:

$$
\begin{aligned}
&z^*_{n-1}(s_{n-1})\\
&= \operatorname*{arg\,inf}_{b_{n-1|n-1} \ll z_{n-1} \ll \Phi(z_{n-2})} D\left(b_{n-1|n-1} \| z_{n-1}\right)\\
&- \mathbb{E}\left[D\left(B_{n|n} \| \Phi(z_{n-1})\right) | S_{n-1} = s_{n-1}, \tilde{E}_n = \tilde{e}_n\right]\\
&= \operatorname*{arg\,inf}_{b_{n-1|n-1} \ll z_{n-1} \ll \Phi(z_{n-2})} D\left(b_{n-1|n-1} \| z_{n-1}\right)\\
&- \mathbb{E}\left[D\left(B_{n|n} \| \Phi(b_{n-1|n-1})\right) - \mathbb{E}_{B_{n|n}}\left[\log \frac{d\Phi(b_{n-1|n-1})}{d\Phi(z_{n-1})} \middle| B_{n|n}\right] \middle| B_{n-1|n-1} = b_{n-1|n-1}\right] \qquad \text{(E.10)}\\
&= \operatorname*{arg\,inf}_{b_{n-1|n-1} \ll z_{n-1} \ll \Phi(z_{n-2})} D\left(b_{n-1|n-1} \| z_{n-1}\right)\\
&- \int_{y \in \mathsf{Y}} \int_{w \in \mathsf{W}} \log \frac{d\Phi(b_{n-1|n-1})}{d\Phi(z_{n-1})}(w) \underbrace{\Lambda(b_{n-1|n-1}, y, \tilde{e}_n)}_{b_{n|n}}(dw) P_\Lambda\left(dy | b_{n-1|n-1}, \tilde{e}_n\right) \qquad \text{(E.11)}\\
&= \operatorname*{arg\,inf}_{b_{n-1|n-1} \ll z_{n-1} \ll \Phi(z_{n-2})} D\left(b_{n-1|n-1} \| z_{n-1}\right)\\
&- \iint_{y \in \mathsf{Y}\, w \in \mathsf{W}} \log\left(\frac{d\Lambda(b_{n-1|n-1}, y, \tilde{e}_n)}{d\Lambda(z_{n-1}, y, \tilde{e}_n)}\right)(w) \Lambda(b_{n-1|n-1}, y, \tilde{e}_n)(dw) P_\Lambda\left(dy | b_{n-1|n-1}, \tilde{e}_n\right)\\
&+ \int_{y \in \mathsf{Y}} \log \frac{dP_\Lambda\left(\cdot | b_{n-1|n-1}, \tilde{e}_n\right)}{dP_\Lambda\left(\cdot | z_{n-1}, \tilde{e}_n\right)}(y) P_\Lambda\left(dy | b_{n-1|n-1}, \tilde{e}_n\right) \qquad \text{(E.12)}\\
&= \operatorname*{arg\,inf}_{b_{n-1|n-1} \ll z_{n-1} \ll \Phi(z_{n-2})} D\left(P_\Lambda\left(\cdot | b_{n-1|n-1}, \tilde{e}_n\right) \| P_\Lambda\left(\cdot | z_{n-1}, \tilde{e}_n\right)\right)\\
&+ \underbrace{D\left(b_{n-1|n-1} \| z_{n-1}\right) - \mathbb{E}_{P_\Lambda\left(\cdot | b_{n-1|n-1}, \tilde{e}_n\right)}\left[D\left(\Lambda\left(b_{n-1|n-1}, Y_n, \tilde{e}_n\right) \| \Lambda(z_{n-1}, Y_n, \tilde{e}_n)\right)\right]}_{\geq 0} \qquad \text{(E.13)}
\end{aligned}
$$

where (E.10) follows because $B_{n|n} \ll \Phi(b_{n-1|n-1}) \ll \Phi(z_{n-1})$ and so $\frac{dB_{n|n}}{d\Phi(z_{n-1})} = \frac{dB_{n|n}}{d\Phi(b_{n-1|n-1})} \frac{d\Phi(b_{n-1|n-1})}{d\Phi(z_{n-1})}$; (E.11) follows from the definition of the nonlinear filter (2.57b); (E.12) follows from the fact that $b_{n-1|n-1} \ll z_{n-1}$ and the definition of the nonlinear filter in (2.54); and the difference in (E.13) being non-negative follows from mapping this scenario to that of the hidden Markov model and the nonlinear filter:

- Here, the latent Markov process is W and one observation Y_n is recorded.

- Because in this dynamic programming problem, while in state s_{n-1} and under a fixed $\tilde{e}_n : \mathsf{W} \to \mathsf{X}$, the noisy channel from W_n to Y_n is the composition of the encoder map $\tilde{e}_n$ and the input to the channel from X_n to Y_n: $P_{Y_n|W_n}(dy|w_n) = P_{Y|X}(dy|\tilde{e}_n(w_n))$.
- Two different decoders both know the statistical dynamics but have different initial beliefs about W_{n-1}. One decoder's initial belief is $b_{n-1|n-1} \in \mathcal{P}(\mathsf{W})$ and the other's is $z_{n-1} \in \mathcal{P}(\mathsf{W})$. The initial 'distance' between the beliefs is measured by the KL divergence, $D\left(b_{n-1|n-1} \| z_{n-1}\right)$.
- Both decoders observe Y_n and update their beliefs about W_n according to the one-step nonlinear filter one updates its belief according to $\Lambda(b_{n-1|n-1}, Y_n, \tilde{e}_n)$ and the other does so according to $\Lambda(z_{n-1}, Y_n, \tilde{e}_n)$. The divergence between their beliefs after the observation is given by $D\left(\Lambda\left(b_{n-1|n-1}, Y_n, \tilde{e}_n\right) \| \Lambda\left(z_{n-1}, Y_n, \tilde{e}_n\right)\right)$ and on average this is smaller than the original due to Jensen's inequality and the second law of thermodynamics for hidden Markov chains. This inequality is thus a manifestation of how the relative entropy is a 'Lyapunov function' for the stability (e.g. insensitivity to initial beliefs) of the nonlinear filter [52, Remark 4.2].

Hence the optimal choice for z_{n-1} is to pick $b_{n-1|n-1}$. Consequently,

$$\begin{aligned}
&V_{n-1}\left(s_{n-1}\right) \\
&= -D\left(b_{n-1|n-1} \| \Phi(z_{n-2})\right) + \inf_{\tilde{e}_n} \alpha\bar{\eta}(s_{n-1}, \tilde{e}_n) \\
&\quad + \mathbb{E}\left[V_n\left(b_{n-1|n-1}, B_{n|n}\right) | B_{n-1|n-1} = b_{n-1|n-1}, \tilde{E}_n = \tilde{e}_n\right] \\
&= -D\left(b_{n-1|n-1} \| \Phi(z_{n-2})\right) + \alpha\bar{\eta}(s_{n-1}, \tilde{e}_n^*[b_{n-1|n-1}]) \\
&+ \mathbb{E}\left[V_n\left(b_{n-1|n-1}, B_{n|n}\right) | B_{n-1|n-1} = b_{n-1|n-1}, \tilde{E}_n = \tilde{e}_n^*[b_{n-1|n-1}]\right]. \quad \text{(E.14)}
\end{aligned}$$

Using an inductive argument and the exact same set of arguments as above, it follows that for any $1 \leq k \leq n-1$, and any encoder policy $\tilde{e}_{k+1}$, the optimal choice for z_k is given by $b_{k|k}$ and that for $s_k = (z_{k-1}, b_{k|k})$,

$$\begin{aligned}
V_k\left(s_k\right) = -D\left(b_{k|k} \| \Phi(z_{k-1})\right) + \alpha\bar{\eta}(s_k, \tilde{e}_{k+1}^*[b_{k|k}]) \\
+ \mathbb{E}\left[V_{k+1}\left((b_{k|k}, B_{k+1|k+1})\right) | B_{k|k} = b_{k|k}, \tilde{E}_{k+1} = \tilde{e}_{k+1}^*[b_{k|k}]\right]. \quad \text{(E.15)}
\end{aligned}$$

For the initial step, $k = 0$, by definition $Z_0(A) = B_{0|0}(A)\mathbb{P}\left(W_0 \in A\right)$ and is known to both encoder and decoder. Thus the minimization is only over $\tilde{e}_1$. For a state $s_0 = (z_{-1}, b_{0|0})$ and control $u_0 = (\tilde{e}_1, z_0) = (\tilde{e}_1, b_{0|0})$, we have:

$$\begin{aligned} V_0\left(s_0\right) &= \inf_{\tilde{e}_1} \alpha\bar{\eta}(s_0, \tilde{e}_1) + \mathbb{E}\left[V_1\left(b_{0|0}, B_{1|1}\right)|B_{0|0} = b_{0|0}, \tilde{E}_1 = \tilde{e}_1\right] \\ &= \alpha\bar{\eta}(s_0, \tilde{e}_1^*[b_{0|0}]) + \mathbb{E}\left[V_1\left(b_{0|0}, B_{1|1}\right)|B_{0|0} = b_{0|0}, \tilde{E}_1 = \tilde{e}_1^*[b_{0|0}]\right]. \end{aligned} \tag{E.16}$$

Next, from (3.10) we have that $I(W_i; Y_i|Y^{i-1}) = \mathbb{E}\left[D\left(B_{i|i}\|\Phi(B_{i-1|i-1})\right)\right]$ and thus from (3.2.1) we have:

$$J^{\alpha}_{n,\gamma^*} = \mathbb{E}\left[V_0(S_0)\right] = \min_{e\in\mathsf{E}} -I(W^n; Y^n) + \alpha\mathbb{E}_e\left[\sum_{i=1}^{n}\eta(X_i)\right]. \tag{E.17}$$

Lastly, it follows directly that a more 'concise' sufficient statistic exists for the encoder - namely that it does not need to maintain z_{i-1} to produce x_{i+1} because under any optimal scheme, $Z_{i-1} = B_{i-1|i-1}$ and thus $\sigma(Z_{i-1}) \subset \sigma(B_{i|i})$ so the state variable $S_i = (Z_{i-1}, B_{i|i})$ can be reduced to $S_i = (B_{i|i})$ with Lemma 3.1.2 still holding. $\square$

APPENDIX F

PROOF OF LEMMA 4.1.2

Proof. Note the following standard set of inequalities:

$$
\begin{aligned}
R_n(\rho, P_{W^n}, D) \leq\ & \frac{1}{n} I(W^n; Z^n) && \text{(F.1a)}\\
\leq\ & \frac{1}{n} I(W^n; Y^n) && \text{(F.1b)}\\
=\ & \frac{1}{n} \sum_{i=1}^{n} I(W_i; Y_i | Y^{i-1}) && \text{(F.1c)}\\
=\ & \frac{1}{n} \sum_{i=1}^{n} D\left(P_{Y_i|W_i,Y^{i-1}} \| P_{Y_i|Y^{i-1}} | P_{W_i,Y^{i-1}}\right) && \\
=\ & \frac{1}{n} \sum_{i=1}^{n} D\left(P_{Y_i|X_i} \| P_{Y_i|Y^{i-1}} | P_{X_i,Y^{i-1}}\right) && \text{(F.1d)}\\
\leq\ & \frac{1}{n} \sum_{i=1}^{n} D\left(P_{Y_i|X_i} \| P_{Y_i} | P_{X_i}\right) && \text{(F.1e)}\\
=\ & \frac{1}{n} \sum_{i=1}^{n} I(X_i; Y_i) && \\
\leq\ & \frac{1}{n} \sum_{i=1}^{n} C\left(\eta, P_{Y|X}, \mathbb{E}[\eta(X_i)]\right) && \text{(F.1f)}\\
\leq\ & C\left(\eta, P_{Y|X}, L\right) && \text{(F.1g)}
\end{aligned}
$$

where (F.1a) follows (4.4); (F.1b) follows from the data processing inequality; (F.1c) follows from Lemma 3.2.1; (F.1d) follows from the definition of conditional mutual information (2.7) and the fact that X_i is a function of W_i and Y^{i-1} under policy $\bar{e}$; (F.1e) follows from the memoryless nature of the channel (2.51) and Jensen's inequality; (F.1f) follows from (2.10); and (F.1g) follows from (4.3) and the concavity of the capacity-cost function [39]. $\square$

APPENDIX G

PROOF OF LEMMA 4.1.4

Proof. To prove (4.6a),

$$\begin{aligned}
& P_{Z_i|Z^{i-1}=z^{i-1},W^n=w^n}(dz_i) \\
= \; & \int_{\mathsf{Y}} P_{Z_i|Z^{i-1}=z^{i-1},W^n=w^b,Y_i=y}(dz_i) \\
& \qquad \times P_{Y_i|Z^{i-1}=z^{i-1},W^n=w^n}(dy) \\
= \; & \int_{\mathsf{Y}} P_{Z_i|Z^{i-1}=z^{i-1},W^n=w^b,Y_i=y}(dz_i) \\
& \quad \times P_{Y_i|Z^{i-1}=z^{i-1},W^n=w^n,X_i=\bar{e}(w_i,z_{i-1})}(dy) && \text{(G.1)} \\
= \; & \int_{\mathsf{Y}} 1_{\left\{z_i=\bar{d}(z_{i-1},y_i)\right\}} P_{Y|X=\bar{e}(w_i,z_{i-1})}(dy) && \text{(G.2)} \\
\triangleq \; & Q_{Z'|Z,W'}(dz_i|z_{i-1},w_i) && \text{(G.3)}
\end{aligned}$$

where (G.1) follows from the stationary Markov encoder policy: $x_i = \bar{e}(w_i, z_{i-1})$; (G.2) follows from defining $1_{\left\{z_i=\bar{d}(z_{i-1},y_i)\right\}}$ as a Dirac measure at the point $\bar{d}(z_{i-1}, y_i)$, the stationary Markov decoder policy $z_i = \bar{d}(z_{i-1}, y)$, and the non-anticipative and memoryless nature of the channel (2.51); and (G.3) simply denotes the time-invariant nature of the conditional distribution.

To prove (4.6b), we exploit the assumption that $\{Y_i\}$ are i.i.d. Because of this, we can denote $(Z_i : i = 1, \ldots, n)$ by the following composition of independent random maps:

$$Z_i = \bar{d}(Z_{i-1}, Y_i) \triangleq \bar{d}_{Y_i}(Z_{i-1}) = \bar{d}_{Y_i} \circ \bar{d}_{Y_{i-1}} \circ \cdots \circ \bar{d}_{Y_2}(Z_1).$$

This is thus an an iterated function system (IFS) [64], which is a time-homogeneous Markov chain over the state space Z. $\square$

APPENDIX H

PROOF OF LEMMA 4.4.1

Proof. Let $E_i \triangleq W_i - \mathbb{E}\left[W_i|Y^{i-1}\right]$ be the error term in estimation. We now select the statistics of W_0 such that $X_i \sim \mathcal{N}(0, L), \forall i$. The normalizing coefficient can be expressed as $\beta_i = \sqrt{\frac{L}{\mathrm{Cov}(E_i, E_i)}}$, where the covariance of the error term can be recursively computed using

$$\mathrm{Cov}(E_i, E_i) = \begin{cases} \frac{\rho^2\sigma_n^2}{L+\sigma_n^2}\mathrm{Cov}(E_{i-1}, E_{i-1}) + \sigma_m^2, & i \geq 1; \\ \mathrm{Cov}(W_0, W_0), & i = 0. \end{cases} \tag{H.1}$$

Let the steady state value of the covariance from (H.1) be denoted by C. Then,

$$C \triangleq \frac{\sigma_m^2}{1 - \rho^2 \frac{\sigma_n^2}{L+\sigma_n^2}}. \tag{H.2}$$

Note that because of the choice of $\tilde{W}_0$ in (4.35b), $\mathrm{Cov}(E_i, E_i) = C$ and $\beta_i = \beta = \sqrt{\frac{L}{C}}$ for all $i \geq 0$.

Since all operations are linear and all primitive random variables $(\tilde{W}_i, V_i : i \geq 1)$ are i.i.d. and Gaussian, and since all other relationships are linear, all random variables are jointly Gaussian. From standard MMSE estimation theory, E_i is thus independent of Y^{i-1}. As such, clearly $I(X_i; Y^{i-1}) = 0$. Since the initial condition $\tilde{W}_0$ is chosen according to (4.35a), $X_i \sim \mathcal{N}(0, L)$ for all i. Therefore, since the variance of V_i's is σ_v^2, this means that Y's are i.i.d. The policies (4.37) are thus stationary-Markov coordination strategies:

$$X_i = \beta\left(W_i - \rho Z_{i-1}\right) \tag{H.3a}$$

$$Z_i = \rho Z_{i-1} + \gamma Y_i \tag{H.3b}$$

where (H.3a) follows because $\mathbb{E}\left[W_i|Y^{i-1}\right] = \mathbb{E}\left[\rho W_{i-1} + \tilde{W}_i|Y^{i-1}\right] = \rho Z_{i-1}$, and (H.3b) follows by expanding $\mathbb{E}\left[W|Y^i\right]$ using the innovation sequence and exploiting how Y_i are i.i.d. The value of the parameters β, γ are given by

$$\beta = \sqrt{\frac{L}{C}}, \gamma = \frac{\beta C}{L+\sigma_n^2}. \tag{H.4}$$

Note that from the definition of C in (H.2), $P_{W_i|Z_{i-1}=z_{i-1}} \sim \mathcal{N}(\rho z_{i-1}, C)$. Hence, using (4.38),

$$\begin{aligned} Q_{Z'|Z,W'}(\cdot|z_{i-1}, w_i) &\sim \mathcal{N}(\rho z_{i-1} + \beta\gamma(w_i - \rho z_{i-1}), \gamma^2\sigma_v^2) \\ Q_{Z'|Z}(\cdot|z_{i-1}) &\sim \mathcal{N}(\rho z_{i-1}, \gamma^2(L+\sigma_v^2)). \end{aligned}$$

From Theorem 4.1.5, the linear stationary Markov coordination strategy (4.37) is inverse control optimal for a ρ of the form

$$\begin{aligned} \rho(w_i, z_{i-1}, z_i) \propto_+ & \quad -\log \frac{dQ_{Z'|Z,W'}(\cdot|z_{i-1}, w_i)}{dQ_{Z'|Z}(\cdot|z_{i-1})}(z_i) \\ &= \frac{\left(z_i - \rho z_{i-1} - \beta\gamma(w_i - \rho z_{i-1})\right)^2}{2\gamma^2\sigma_v^2} \\ &- \frac{\left(z_i - \rho z_{i-1}\right)^2}{2\gamma^2(L+\sigma_v^2)} - \log\sqrt{\frac{L+\sigma_v^2}{\sigma_v^2}} \\ & \propto_+ (z_i - w_i)^2 - \frac{\sigma_v^2}{L+\sigma_v^2}(w_i - \rho z_{i-1})^2 \end{aligned} \tag{H.5}$$

where (H.5) follows from (H.4). Similarly, the power-like cost for inverse control optimality is given $\eta(x) \propto_+ D\left(P_{Y|X=x} \| P_Y\right) = D\left(P_V(\cdot - x) \| P_Y(\cdot)\right) \propto_+ x^2$.

Thus we have

$$
\begin{aligned}
& \mathbb{E}\left[\sum_{i=1}^{n} \rho(W_i, Z_{i-1}, Z_i)\right] \\
\propto_{+} \quad & \mathbb{E}\left[\sum_{i=1}^{n} (W_i - Z_i)^2 - \left(\frac{\sigma_v^2}{L+\sigma_v^2}\right)(W_i - \rho Z_{i-1})^2\right] && \text{(H.6)} \\
= \quad & \mathbb{E}\left[\sum_{i=1}^{n} (W_i - Z_i)^2\right] \\
- \quad & \mathbb{E}\left[\left(\frac{\sigma_v^2}{L+\sigma_v^2}\right)\left(\rho W_{i-1} - \rho Z_{i-1} + \tilde{W}_i\right)^2\right] && \text{(H.7)} \\
= \quad & \mathbb{E}\left[\sum_{i=1}^{n} (W_i - Z_i)^2 - \left(\frac{\sigma_v^2\rho^2}{L+\sigma_v^2}\right)(W_{i-1} - \rho Z_{i-1})^2\right] \\
- \quad & \mathbb{E}\left[\left(\frac{\sigma_v^2}{L+\sigma_v^2}\right)\tilde{W}_i^2\right] && \text{(H.8)} \\
= \quad & \mathbb{E}\left[\sum_{i=1}^{n} \left(1 - \frac{\sigma_v^2\rho^2}{L+\sigma_v^2}\right)(W_i - Z_i)^2 - \left(\frac{\sigma_v^2}{L+\sigma_v^2}\right)\tilde{W}_i^2\right] \\
& - \frac{\sigma_v^2\rho^2}{L+\sigma_v^2}\mathbb{E}\left[W_0^2\right] + \frac{\sigma_v^2\rho^2}{L+\sigma_v^2}\mathbb{E}\left[(Z_n - W_n)^2\right] \\
\propto_{+} \quad & \mathbb{E}\left[\sum_{i=1}^{n} (W_i - Z_i)^2\right] + \left(\frac{1}{1 - \frac{\sigma_v^2\rho^2}{L+\sigma_v^2}}\right)\mathbb{E}\left[(Z_n - W_n)^2\right]
\end{aligned}
$$

where (H.6) follows from (4.39b); (H.7) follows from (4.35b); (H.8) follows from (4.35c). □

APPENDIX I

PROOF OF LEMMA 5.4.2

Proof. We will work on the space $\Omega' = \mathsf{W}^{\mathbb{Z}_+} \times \mathsf{W}^{\mathbb{Z}_+} \times \mathsf{Y}^{\mathbb{Z}_+}$, where we write $\tilde{W}_n(w, w', y) = w(n)$, $\tilde{W}'_n(w, w', y) = w'(n)$, and $Y_n(w, w', y) = y(n)$.

Since $||\mathbf{P}^\nu(\tilde{W}_n \in \cdot) - \mathbf{P}^{\overline{\nu}}(\tilde{W}_n \in \cdot)||_{\text{TV}} \overset{n\to\infty}{\longrightarrow} 0$, we can construct a probability measure $\mathbf{Q} : \mathcal{B}\left(\mathsf{W}^{\mathbb{Z}_+} \times \mathsf{Y}^{\mathbb{Z}_+} \times \mathsf{W}^{\mathbb{Z}_+}\right) \to [0, 1]$ such that:

We make use of the well-known fact [75, Theorem III.14.10 and (III.20.7)], that $||\mathbf{P}^\nu(\tilde{W}_n \in \cdot) - \mathbf{P}^{\overline{\nu}}(\tilde{W}_n \in \cdot)||_{\text{TV}} \overset{n\to\infty}{\longrightarrow} 0$ as $n \to \infty$ implies the existence of a successful coupling of the laws of $(\tilde{W}_n)_{n\geq 0}$ under $\mathbf{P}^\nu$ and $\mathbf{P}^{\overline{\nu}}$. We can thus construct a probability measure $\mathbf{Q} : \mathcal{B}\left(\mathsf{W}^{\mathbb{Z}_+} \times \mathsf{Y}^{\mathbb{Z}_+} \times \mathsf{W}^{\mathbb{Z}_+}\right) \to [0, 1]$ such that:

- The law of $(\tilde{W}_n)_{n\geq 0}$ under $\mathbf{Q}$ coincides with the law of $(\tilde{W}_n)_{n\geq 0}$ under $\mathbf{P}^\nu$.
- The law of $(\tilde{W}'_n)_{n\geq 0}$ under $\mathbf{Q}$ coincides with the law of $(\tilde{W}_n)_{n\geq 0}$ under $\mathbf{P}^{\overline{\nu}}$.
- There is a finite random time τ such that a.s. $\tilde{W}_n = \tilde{W}'_n$ for all $n \geq \tau$.

In addition, define a probability kernel $\mathbf{Q}^Y : \mathsf{W}^{\mathbb{Z}_+} \times \mathcal{B}\left(\mathsf{Y}^{\mathbb{Z}_+}\right) \to [0, 1]$ such that $(Y_n)_{n\geq 0}$ are independent under $\mathbf{Q}^Y(w, \cdot)$ and $\mathbf{Q}^Y(w, Y_n \in \cdot) = P_{Y|\tilde{W}}(\cdot|w)$.

Now consider the following probability measures on Ω':

$$\mathbf{Q}_1(A) = \int 1_{\{(w,w',y)\in A\}} \mathbf{Q}^Y(w, dy) \mathbf{Q}(dw, dy, dw') \tag{I.1}$$

$$\mathbf{Q}_2(A) = \int 1_{\{(w,w',y)\in A\}} \mathbf{Q}^Y(w', dy) \mathbf{Q}(dw, dy, dw'). \tag{I.2}$$

It is easily seen that $\mathbf{P}^\nu\Big|_{\mathcal{F}^Y_{0,\infty}} = \mathbf{Q}_1\Big|_{\mathcal{F}^Y_{0,\infty}}$ and $\mathbf{P}^{\overline{\nu}}\Big|_{\mathcal{F}^Y_{0,\infty}} = \mathbf{Q}_2\Big|_{\mathcal{F}^Y_{0,\infty}}$. To complete the proof, it therefore suffices to show that $\mathbf{Q}_1 \sim \mathbf{Q}_2$. It is immediate, however, that

$$\frac{d\mathbf{Q}^Y(w', \cdot)}{d\mathbf{Q}^Y(w, \cdot)} = \prod_{k=0}^{N} \frac{g(w'(k), y(k))}{g(w(k), y(k))}, \qquad \text{whenever } w(n) = w'(n) \text{ for all } n > N \tag{I.3}$$

where $g(z, y)$ is the observation density defined in Defn 2. Thus, evidently

$$\mathbf{Q}_1 \sim \mathbf{Q}_2 \qquad \text{with } \frac{d\mathbf{Q}_2}{d\mathbf{Q}_1} = \prod_{k=0}^{\tau} \frac{g(\tilde{W}'_k, Y_k)}{g(\tilde{W}_k, Y_k)}. \tag{I.4}$$

$\square$

APPENDIX J

PROOF OF LEMMA 5.4.3

Proof. By the Markov property, P and P^S are versions of the regular conditional probabilities $\mathbf{P}(S_1 \in \cdot|\sigma(S_0))$ and $\mathbf{P}(S_1 \in \cdot|\sigma(S_0) \vee \mathcal{F}^Y_{0,\infty})$, respectively. By the Polish assumption, we can also introduce regular conditional probabilities $R : \mathsf{U} \times \mathcal{F}^Y_{0,\infty} \to [0,1]$ and $R^S : \mathsf{U} \times \mathsf{U} \times \mathcal{F}^Y_{0,\infty} \to [0,1]$ of the form $\mathbf{P}((Y_k)_{k\geq 0} \in \cdot|\sigma(S_0))$ and$\mathbf{P}((Y_k)_{k\geq 0} \in \cdot|\sigma(S_0, S_1))$, respectively. Applying [1, Lemma 3.6] to the law of the triple $(S_0, S_1, (Y_k)_{k\geq 0})$, it evidently suffices to show that there is a strictly positive measurable $h : \mathsf{U} \times \Omega^Y \times \mathsf{U} \to (0,\infty)$ such that

$$R^S(z, z', A) = \int 1_{\{A\in y\}} h(z, y, z') R(z, dy) \qquad \forall A \in \mathcal{F}^Y_{0,\infty}$$

for $(z, z') \in H$ with $\mathbf{P}((S_0, S_1) \in H) = 1$.

By a well-known result on kernels [76, Section V.58] there exists a non-negative measurable function $\tilde{h} : \mathsf{U} \times \Omega^Y \times \mathsf{U} \to (0,\infty)$, for all $z, z' \in \mathsf{U}$,

$$R^S(z, z', A) = \int 1_{\{A\in y\}} \tilde{h}(z, y, z') R(z, dy) + R^{\perp}(z, z', A) \qquad \forall A \in \mathcal{F}^Y_{0,\infty},$$

where the kernel $R^{\perp}$ is such that $R^{\perp}(z, z', \cdot) \perp R(z, \cdot)$ for every $z, z' \in \mathsf{U}$. Now suppose we can establish that $R^S(z, z', \cdot) \sim R(z, \cdot)$ for $(z, z') \in H$ with $\mathbf{P}((S_0, S_1) \in H) = 1$. Then $R^{\perp}(z, z', \cdot) = 0$ for $(z, z') \in H$, and $\tilde{h}(z, y, z') > 0$ except on a null set. We can then set $h(z, y, z') = 1$ whenever $\tilde{h}(z, y, z') = 0$, and set $h(z, y, z') = \tilde{h}(z, y, z')$ otherwise; this gives a function h with the desired properties, completing the proof. It thus remains to show that $R^S(z, z', \cdot) \sim R(z, \cdot)$ for $(z, z') \in H$ with $\mathbf{P}((S_0, S_1) \in H) = 1$.

To this end, let us introduce convenient versions of the regular conditional probabilities R and R^S . Note that we may set

$$\int f_0(y(0)) \cdots f_n(y(n)) R^S(z, z', dy) = \int f_0(u) P_{Y|S}(du|z) \times \mathbf{E}^{z'}(f_1(Y_0) \cdots f_n(Y_{n-1}))$$

for all bounded measurable $f_0, \cdots, f_n$ and $n < \infty$. Similarly, we may get

$$
\begin{aligned}
&\int f_0(y(0)) \cdots f_n(y(n)) R(z, dy) \\
&= \int f_0(u) P_{Y|S}(du|z)) \times \int \mathbf{E}^{\tilde{z}}(f_1(Y_0) \cdots f_n(Y_{n-1})) P(z, d\tilde{z}) \\
&= \int f_0(u) P_{Y|S}(du|z)) \times \mathbf{E}^{P(z,\cdot)}(f_1(Y_0) \cdots f_n(Y_{n-1})).
\end{aligned}
$$

It thus suffices to show that

$$
\mathbf{P}^{z'}\Big|_{\mathcal{F}^Y_{0,\infty}} \sim \mathbf{P}^{P(z,\cdot}\Big|_{\mathcal{F}^Y_{0,\infty}} \quad \text{for } (z, z') \in H \text{ with } \mathbf{P}((S_0, S_1) \in H) = 1.
$$

Since $P_{Y|S}$ is non-degenerate and from Lemma 5.4.2, it suffices to show that

$$
||\mathbf{P}^{z'}(S_n \in \cdot) - \mathbf{P}^{P(z,\cdot}(S_n \in \cdot)||_{\text{TV}} \stackrel{n\to\infty}{\longrightarrow} 0
$$

for $(z, z') \in H$ with $\mathbf{P}((S_0, S_1) \in H) = 1$.

Now note that by ergodicity assumption on $\{S_n\}$, we may choose a set H_1 of $\varpi-$full measure such that $||\mathbf{P}^{z'}(S_n \in \cdot) - \varpi||_{\text{TV}} \to 0$ as $n \to \infty$ for all $z \in H_1$. By [1, Lemma 2.6], there is a subset $H_2 \subset H_1$ of ϖ-full measure such that for every $z \in H_2$ we have $\mathbf{P}^z(S_n \in H_2 \text{ for all } n \geq 0) = 1$. In particular, for $(z, z') \in H$, we than have

$$
\begin{aligned}
&||\mathbf{P}^{z'}(S_n \in \cdot) - \mathbf{P}^{P(z,\cdot}(S_n \in \cdot)||_{\text{TV}} \\
&\leq ||\mathbf{P}^{z'}(S_n \in \cdot) - \varpi||_{\text{TV}} + \int ||\mathbf{P}^{z''}(S_n \in \cdot) - \varpi||_{\text{TV}} P(z, dz'') \\
&\stackrel{n\to\infty}{\longrightarrow} 0.
\end{aligned}
$$

But $H = H_2 \times H_2$ satisfies $\mathbf{P}((S_0, S_1) \in H) = 1$ by construction.

□

REFERENCES

[1] R. Van Handel, "The stability of conditional Markov processes and Markov chains in random environments," *The Annals of Probability*, vol. 37, no. 5, pp. 1876–1925, 2009.

[2] C. Papadimitriou and J. Tsitsiklis, "Intractable problems in control theory," *SIAM Journal on Control and Optimization*, vol. 24, no. 4, pp. 639–654, 1986.

[3] T. M. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.

[4] D. Bertsekas and S. Shreve, *Stochastic Optimal Control: The Discrete Time Case*. Athena Scientific, 1978.

[5] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: Lossy source-channel communication revisited," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

[6] R. Kalman, "When is a linear control system optimal?" *Trans. ASME Ser. D:J: Basic Eng*, pp. 1–10, 1964.

[7] K. Körding and D. Wolpert, "The loss function of sensorimotor learning," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 26, p. 9839, 2004.

[8] E. Todorov, "Optimality principles in sensorimotor control," *Nature neuroscience*, vol. 7, no. 9, pp. 907–915, 2004.

[9] C. Baker, J. Tenenbaum, and R. Saxe, "Goal inference as inverse planning," in *Proceedings of the 29th Annual Meeting of the Cognitive Science Society*, 2007, pp. 779–784.

[10] C. Shannon, "A mathematical theory of Communication," *Bell System Technical Journal*, vol. 22, pp. 379–423, 1948.

[11] H. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proceedings of the IEEE*, vol. 59, pp. 1557–1566, 1971.

[12] H. Witsenhausen, “Informational aspects of stochastic control,” *Analysis and Optimization of Stochastic Systems*, pp. 272–284, 1980.

[13] D. Teneketzis, “Communication in decentralized control,” Ph.D. dissertation, Massachusetts Institute of Technology, 1980.

[14] O. Shayevitz and M. Feder, “Communication with feedback via posterior matching,” in *IEEE International Symposium on Information Theory*, June 2007, pp. 391 –395.

[15] O. Shayevitz and M. Feder, “The posterior matching feedback scheme: Capacity achieving and error analysis,” in *IEEE International Symposium on Information Theory*, July 2008, pp. 900–904.

[16] O. Shayevitz and M. Feder, “Optimal feedback communication via posterior matching,” *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1186–1222, March 2011.

[17] M. Horstein, “Sequential transmission using noiseless feedback,” *IEEE Transactions on Information Theory*, vol. 9, no. 3, pp. 136–143, 1963.

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

[19] J. Schalkwijk, “A coding scheme for additive noise channels with feedback–II: Band-limited signals,” *IEEE Transactions on Information Theory*, vol. 12, no. 2, pp. 183–189, 1966.

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

[21] C. Omar, A. Akce, M. Johnson, T. Bretl, R. Ma, E. Maclin, M. McCormick, and T. P. Coleman, “A feedback information-theoretic approach to the design of brain-computer interfaces,” *Int’l Journal on Human-Computer Interaction*, vol. 27, no. 1, pp. 5–23, January 2011, special issue on Current Trends in Brain-Computer Interface (BCI) Research.

[22] R. Elliott, L. Aggoun, and J. Moore, *Hidden Markov Models: Estimation and Control*. Springer, 1995.

[23] F. P. Kelly, *Reversibility and Stochastic Networks*. New York: Kluwer Academic Publishers, 1979.

[24] R. G. Gallager, *Discrete Stochastic Processes*, 1st ed. Boston: Kluwer Academic Publishers, 1996.

[25] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*, 2nd ed. New York: Cambridge University Press, 2009.

[26] J. Alvarez and B. Hajek, "Equivalence of *trans* paths in ion channels," *Phys. Rev. E*, vol. 73, no. 4, pp. 1–12, Apr 2006.

[27] M. Propp, "The thermodynamic properties of Markov processes," Ph.D. dissertation, M.I.T, 1985.

[28] S. Mitter and N. Newton, "Information and entropy flow in the Kalman–Bucy filter," *Journal of Statistical Physics*, vol. 118, no. 1, pp. 145–176, 2005.

[29] J. Hsu and P. Burke, "Behavior of tandem buffers with geometric input and Markovian output," *IEEE Transactions on Communications*, vol. 24, no. 3, pp. 358–361, 1976.

[30] V. Anantharam and S. Verdú, "Bits through queues," *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 4–18, 1996.

[31] R. Sundaresan and S. Verdú, "Capacity of queues via point-process channels," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2697–2709, June 2006.

[32] T. P. Coleman, "A simple memoryless proof of the capacity of the exponential server timing channel," in *Proceedings of the IEEE Information Theory Workshop (ITW)*, Volos, Greece, June 2009, pp. 101–105.

[33] T. P. Coleman, "A stochastic control approach to 'posterior matching'-style feedback communication schemes," in *IEEE International Symposium on Information Theory*, Seoul, Korea, July 2009, pp. 1520–1524.

[34] S. Tatikonda, "Control under communication constraints," Ph.D. dissertation, Massachusetts Institute of Technology, 2000.

[35] R. Bansal and T. Basar, "Simultaneous design of communication and control strategies for stochastic systems with feedback," *Automatica*, vol. 45, no. 5, pp. 679–694, 1989.

[36] D. Blackwell, *Information Theory, in Modern Mathematics for the Engineer*, 2nd ed. McGraw-Hill, 1961.

[37] R. Ahlswede and A. Kaspi, "Optimal coding strategies for certain permuting channels," *IEEE Trans. on Information Theory*, vol. 33, no. 3, pp. 310–314, 1987.

[38] H. Permuter, P. Cuff, B. Roy, and T. Weissman, "Capacity of the trapdoor channel with feedback," *IEEE Trans. on Information Theory*, pp. 3150–65, July, 2008.

[39] R. McEliece, *The Theory of Information and Coding*. Cambridge Univ Press, 2002.

[40] R. Kalman, "When is a linear control system optimal?" *Transactions of the ASME, Journal of Basic Engineering, Series D*, vol. 86, pp. 81–90, 1964.

[41] J. Casti, "On the general inverse problem of optimal control theory," *Journal of Optimization theory and applications*, vol. 32, no. 4, pp. 491–497, December 1980.

[42] H. Deng and M. Krstic, "Stochastic nonlinear stabilization—part ii: Inverse optimality," *Systems and Control Letters*, vol. 32, pp. 151–159, 1997.

[43] K. Dvijotham and E. Todorov, "Inverse optimal control with linearly-solvable MDPs," *International Conference on Machine Learning*, vol. 27, pp. 335–342, 2010.

[44] R. Atar and O. Zeitouni, "Exponential stability for nonlinear filtering," *Ann. Inst. H. Poincare Probab. Statist.*, vol. 33, no. 6, pp. 697–725, 1997.

[45] A. Budhiraja and D. Ocone, "Exponential stability of discrete-time filters for bounded observation noise," *Systems Control Lett.*, vol. 30, no. 4, pp. 185–193, 1997.

[46] F. LeGland and N. Oudjane, "A robustification approach to stability and to uniform particle approximation of nonlinear filters: the example of pseudo-mixing signals," *Stochastic Process. Appl.*, vol. 106, no. 2, pp. 279–316, 2003.

[47] F. LeGland and N. Oudjane, "Stability and uniform approximation of nonlinear filters using the Hilbert metric and application to particle filters," *Ann. Appl. Probab.*, vol. 14, no. 1, pp. 144–187, 2004.

[48] R. Atar and O. Zeitouni, "Lyapunov exponents for finite state nonlinear filtering," *SIAM J. Control Optim.*, vol. 35, no. 1, pp. 36–55, 1997.

[49] A. Budhiraja and D. Ocone, "Lyapunov exponents for finite state nonlinear filtering," *Stochastic Process. Appl.*, vol. 82, no. 2, pp. 245–257, 1999.

[50] P. Chigansky, "Stability of the nonlinear filter for slowly switching Markov chains," *Stochastic Process. Appl.*, vol. 116, no. 8, pp. 1185–1194, 2006.

[51] P. Chigansky, "An ergodic theorem for filtering with applications to stability," *Systems Control Lett.*, vol. 55, no. 11, pp. 908–917, 2006.

[52] P. Chigansky, R. Liptser, and R. Van Handel, "Intrinsic methods in filter stability," in *Oxford Handbook of Nonlinear Filtering*. Oxford University Press, 2011.

[53] N. Elia, "When Bode meets Shannon: Control-oriented feedback communication schemes," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1477–1488, 2004.

[54] S. Mitter and N. Newton, "A variational approach to nonlinear estimation," *SIAM Journal on Control and Optimization*, vol. 42, no. 5, pp. 1813–1833, 2004.

[55] A. Mahajan and D. Teneketzis, "On the design of globally optimal communication strategies for real-time noisy communication systems with noisy feedback," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 580–595, 2008.

[56] M. Anand and P. R. Kumar, "Estimating the state of a Markov chain over a noisy communication channel: a bound and an encoder," in *IEEE Conference on Decision and Control*, Atlanta, GA, December 2010, to appear.

[57] N. Merhav and M. Feder, "Universal prediction," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2124–2147, 2002.

[58] A. Aiyer, K. Pyun, Y. Huang, D. O'Brien, and R. Gray, "Lloyd clustering of Gauss mixture models for image compression and classification," *Signal Processing: Image Communication*, vol. 20, no. 5, pp. 459–485, 2005.

[59] M. S. McCormick, R. Ma, and T. P. Coleman, "An analytic spatial filter and hidden Markov model for enhanced information transfer rate in EEG-based brain computer interfaces," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, TX, 2010, pp. 602–605.

[60] S. Wills and D. MacKay, "DASHER-an efficient writing system for brain-computer interfaces?" *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 14, no. 2, pp. 244–246, 2006.

[61] A. Akce, M. Johnson, and T. Bretl, "Remote teleoperation of an unmanned aircraft with a brain-machine interface: Theory and preliminary results," in *Robot. and Autom. Proceedings. IEEE Int. Conf. on*, 2010, pp. 5322–5327.

[62] A. Mahajan, "Sequential decomposition of sequential dynamic teams: applications to real-time communication and networked control systems," Ph.D. dissertation, University of Michigan, 2008.

[63] D. Bertsekas, *Nonlinear Programming*. Athena Scientific Belmont, MA, 1999.

[64] Y. Kifer, *Ergodic Theory of Random Transformations*. Boston: Birkhäuser, 1986.

[65] I. Csiszár, "On an extremum problem of information theory," *Studia Scientiarum Mathematicarum Hungarica*, vol. 9, no. 1, pp. 57–71, 1974.

[66] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Academic Press, Inc. Orlando, FL, USA, 1982.

[67] M. Gastpar, "To Code or Not to Code," Ph.D. dissertation, École Polytechnique Fédérale de Lausanne (EPFL), 2002.

[68] J. Martin, "Batch queues, reversibility and first-passage percolation," *Queueing Systems*, vol. 62, no. 4, pp. 411–427, 2009.

[69] S. Gorantla and T. Coleman, "On reversible Markov chains and maximization of directed information," in *Proceedings of IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2010, pp. 216–220.

[70] L. Rogge, "Uniform inequalities for conditional expectations," *The Annals of Probability*, vol. 2, no. 3, pp. 486–489, 1974.

[71] H. Mukerjee, "Almost sure equiconvergence of conditional expectations," *The Annals of Probability*, vol. 12, no. 3, pp. 733–741, Aug 1984.

[72] A. E. Gamal, "The feedback capacity of degraded broadcast channels," *IEEE Transactions on Information Theory*, vol. 24, no. 6, pp. 379–381, 1978.

[73] A. Nayyar, A. Mahajan, and D. Teneketzis, "Optimal control strategies in delayed sharing information structures," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1606–1620, July 2011.

[74] R. Dudley, *Real Analysis and Probability*. Cambridge Univ Press, 2002.

[75] T. Lindvall, *Lectures on the Coupling Method*. Wiley, New York, 1992.

[76] C. Dellacherie and P.-A. Meyer, *Probabilities and Potential*. B. North-Holland Mathematics Studies 72. North-Holland, Amsterdam, 1982.