RATIONAL INATTENTION IN CONTROL OF MARKOV CHAINS

BY

EHSAN SHAFIEEPOORFARD

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Adviser:

Assistant Professor Maxim Raginsky

# ABSTRACT

This thesis poses a general model for optimal control subject to information constraint, motivated in part by recent work on information-constrained decision-making by economic agents.

In the average-cost optimal control framework, the general model introduced in this paper reduces to a variant of the linear-programming representation of the average-cost optimal control problem, subject to an additional mutual information constraint on the randomized stationary policy. The resulting infinite-dimensional convex program admits a decomposition based on the Bellman error, which is the subject of study in approximate dynamic programming.

Later, we apply the general theory to an information-constrained variant of the scalar Linear-Quadratic-Gaussian (LQG) control problem. We give an upper bound on the optimal steady-state value of the quadratic performance objective and present explicit constructions of controllers that achieve this bound. We show that the obvious certainty-equivalent control policy is suboptimal when the information constraints are very severe, and propose another policy that performs better in this low-information regime. In the two extreme cases of no information (open-loop) and perfect information, these two policies coincide with the optimum.

*To my parents, for their love and support.*

# ACKNOWLEDGMENTS

Thanks to Prof. Maxim Raginsky for several hints and contributions in derivations of the results. Thanks to my former advisor, Prof. Sean Meyn, for his valuable comments and intuitions.

# TABLE OF CONTENTS

# CHAPTER 1

# INTRODUCTION

In many applications of stochastic dynamic programming, the controller has access only to limited information about the state of the system.

Unlike much of the existing literature on problems with imperfect state information, in this paper it is assumed that the system designer has to decide not only about the control policy, but also about the observation channel based on which the control is derived. There are various applications which fit in this framework, either in engineering or economics.

In economic decision making, the amount of information required to make a truly optimal decision will typically exceed what an agent can handle. In his seminal work [1, 2], Christopher Sims [1] adds an information-processing constraint to a specific kind of dynamic programming problem which is frequently used in macroeconomic models. Sims uses the term "rational inattention" to describe the setting in which information-constrained agents strive to make the best use of whatever information they are able to handle. Sims considers a model in which a representative agent decides about his consumption over subsequent periods of time, while his computational ability to reckon his wealth – the state of the dynamic system – is limited. A special case is considered in which income in one period adds uncertainty of wealth in the next period. Other modeling assumptions reduce the model to an LQG control problem.

Quantitatively, the constraint on observation channel is stated in terms of an upper bound on the *mutual information* in the sense of Shannon [3] between the state of the system and the observation available to the agent. As one justification for introducing the information constraint, Sims remarks [2] that "most people only vaguely aware of their net worth, are little influenced in their current behavior by the status of their retirement account, and can

[1] Christopher Sims shared the 2011 Nobel Memorial Prize in Economics with Thomas Sargent.

be induced to make large changes in savings behavior by minor informational changes, like changes in default options on retirement plans." From a broader perspective, as he expresses, this model aims to support John Maynard Keynes' well-known view that real economic behavior is inconsistent with the idea of continuously optimizing agents interacting in continuously clearing markets.

Extensions of this work include Matejka and McKay [4], who recovered the well-known multinomial logit model for a situation where an individual must choose among discrete alternatives yielding different values not perfectly known ex-ante. In the sequel [5], they extend this model to study Nash equilibria in a Bertrand oligopoly setting where $N$ firms produce the same commodity with different qualities not perfectly known ex-ante; it is shown that this information friction leads to increased prices for the commodity. On a similar theme, Peng [6] is motivated by the observation that "*investors have limited time and attention to process information.*" In a factor model for asset returns, he investigates the dynamics of prices and mutual information when there is uncertainty in the decision making process. Other works using this setup in economics include [6–9]. These works have offered compelling information-theoretic explanations of certain empirically observed features of economic behavior of individuals, firms or institutions; however, most of them rely on heuristic considerations or on simplifying assumptions pertaining to the structure of observation channels.

On the other hand, a parallel line of research on dynamical decision-making with limited information can be found in the control theory literature (a very partial list of references is [10–15]).

Given the appeal and generality of these questions, there is ample motivation for the creation of a general theory for optimal control subject to information constraints. The contribution of this thesis is to initiate the development of such a theory, which would in turn enable us to address many problems in macroeconomics and engineering in a systematic fashion.

In this thesis, we focus on the average-cost optimal control framework and show that the construction of an optimal information-constrained controller reduces to a variant of the linear-programming representation of the average-cost optimal control problem, subject to an additional mutual information constraint on the randomized stationary policy. The resulting optimization problem is convex, and admits a decomposition in terms of the

Bellman error, which is the subject of study in approximate dynamic programming. This decomposition reveals a fundamental connection between information-constrained controller design and *rate-distortion theory* [16], a branch of information theory that deals with optimal compression of data subject to information constraints.

The theoretical methodology developed is then used to analyze the classic linear-quadratic-Gaussian (LQG) control problem [17, 18] in the rational inattention framework. Various information- or communication-constrained versions of the LQG problem have been studied in the literature (see, e.g. [1, 11, 13, 14]). In particular, Sims [1] constructed an information-constrained control law for the LQG problem with discounted cost. His solution relies on the certainty equivalence principle — let the control be the same linear function of a suitable noisy state estimate as one would use in the perfect-information case, and then optimize the observation channel to satisfy the information constraint in steady state. However, the derivation in [1] is based on several *ad hoc* assumptions and leaves open the question of closed-loop stability when the information constraint is so severe that the control must be nearly independent of the state.

The next contribution is explicit construction of rationally inattentive control laws for the LQG problem from first principles, using the convex-analytic approach we have developed. In particular, we show the following:

1. If the controlled linear system is open-loop stable, then the certainty-equivalent control law of the type proposed by Sims [1] induces stable closed-loop dynamics for all values of the mutual information constraint.

2. This control law is *suboptimal* in the regime of very low information. In this regime, it is outperformed by another control law that has similar structure (a linear noisy observation channel followed by linear gain), but both the linear gain and the noise characteristics of the channel depend explicitly on the value of the information constraint.

3. When the controlled system is unstable, we give a simple sufficient condition (lower bound) on the value of information constraint to guarantee that the certainty-equivalent control law will stabilize the system.

# CHAPTER 2

# PRELIMINARIES AND NOTATION

All spaces are assumed to be standard Borel (i.e., isomorphic to a Borel subset of a complete separable metric space), and will be equipped with their Borel $\sigma$-algebras. If $\mathsf{X}$ is such a space, then $\mathcal{B}(\mathsf{X})$ will denote its Borel $\sigma$-algebra, and $\mathcal{P}(\mathsf{X})$ will denote the space of all probability measures on $(\mathsf{X}, \mathcal{B}(\mathsf{X}))$. We will denote by $M(\mathsf{X})$ the space of all measurable functions $f : \mathsf{X} \to \mathbb{R}$ and by $C_b(\mathsf{X}) \subseteq M(\mathsf{X})$ the space of all bounded continuous functions. We will often use bilinear form notation for expectations: for any $f \in L^1(\mu)$,

$$\langle \mu, f \rangle \triangleq \int_{\mathsf{X}} f(x)\mu(\mathrm{d}x) = \mathbb{E}[f(X)],$$

where in the last expression it is understood that $X$ is an $\mathsf{X}$-valued random object with $\mathrm{Law}(X) = \mu$.

Given two spaces $\mathsf{X}$ and $\mathsf{Y}$, a mapping $K(\cdot|\cdot) : \mathcal{B}(\mathsf{Y}) \times \mathsf{X} \to [0,1]$ is a *Markov* (or *stochastic*) *kernel* if $K(\cdot|x) \in \mathcal{P}(\mathsf{Y})$ for all $x \in \mathsf{X}$ and $x \mapsto K(B|x)$ is measurable for every $B \in \mathcal{B}(\mathsf{Y})$. The space of all such Markov kernels will be denoted by $\mathcal{M}(\mathsf{Y}|\mathsf{X})$. Markov kernels $K \in \mathcal{M}(\mathsf{Y}|\mathsf{X})$ act on measurable functions $f \in M(\mathsf{Y})$ from the left as

$$Kf(x) \triangleq \int_{\mathsf{Y}} f(y)K(\mathrm{d}y|x), \qquad \forall x \in \mathsf{X}$$

and on probability measures $\mu \in \mathcal{P}(\mathsf{X})$ from the right as

$$\mu K(B) \triangleq \int_{\mathsf{X}} K(B|x)\mu(\mathrm{d}x), \qquad \forall B \in \mathcal{B}(\mathsf{Y}).$$

Moreover, $Kf \in M(\mathsf{X})$ for any $f \in M(\mathsf{Y})$, and $\mu K \in \mathcal{P}(\mathsf{Y})$ for any $\mu \in \mathcal{P}(\mathsf{X})$. The *relative entropy* (or *information divergence*) between any two $\mu, \nu \in \mathcal{P}(\mathsf{X})$

[3] is defined as

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

Given any probability measure $\mu \in \mathcal{P}(\mathsf{X})$ and any Markov kernel $K \in \mathcal{M}(\mathsf{Y}|\mathsf{X})$, we can define a probability measure $\mu \otimes K$ on the product space $(\mathsf{X} \times \mathsf{Y}, \mathcal{B}(\mathsf{X}) \otimes \mathcal{B}(\mathsf{Y}))$ via its action on the rectangles $A \times B$, $A \in \mathcal{B}(\mathsf{X}), B \in \mathcal{B}(\mathsf{Y})$:

$$(\mu \otimes K)(A \times B) \triangleq \int_A K(B|x)\mu(\mathrm{d}x).$$

Note that $\mu \otimes K(\mathsf{X} \times B) = \mu K(B)$ for all $B \in \mathcal{B}(\mathsf{X})$. The *Shannon mutual information* [3] in the pair $(\mu, K)$ is

$$I(\mu, K) \triangleq D(\mu \otimes K \| \mu \otimes \mu K), \tag{2.1}$$

where, for any $\mu \in \mathcal{P}(\mathsf{X})$ and $\nu \in \mathcal{P}(\mathsf{Y})$, $\mu \otimes \nu$ denotes the product measure defined via $(\mu \otimes \nu)(A \times B) \triangleq \mu(A)\nu(B)$ for all $A \in \mathcal{B}(\mathsf{X}), B \in \mathcal{B}(\mathsf{Y})$. The functional $I(\mu, K)$ is concave in $\mu$ and convex in $K$. If $(X, Y)$ is a pair of random objects with $\mathrm{Law}(X, Y) = \Gamma = \mu \otimes K$, then we will also write $I(X; Y)$ or $I(\Gamma)$ for $I(\mu, K)$.

Finally, given a triple of jointly distributed random objects $(X, Y, Z)$ with $\Gamma = \mathrm{Law}(X, Y, Z)$, we will say that they form a *Markov chain* $X \to Y \to Z$ if there exist Markov kernels $K_1 \in \mathcal{M}(\mathsf{Y}|\mathsf{X})$ and $K_2 \in \mathcal{M}(\mathsf{Z}|\mathsf{Y})$ so that $\Gamma$ can be disintegrated as $\Gamma = \mu \otimes K_1 \otimes K_2$ (in words, if $X$ and $Z$ are conditionally independent given $Y$). The mutual information satisfies the *data processing inequality*: if $X \to Y \to Z$ is a Markov chain, then

$$I(X; Z) \le I(X; Y). \tag{2.2}$$

In words, no additional processing can increase information.

# CHAPTER 3

# SOME GENERAL CONSIDERATIONS

Consider a model in which the controller is constrained to observe the system being controlled through an information-limited channel. This is illustrated in the block diagram shown in Figure 3.1, consisting of:

- the (time-invariant) controlled system, specified by an initial condition $\mu \in \mathcal{P}(\mathsf{X})$ and a stochastic kernel $Q \in \mathcal{M}(\mathsf{X}|\mathsf{X} \times \mathsf{U})$, where $\mathsf{X}$ is the state space and $\mathsf{U}$ is the control (or action) space;

- the observation channel, specified by a sequence $\underline{W}$ of stochastic kernels $W_t \in \mathcal{M}(\mathsf{Z}|\mathsf{X}^t \times \mathsf{Z}^{t-1} \times \mathsf{U}^{t-1})$, $t = 1, 2, \ldots$, where $\mathsf{Z}$ is some observation space; and

- the feedback controller, specified by a sequence $\underline{\Phi}$ of stochastic kernels $\Phi_t \in \mathcal{M}(\mathsf{U}|\mathsf{Z})$, $t = 1, 2, \ldots$.

The $\mathsf{X}$-valued state process $\{X_t\}_{t=1}^{\infty}$, the $\mathsf{Z}$-valued observation process $\{Z_t\}_{t=1}^{\infty}$, and the $\mathsf{U}$-valued control process $\{U_t\}_{t=1}^{\infty}$ are defined on a common probabil-

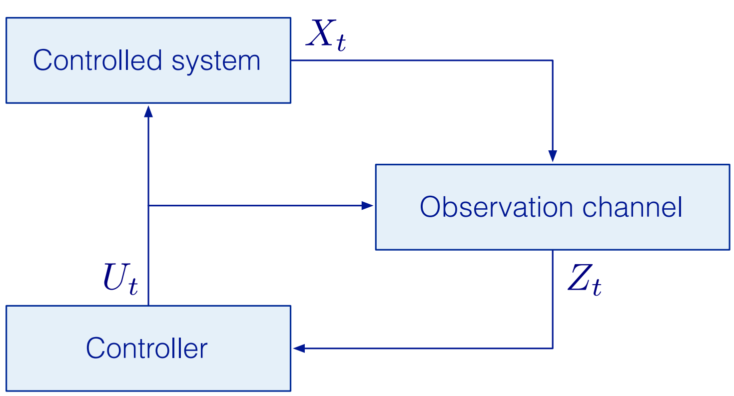


Figure 3.1: System model.

ity space $(\Omega, \mathcal{F}, \mathbb{P})$ and have the causal ordering

$$X_1, Z_1, U_1, \dots, X_t, Z_t, U_t, \dots \tag{3.1}$$

where, $\mathbb{P}$-almost surely, $\mathbb{P}(X_1 \in A) = \mu(A)$ for all $A \in \mathcal{B}(\mathsf{X})$, and for all $t = 1, 2, \dots$, $B \in \mathcal{B}(\mathsf{Z})$, $C \in \mathcal{B}(\mathsf{U})$, $D \in \mathcal{B}(\mathsf{X})$ we have

$$\mathbb{P}(Z_t \in B | X^t, Z^{t-1}, U^{t-1}) = W_t(B | X^t, Z^{t-1}, U^{t-1}) \tag{3.2a}$$

$$\mathbb{P}(U_t \in C | X^t, Z^t, U^{t-1}) = \Phi_t(C | Z_t) \tag{3.2b}$$

$$\mathbb{P}(X_{t+1} \in D | X^t, Z^t, U^t) = Q(D | X_t, U_t). \tag{3.2c}$$

This specification ensures that, for each $t$, the next state $X_{t+1}$ is conditionally independent of $X^{t-1}, Z^t, U^{t-1}$ given $X_t, U_t$ (which is the usual case of a controlled Markov process), and that the control $U_t$ is conditionally independent of $X^t, Z^{t-1}, U^{t-1}$ given $Z_t$. In other words, at each time $t$ the controller takes as input only the most recent observation $Z_t$, which amounts to the assumption that there is a *separation structure* between the observation channel and the controller. This assumption is common in the literature [10, 12, 13].

In the spirit of the rational inattention framework, we assume that the amount of information flow that can be maintained across the observation channel per time step is constrained, and wish to *design* a suitable channel $\underline{W}$ and a controller $\underline{\Phi}$ to minimize a given performance objective under the information constraint. For maximum flexibility, we grant the system designer the freedom to choose the observation space $\mathsf{Z}$ as well. In other words, the designer is allowed to choose an optimal *representation* for the data supplied to the controller.

As we will demonstrate shortly, the choice $\mathsf{Z} = \mathcal{P}(\mathsf{X})$ (i.e., letting $\mathsf{Z}$ be the space of beliefs on the state space) is information-theoretically optimal. For now, let us see how much insight we can gain in the case of a fixed $\mathsf{Z}$. Then the problem of *optimal control with rational inattention* can be formulated as follows. Let $c : \mathsf{X} \times \mathsf{U} \to \mathbb{R}^+$ be a given measurable one-step state-action cost function. Given a fixed finite horizon $T$ and some $R \geq 0$, we seek an observation channel $\underline{W}$ and a controller $\underline{\Phi}$ in order to

$$\text{minimize } \mathbb{E}\left[\sum_{t=1}^{T} c(X_t, U_t)\right] \tag{3.3a}$$

$$\text{subject to } I(X_t; Z_t) \le R,\ t = 1, 2, \ldots, T. \tag{3.3b}$$

Later we focus on a related average-cost performance criterion.

## 3.1 Key structural result

The optimization problem (3.3) seems formidable: for each time step $t = 1, \ldots, T$ we must design stochastic kernels $W_t(\mathrm{d}z_t|x^t, z^{t-1}, u^{t-1})$ and $\Phi_t(\mathrm{d}u_t|z_t)$ for the observation channel and the controller, and the complexity of the feasible set of $W_t$'s grows with $t$. However, the facts that (a) both the controlled system and the controller are Markov, and (b) the cost function at each stage depends only on the current state-action pair, permit a drastic simplification — at each time $t$, we can limit our search to *memoryless* channels $W_t(\mathrm{d}z_t|x_t)$ without impacting either the average cost in (3.3a) or the information flow constraint in (3.3b):

**Theorem 3.1.1 (Memoryless observation channels suffice)** *Under the specified information pattern* (3.2)*, for any controller specification $\underline{\Phi}$ and any channel specification $\underline{W}$, there exists another channel specification $\underline{W}'$ consisting of stochastic kernels $W_t(\mathrm{d}z_t|x_t)$, $t = 1, 2, \ldots$, such that*

$$\mathbb{E}\left[\sum_{t=1}^{T} c(X'_t, U'_t)\right] = \mathbb{E}\left[\sum_{t=1}^{T} c(X_t, U_t)\right]$$

*and*

$$I(X'_t; Z'_t) = I(X_t; Z_t), \qquad t = 1, 2, \ldots, T$$

*where $\{(X_t, U_t, Z_t)\}$ is the original process with $(\mu, Q, \underline{W}, \underline{\Phi})$, while $\{X'_t, U'_t, Z'_t)\}$ is the one with $(\mu, Q, \underline{W}', \underline{\Phi})$.*

**Proof** To prove the theorem, we follow the approach used by Wistenhausen in [19]. We start with the following simple observation that can be regarded as an instance of the Shannon–Mori–Zwanzig Markov model [20]:

**Principle of Irrelevant Information.** Let $\Xi, \Theta, \Psi, \Upsilon$ be four random variables defined on a common probability space, such that $\Upsilon$ is conditionally independent of $(\Theta, \Xi)$ given $\Psi$. Then there exist four random variables $\Xi', \Theta', \Psi', \Upsilon'$ defined on the same spaces as the original tuple, such that $\Xi' \to \Theta' \to \Psi' \to \Upsilon'$ is a *Markov chain*, and moreover the bivariate marginals agree:

$$\begin{aligned}
\mathrm{Law}(\Xi, \Theta) &= \mathrm{Law}(\Xi', \Theta') \\
\mathrm{Law}(\Theta, \Psi) &= \mathrm{Law}(\Theta', \Psi') \\
\mathrm{Law}(\Psi, \Upsilon) &= \mathrm{Law}(\Psi', \Upsilon').
\end{aligned}$$

Using this principle, we can prove the following two lemmas (see Appendices for the proofs):

**Lemma 3.1.2 (Two-Stage Lemma)** *Suppose* $T = 2$*. Then the kernel* $W_2(\mathrm{d}z_2|x^2, z_1, u_1)$ *can be replaced by another kernel* $W_2'(\mathrm{d}z_2|x_2)$*, such that the resulting variables* $(X_t', Z_t', U_t')$*,* $t = 1, 2$*, satisfy*

$$\mathbb{E}[c(X_1', U_1') + c(X_2', U_2')] = \mathbb{E}[c(X_1, U_1) + c(X_2, U_2)]$$

*and* $I(X_t'; Z_t') = I(X_t; Z_t)$*,* $t = 1, 2$*.*

**Lemma 3.1.3 (Three-Stage Lemma)** *Suppose* $T = 2$*, and* $Z_3$ *is conditionally independent of* $(X_i, Z_i, U_i)$*,* $i = 1, 2$*, given* $X_3$*. Then the kernel* $W_2(\mathrm{d}z_2|x^2, z_1, u_1)$ *can be replaced by another kernel* $W_2'(\mathrm{d}z_2|x_2)$*, such that the resulting variables* $(X_i', Z_i', U_i')$*,* $i = 1, 2, 3$*, satisfy*

$$\mathbb{E}\left[\sum_{t=1}^{3} c(X_t', U_t')\right] = \mathbb{E}\left[\sum_{t=1}^{3} c(X_t, U_t)\right]$$

*and* $I(X_t'; Z_t') = I(X_t; Z_t)$ *for* $t = 1, 2, 3$*.*

**Proof** Armed with these two lemmas, we can now prove the theorem by backward induction and grouping of variables. Fix any $T$. By the Two-Stage-Lemma, we may assume that $W_T$ is memoryless, i.e., $Z_T$ is conditionally independent of $X^{T-1}, Z^{T-1}, U^{T-1}$ given $X_T$. Now we apply the Three-Stage

Lemma to

$$\Big|\underbrace{X^{T-3},Z^{T-3},U^{T-3},X_{T-2}}_{\substack{\text{Stage 1}\\ \text{state}}},\ \underbrace{Z_{T-2}}_{\substack{\text{Stage 1}\\ \text{observation}}},\underbrace{U_{T-2}}_{\substack{\text{Stage 1}\\ \text{control}}}\Big|$$
$$\Big|\underbrace{X_{T-1}}_{\substack{\text{Stage 2}\\ \text{state}}},\ \underbrace{Z_{T-1}}_{\substack{\text{Stage 2}\\ \text{observation}}},\underbrace{U_{T-1}}_{\substack{\text{Stage 2}\\ \text{control}}}\Big|\underbrace{X_T}_{\substack{\text{Stage 3}\\ \text{state}}},\ \underbrace{Z_T}_{\substack{\text{Stage 3}\\ \text{observation}}},\ \underbrace{U_T}_{\substack{\text{Stage 3}\\ \text{control}}}\Big| \tag{3.4}$$

to replace $W_{T-1}(\mathrm{d}z_{T-1}|x^{T-1},z^{T-2},u^{T-2})$ with $W'_{T-1}(\mathrm{d}z_{T-1}|x_{T-1})$ without affecting the expected cost or the mutual information between the state and the observation at time $T-1$. We proceed inductively by merging the second and the third stages in (3.4), splitting the first stage in (3.4) into two, and then applying the Three-Stage Lemma to replace the original observation kernel $W_{T-2}$ with a memoryless one.

## 3.2 Long-term average cost criterion

Despite the simplification afforded by Theorem 3.1.1, the optimization problem (3.3) is still difficult even when the observation space $\mathsf{Z}$ is fixed, and the only general way of solving it is via infinite-dimensional dynamic programming.

Our goal is to gain theoretical insight into the structure of optimal control policies in the rational inattention framework. To that end, we make several simplifications:

1. We replace the finite-horizon cost criterion in (3.3a) with the one based on the long-term average cost.

2. We consider only stationary (time-invariant) observation channels and controllers.

3. Instead of separately optimizing over the observation space, the observation channel, and the controller, we collapse these decision variables into the choice of a Markov randomized stationary (MRS) control law $\Phi \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ satisfying the information constraint.

Of these, Item 3) requires some explanation; in essence, it is justified by the simple fact that the mutual information $I(X;U)$ between two random variables $(X, U)$ can be expressed as

$$I(X;U) = \inf \Big\{ I(X;Z) : X \to Z \to U \Big\},$$

where the infimum is over all standard Borel spaces $\mathsf{Z}$ and all $\mathsf{Z}$-valued random objects $Z$ jointly distributed with $(X, U)$, such that $X \to Z \to U$ forms a Markov chain. Indeed, for any such triple we have $I(X;U) \le I(X;Z)$ by the data processing inequality; the other direction follows by considering the "degenerate" Markov chain $X \to U \to U$. Consequently, for any choice of $\mu \in \mathcal{P}(\mathsf{X})$, $\mathsf{Z}$, $W \in \mathcal{M}(\mathsf{Z}|\mathsf{X})$ and $\Phi \in \mathcal{M}(\mathsf{U}|\mathsf{Z})$, the kernel $\Psi = \Phi \circ W \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ will satisfy $I(\mu, \Psi) \le I(\mu, W)$ (recall notation from (2.1)). Conversely, we can always factor any given $\Psi \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ through some standard Borel space $\mathsf{Z}$ as $\Psi = \Phi \circ W$ with $\Phi \in \mathcal{M}(\mathsf{U}|\mathsf{Z})$ and $W \in \mathcal{M}(\mathsf{Z}|\mathsf{X})$, such that $I(\mu, W) = I(\mu, \Psi)$.

In view of the above, we focus on the following problem: find an MRS control law $\Phi \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ to

$$\text{minimize } \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}\left[ \sum_{t=1}^{T} c(X_t, U_t) \right] \tag{3.5a}$$

$$\text{subject to } \limsup_{t \to \infty} I(X_t; U_t) \le R, \ t = 1, 2, \ldots \tag{3.5b}$$

# CHAPTER 4

# ONE-STAGE PROBLEM: SOLUTION VIA RATE-DISTORTION THEORY

Before we analyze the infinite-horizon problem (3.5), let us show that the one-stage case can be solved completely using *rate-distortion theory* [16] (a branch of information theory that deals with optimal compression of data subject to information constraints). To the best of our knowledge, this solution is originally due to Stratonovich [21] (see also [22, Sec. 9.7]). Because our subsequent development builds on these ideas, we briefly describe them here. Moreover, this analysis will provide additional justification for eliminating the decision variables $\mathsf{Z}$ and $W \in \mathcal{M}(\mathsf{Z}|\mathsf{X})$ in favor of an information-constrained controller $\Phi \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ directly connected to the system being controlled.

When $T = 1$, the problem in (3.3) becomes

$$\text{minimize } \mathbb{E}\left[c(X,U)\right] \tag{4.1a}$$

$$\text{subject to } I(X;Z) \le R \tag{4.1b}$$

for a given law $\mu \in \mathcal{P}(\mathsf{X})$ for $X$, where the minimization is over all observation channels $W \in \mathcal{M}(\mathsf{Z}|\mathsf{X})$ and controllers $\Phi \in \mathcal{M}(\mathsf{U}|\mathsf{Z})$. If we denote the optimum value attained in (4.1) by $V(R,\mathsf{Z})$, then the quantity of interest is

$$V(R) \triangleq \inf_{\mathsf{Z}} V(R,\mathsf{Z}), \tag{4.2}$$

where the infimum is over all standard Borel observation spaces. In other words, we seek an observation space $\mathsf{Z}^*$, an observation channel $W^* \in \mathcal{M}(\mathsf{Z}^*|\mathsf{X})$, and a controller $\Phi^* \in \mathcal{M}(\mathsf{U}|\mathsf{Z}^*)$, such that $I(X;Z^*) \le R$ and the resulting expected cost $\mathbb{E}[c(X,U^*)]$ is minimized.

In order to properly frame the main result of [21], we need to make some preliminary observations. If we fix $\mathsf{Z}$ and an observation channel $W \in \mathcal{M}(\mathsf{Z}|\mathsf{X})$, then the optimal choice of the controller $\Phi^*_W \in \mathcal{M}(\mathsf{U}|\mathsf{Z})$

is to let $\Phi^*_W(\mathrm{d}u|z)$ be supported on the set of all $u^* \in \mathsf{U}$, such that

$$\mathbb{E}[c(X,u^*)|Z=z] = \min_{u\in\mathsf{U}} \mathbb{E}[c(X,u)|Z=z].$$

In fact, under suitable assumptions on $c$, there exists a deterministic *measurable selector* $\varphi^*_W : \mathsf{Z} \to \mathsf{U}$, so that

$$\mathbb{E}[c(X,\varphi^*_W(z))|Z=z] = \min_{u\in\mathsf{U}} \mathbb{E}[c(X,u)|Z=z].$$

With this, we would then use the deterministic controller $\Phi^*_W(\mathrm{d}u|z) = \delta_{\varphi^*_W(z)}(\mathrm{d}u)$, where $\delta_u$ denotes the Dirac measure concentrated at $u \in \mathsf{U}$. Thus,

$$\begin{aligned} V(R,\mathsf{Z}) &= \inf_{\substack{W\in\mathcal{M}(\mathsf{Z}|\mathsf{X})\\ I(X;Z)\le R}} \mathbb{E}[c(X,\varphi^*_W(Z))] \\ &= \inf_{\substack{W\in\mathcal{M}(\mathsf{Z}|\mathsf{X})\\ I(X;Z)\le R}} \mathbb{E}\left[\min_{u\in\mathsf{U}} \mathbb{E}[c(X,u)|Z]\right]. \end{aligned}$$

We also need to introduce some notions from rate-distortion theory [16]. For a given $\mu \in \mathcal{P}(\mathsf{X})$ and a given $R \ge 0$, consider the set

$$\mathcal{I}_\mu(R) \triangleq \Big\{K \in \mathcal{M}(\mathsf{U}|\mathsf{X}) : I(\mu,K) \le R\Big\}.$$

The set $\mathcal{I}_\mu(R)$ is nonempty for every $R \ge 0$. To see this, note that any kernel $K_\diamond \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ for which the function $x \mapsto K_\diamond(B|x)$ is constant ($\mu$-a.e. for any $B \in \mathcal{B}(\mathsf{U})$) satisfies $I(\mu,K_\diamond) = 0$.

The *Shannon distortion-rate function* (DRF) of $\mu$ is defined as

$$D_\mu(R) \triangleq \inf_{K\in\mathcal{I}_\mu(R)} \langle \mu\otimes K, c\rangle. \tag{4.3}$$

We use the more cumbersome notation $D_\mu(R;c)$ when we need to specify the dependence of the DRF on the underlying cost function $c$. Starting with the easily proved variational expression

$$I(\mu,K) = \inf_{\nu\in\mathcal{P}(\mathsf{U})} D(\mu\otimes K \| \mu\otimes\nu)$$

(where the infimum is achieved uniquely by $\nu = \mu K$), we can introduce the

Lagrangian relaxation

$$\mathcal{L}_\mu(K,\nu,s) \triangleq sD(\mu\otimes K\|\mu\otimes\nu) + \langle \mu\otimes K, c\rangle$$

for $s \ge 0$ and $\nu \in \mathcal{P}(\mathsf{U})$, and establish the following key results [16, 23]:

**Proposition 4.0.1** *The DRF $D_\mu(R)$ is convex and nondecreasing in $R$. If $D_\mu(R) < \infty$, then a Markov kernel $K^* \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ attains the infimum in (4.3) if and only if $I(\mu, K^*) = R$ and the Radon–Nikodym derivative of $\mu\otimes K^*$ w.r.t. $\mu\otimes\mu K^*$ takes the form*

$$\frac{\mathrm{d}(\mu\otimes K^*)}{\mathrm{d}(\mu\otimes\mu K^*)}(x,y) = \alpha(x)e^{-\frac{1}{s}c(x,u)}, \tag{4.4}$$

*where $\alpha : \mathsf{X} \to \mathbb{R}^+$ and $s \ge 0$ are such that*

$$\int_{\mathsf{X}} \alpha(x)e^{-\frac{1}{s}c(x,u)}\mu(\mathrm{d}x) \le 1, \qquad \forall u \in \mathsf{U} \tag{4.5}$$

*and $-s$ is the slope of a line tangent to the graph of $D_\mu(R)$ at $R$:*

$$D_\mu(R') + sR' \ge D_\mu(R) + sR, \qquad \forall R' \ge 0. \tag{4.6}$$

**Proposition 4.0.2** *The DRF $D_\mu(R)$ can be expressed as*

$$D_\mu(R) = \sup_{s\ge 0}\inf_{\nu\in\mathcal{P}(\mathsf{U})} s\left[\left\langle \mu, \log\frac{1}{\int_{\mathsf{U}} e^{-\frac{1}{s}c(x,u)}\nu(\mathrm{d}u)}\right\rangle - R\right].$$

We are now in a position to state and prove the main result of this section:

**Theorem 4.0.3 (Stratonovich)** *For any $R \ge 0$ such that $D_\mu(R) < \infty$, we have $V(R) = D_\mu(R)$, and the infimum over $\mathsf{Z}$ in (4.2) is attained at $\mathsf{Z}^* = \mathcal{P}(\mathsf{X})$.*

**Proof (Sketch)** One direction, $V(R) \ge D_\mu(R)$, is relatively straightforward. Fix $\mathsf{Z}$, and suppose that $W^* \in \mathcal{M}(\mathsf{Z}|\mathsf{X})$ and $\Phi^* \in \mathcal{M}(\mathsf{U}|\mathsf{Z})$ attain the optimal value $V(R,\mathsf{Z})$. Consider the Markov kernel $K \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ obtained by composing $\Phi^*$ and $W^*$: $K = \Phi^* \circ W^*$. The joint action of $W^*$ and $\Phi^*$ can be described by a Markov chain $X \to Z^* \to U^*$, where $\mathrm{Law}(X) = \mu$, $\mathrm{Law}(Z^*|X = x) = W^*(\cdot|x)$, and $\mathrm{Law}(U^*|X = x, Z^* = z) = \mathrm{Law}(U|Z^* = z) = \Phi^*(\cdot|z)$. Moreover, $\mathrm{Law}(X,U^*) = \mu\otimes K$. Since

$I(X;Z^*) \le R$, we have $K \in \mathcal{I}_\mu(R)$ by the data processing inequality (2.2). Consequently, $D_\mu(R) \le \langle \mu \otimes K, c\rangle = V(R,\mathsf{Z})$. Taking the infimum over $\mathsf{Z}$, we get $D_\mu(R) \le V(R)$.

To prove the other direction, let $\mathsf{Z} = \mathcal{P}(\mathsf{X})$[1] and consider the optimal kernel $K^* \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ that achieves the infimum in (4.3). Using (4.4) and the Bayes' rule, we can compute the posterior distribution (belief state)

$$\check{K}^*(\mathrm{d}x|u) = \frac{e^{-\frac{1}{s}c(x,u)}\mu(\mathrm{d}x)}{\int_{\mathsf{X}} e^{-\frac{1}{s}c(x,u)}\mu(\mathrm{d}x)}.$$

Using the minimal sufficiency property of the belief state [24,25] and the fact that $K^*$ attains the DRF, it can be shown that the kernel $W \in \mathcal{M}(\mathsf{Z}|\mathsf{X})$ given by the composition of $K^*$ and the deterministic mapping $u \mapsto \check{K}^*(\cdot|u)$ is feasible for the problem (4.1) with $\mathsf{Z} = \mathcal{P}(\mathsf{X})$. Moreover, if we choose the controller $\Phi \in \mathcal{M}(\mathsf{U}|\mathsf{Z})$ in such a way that $\Phi(\mathrm{d}u|z)$ is supported on the set $\{u \in \mathsf{U} : \check{K}^*(\cdot|u) = z\}$, then the resulting cost will not exceed $D_\mu(R)$. Again, taking the infimum over $\mathsf{Z}$, we get $V(R) \le D_\mu(R)$.

[1]Note that because $\mathsf{X}$ is standard Borel, the space $\mathsf{Z} = \mathcal{P}(\mathsf{X})$ is a complete separable metric space w.r.t. any metric that compatible with weak convergence of probability measures, and so $\mathsf{Z}$ is standard Borel as well.

# CHAPTER 5

# AVERAGE-COST OPTIMAL CONTROL WITH RATIONAL INATTENTION

We now turn to the analysis of the infinite-horizon control problem (3.5) with an information constraint. In multi-stage control problems, such as this one, the control law has a *dual effect* [26]: it affects both the cost at the current stage and the uncertainty about the state at future stages. The presence of the mutual information constraint (3.5b) enhances this dual effect, since it prevents the controller from ever learning "too much" about the state. This, in turn, limits the controller's future ability to keep the average cost low. These considerations suggest that, in order to bring rate-distortion theory to bear on the problem (3.5), we cannot use the one-stage cost $c$ as the distortion function. Instead, we must modify it to account for the effect of the control action on future costs. As we will see, this modification implies a certain stochastic relaxation of the Average Cost Optimality Equation (ACOE) that characterizes optimal performance achievable by any MRS control law in the absence of information constraints.

## 5.1 Reduction to single-stage optimization

To construct this modification in a principled manner, we first reduce the dynamic optimization problem (3.5) to a suitable static (single-stage) problem. Once this has been carried out, we will be able to take advantage of the results of Section 4. The reduction is based on the so-called *convex-analytic* approach to controlled Markov processes [27] (see also [28, 29]), which we briefly summarize here. Given a Markov control problem with initial state distribution $\mu$ and transition kernel $Q$, the long-run expected average cost of an MRS control law $\Phi \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ is given by

$$J_\mu(\Phi) = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T} c(X_t, U_t)\right]. \tag{5.1}$$

We wish to find an MRS control law $\Phi^*$ that would minimize $J_\mu(\Phi)$ simultaneously for all $\mu$. Any MRS control law $\Phi$ induces a transition kernel $Q_\Phi$ on the state space $\mathsf{X}$:

$$Q_\Phi(A|x) \triangleq \int_{\mathsf{U}} Q(A|x,u)\Phi(\mathrm{d}u|x), \quad \forall A \in \mathcal{B}(\mathsf{X}).$$

We say that $\Phi$ is *stable* if:

1. There exists a probability measure $\pi_\Phi \in \mathcal{P}(\mathsf{X})$ which is invariant to $Q_\Phi$, i.e., $\pi_\Phi = \pi_\Phi Q_\Phi$.

2. The average cost $J_{\pi_\Phi}(\Phi)$ is finite, and moreover

$$J_{\pi_\Phi}(\Phi) = \langle \Gamma_\Phi, c\rangle = \int_{\mathsf{X}\times\mathsf{U}} c(x,u)\Gamma_\Phi(\mathrm{d}x, \mathrm{d}u),$$

   where $\Gamma_\Phi \triangleq \pi_\Phi \otimes \Phi$.

Let $\mathcal{K} \subset \mathcal{M}(\mathsf{U}|\mathsf{X})$ denote the space of all such stable laws.

**Theorem 5.1.1** *Suppose that the following assumptions are satisfied:*

- *(**A.1**) The one-stage cost function $c$ is nonnegative, lower semicontinuous, and coercive, i.e., there exist two sequences of compact sets $\mathsf{X}_n \uparrow \mathsf{X}$ and $\mathsf{U}_n \uparrow \mathsf{U}$ such that*

$$\lim_{n\to\infty} \inf_{x\in\mathsf{X}_n^c, u\in\mathsf{U}_n^c} c(x,u) = +\infty.$$

- *(**A.2**) The transition kernel $Q$ is continuous, i.e., $Qf \in C_b(\mathsf{X}\times\mathsf{U})$ for any $f \in C_b(\mathsf{X})$.*

*Let $J^* \triangleq \inf_\Phi \inf_\mu J_\mu(\Phi)$. Then there exists a control law $\Phi^* \in \mathcal{K}$, such that*

$$J_{\pi_{\Phi^*}}(\Phi^*) = J^* = \inf_{\Phi\in\mathcal{K}} \langle \Gamma_\Phi, c\rangle. \tag{5.2}$$

## 5.2 Bellman error minimization via marginal decomposition

For our purposes, it is convenient to decompose the infimum over $\Phi$ in (5.2) by first fixing the marginal $\pi \in \mathcal{P}(\mathsf{X})$. Consider the set of all stable control laws that leave $\pi$ invariant,

$$\mathcal{K}_\pi \triangleq \Big\{ \Phi \in \mathcal{K} : \pi = \pi_\Phi \Big\}.$$

This set may very well be empty for some $\pi$. Regardless, assuming that the conditions of Theorem 5.1.1 are satisfied, we can write

$$J^* = \inf_{\Phi \in \mathcal{K}} \langle \Gamma_\Phi, c \rangle = \inf_{\pi \in \mathcal{P}(\mathsf{X})} \inf_{\Phi \in \mathcal{K}_\pi} \langle \pi \otimes \Phi, c \rangle.$$

We are now in a position to introduce the information constraint. Let $\mathcal{K}_\pi(R) \triangleq \mathcal{K}_\pi \cap \mathcal{I}_\pi(R)$. Then we define the optimal steady-state value of the information-constrained average-cost control problem as

$$J^*(R) \triangleq \inf_{\pi \in \mathcal{P}(\mathsf{X})} \inf_{\Phi \in \mathcal{K}_\pi(R)} \langle \pi \otimes \Phi, c \rangle. \tag{5.3}$$

As a first step to understanding solutions to (5.3), we consider each candidate invariant distribution $\pi \in \mathcal{P}(\mathsf{X})$ separately:

**Proposition 5.2.1** *For any* $\pi \in \mathcal{P}(\mathsf{X})$*, let*

$$J^*_\pi(R) \triangleq \inf_{\Phi \in \mathcal{K}_\pi(R)} \langle \pi \otimes \Phi, c \rangle.$$

*Then*

$$J^*_\pi(R) = \inf_{\Phi \in \mathcal{I}_\pi(R)} \sup_{h \in C_b(\mathsf{X})} \langle \pi \otimes \Phi, c + Qh - h \rangle. \tag{5.4}$$

**Remark** The function $h \in C_b(\mathsf{X})$ plays the role of a Lagrange multiplier associated with the constraint $\Phi \in \mathcal{K}_\pi$, which is what can be expected from the theory of average-cost optimal control [28, Ch. 9].

On setting $\eta = \langle \pi \otimes \Phi, c \rangle$, the function $c + Qh - h - \eta$ is the *Bellman error* associated with $h$ that is central to approximate dynamic programming.

**Remark** Both in (5.4) and elsewhere, we can extend the supremum over $h$

to all $h \in L^1(\pi)$ without affecting the value of $J^*_\pi(R)$.

**Proof** Define the function

$$\iota_\pi(\Phi) \triangleq \begin{cases} 0, & \Phi \in K_\pi \\ +\infty, & \Phi \notin K_\pi \end{cases}.$$

Then we can write

$$J^*_\pi(R) = \inf_{\Phi \in \mathcal{I}_\pi(R)} \left[ \big\langle \pi \otimes \Phi, c \big\rangle + \iota_\pi(\Phi) \right]. \tag{5.5}$$

Moreover,

$$\iota_\pi(\Phi) = \sup_{h \in C_b(\mathsf{X})} \left[ \big\langle \pi Q_\Phi, h \big\rangle - \big\langle \pi, h \big\rangle \right]. \tag{5.6}$$

Indeed, if $\Phi \in K_\pi$, then the right-hand side of (5.6) is zero. On the other hand, suppose that $\Phi \notin K_\pi$. Since $\mathsf{X}$ is standard Borel, any two probability measures $\mu, \nu \in \mathcal{P}(\mathsf{X})$ are equal if and only if $\langle \mu, h \rangle = \langle \nu, h \rangle$ for all $h \in C_b(\mathsf{X})$. Consequently, there exists some $h_0 \in C_b(\mathsf{X})$ such that $\langle \pi, h_0 \rangle \neq \langle \pi Q_\Phi, h_0 \rangle$. There is no loss of generality if we assume that $\langle \pi Q_\Phi, h_0 \rangle - \langle \pi, h_0 \rangle > 0$. Then by considering functions $h^n_0 = n h_0$ for all $n = 1, 2, \ldots$ and taking the limit as $n \to \infty$, we can make the right-hand side of (5.6) grow without bound. Thus, we have proved (5.6). Substituting it into (5.5), we get (5.4).

To analyze the optimization problem (5.3), let us fix some $\pi$ and consider the *dual value*

$$J_{*,\pi}(R) \triangleq \sup_{h \in C_b(\mathsf{X})} \inf_{\Phi \in \mathcal{I}_\pi(R)} \langle \pi \otimes \Phi, c + Qh - h \rangle. \tag{5.7}$$

Clearly, $J^*_\pi(R) \geq J_{*,\pi}(R)$ for all $\pi$ and $R$. Moreover:

**Proposition 5.2.2** *Suppose that assumption* (**A.1**) *above is satisfied, and that $J^*_\pi(R) < \infty$. Then the primal value $J^*_\pi(R)$ and the dual value $J_{*,\pi}(R)$ are equal.*

**Proof** Let $\mathcal{P}^0_{\pi,c}(R) \subset \mathcal{P}(\mathsf{X} \times \mathsf{U})$ be the closure, in the weak topology, of the set of all probability measures $\Gamma \in \mathcal{P}(\mathsf{X} \times \mathsf{U})$, such that $\Gamma(A \times \mathsf{U}) = \pi(A)$,

$I(\Gamma) \leq R$, and $\langle \Gamma, c \rangle < \infty$. Since $J^*_\pi(R) < \infty$ by hypothesis, we can write

$$J^*_\pi(R) = \inf_{\Gamma \in \mathcal{P}^0_{\pi,c}(R)} \sup_{h \in C_b(\mathsf{X})} \langle \Gamma, c + Qh - h \rangle. \tag{5.8}$$

Because $c$ is coercive and nonnegative, the set $\{\Gamma \in \mathcal{P}(\mathsf{X} \times \mathsf{U}) : \langle \Gamma, c \rangle < \infty\}$ is tight [30, Proposition 1.4.15], so its closure is weakly sequentially compact by Prohorov's theorem. Moreover, because the function $\Gamma \mapsto I(\Gamma)$ is weakly lower semicontinuous [3], the set $\{\Gamma : I(\Gamma) \leq R\}$ is closed. Therefore, the set $\mathcal{P}^0_{\pi,c}(R)$ is closed and tight, hence weakly sequentially compact. Moreover, the sets $\mathcal{P}^0_{\pi,c}(R)$ and $C_b(\mathsf{X})$ are both convex, and the objective function on the right-hand side of (5.8) is affine in $\Gamma$ and linear in $h$. Therefore, by Sion's minimax theorem [31] we may interchange the supremum and the infimum to conclude that $J^*_\pi(R) = J_{*,\pi}(R)$.

We are now in a position to relate the optimal value $J^*_\pi(R)$ to a suitable rate-distortion problem. For the sake of conciseness, we remind ourselves of the notation in (4.3) and denote the DRF of $\pi$ w.r.t. the distortion function $c + Qh_\pi$ with

$$D_\pi(R; c + Qh_\pi) \triangleq \inf_{\Phi \in \mathcal{I}_\pi(R)} \langle \pi \otimes \Phi, c + Qh_\pi \rangle. \tag{5.9}$$

The key to our set-up is the existence of a $h^*_\pi$ which is the sufficient condition to assure that $J^*(R)$ has a solution corresponding to $\pi$.

**Theorem 5.2.3** *Consider a probability measure $\pi \in \mathcal{P}(\mathsf{X})$ such that $J^*_\pi(R) < \infty$, and the supremum over $h_\pi \in C_b(\mathsf{X})$ in* (5.7) *is attained by some $h^*_\pi$. Then there exists an MRS control law $\Phi^* \in \mathcal{M}(\mathsf{U}|\mathsf{X})$, such that $I(\pi, \Phi^*) = R$, and we have*

$$J^*_\pi(R) + \langle \pi, h^*_\pi \rangle = \langle \pi \otimes \Phi^*, c + Qh^*_\pi \rangle = D_\pi(R; c + Qh^*_\pi). \tag{5.10}$$

**Proof** Using Proposition 5.2.2 and the definition (5.7) of the dual value $J_{*,\pi}(R)$, we can express $J^*_\pi(R)$ as a pointwise supremum of a family of DRF's:

$$J^*_\pi(R) = \sup_{h_\pi \in C_b(\mathsf{X})} \left[ D_\pi(R; c + Qh_\pi) - \langle \pi, h_\pi \rangle \right]. \tag{5.11}$$

Since $J^*_\pi(R) < \infty$, we can apply Proposition 4.0.1 separately for each $h_\pi \in$

$C_b(\mathsf{X})$. In particular, we can take $h_\pi^* \in C_b(\mathsf{X})$ that achieves the supremum in (5.11) (the existence is assumed).

**Remark:** Upon substituting $J_\pi^*(R) < \infty$ in Equation (5.10) with $\lambda_\pi$, Equation (5.12) as below is derived which can be interpreted as *Information-Constrained Bellman Equation* (IC-BE) in contrast to standard *Bellman Equation*(BE) for the average cost [18]. One can notice the intuitive similarity between the two:

$$\begin{aligned} &\text{BE:} && h(x) + \lambda = \inf_{u \in \mathsf{U}} [c(x,u) + Qh(x,u)] \\ &\text{IC-BE:} && \langle \pi, h_\pi \rangle + \lambda_\pi = \inf_{\Phi \in \mathcal{I}_\pi(R)} \langle \pi \otimes \Phi, c + Qh_\pi \rangle. \end{aligned} \tag{5.12}$$

Note that while in standard BE the controller has access to the exact value of the state, in IC-BE it just has access to the ergodic distribution of the state $\pi$ which is the prior based on which the decision is made. Moreover, while the control belongs to $\mathsf{U}$ in BE, it has to be constrained to $\mathcal{I}_\pi(R)$ in IC-BE. In part 5.2.1 we show how the standard BE can be recovered in the limit $R \to \infty$.

Utilizing the results in Propositions 4.0.1 and 4.0.2 to substitute for $D_\pi(R; c + Qh_\pi)$, we can have the form of optimal randomized Markov policy under the condition of Theorem 5.2.3.

**Theorem 5.2.4** *Under the condition of Theorem 5.2.3, the Radon–Nikodym derivative of $\pi \otimes \Phi^*$ w.r.t. $\pi \otimes \pi\Phi^*$ takes the form*

$$\frac{\mathrm{d}(\pi \otimes \Phi^*)}{\mathrm{d}(\pi \otimes \pi\Phi^*)}(x,u) = \frac{e^{-\frac{1}{s}d(x,u)}}{\int_{\mathsf{U}} e^{-\frac{1}{s}d(x,u)} \pi\Phi^*(\mathrm{d}u)}, \tag{5.13}$$

*where $d(x,u) \triangleq c(x,u) + Qh_\pi^*(x,u)$, and $s \geq 0$ satisfies*

$$D_\pi(R'; c + Qh_\pi^*) + sR' \geq D_\pi(R; c + Qh_\pi^*) + sR \tag{5.14}$$

*for all $R'$. Moreover, the optimal value $J_\pi^*(R)$ admits the following variational* *representation:*

$$J_\pi^*(R) = \sup_{s\geq 0} \sup_{h\in C_b(\mathsf{X})} \inf_{\nu\in\mathcal{P}(\mathsf{U})} \Bigg\{ -\langle \pi, h\rangle + s\left[\left\langle \pi, \log \frac{1}{\int_{\mathsf{U}} e^{-\frac{1}{s}[c(x,u)+Qh(x,u)]}\nu(\mathrm{d}u)}\right\rangle - R\right]\Bigg\}. \tag{5.15}$$

**Proof** For the first part, using (4.4) with

$$\alpha(x) = \frac{1}{\int_{\mathsf{U}} e^{-\frac{1}{s}d(x,u)}\pi\Phi^*(\mathrm{d}u)}$$

results in (5.13). In the same way, (5.14) follows from (4.6) in Proposition 4.0.1. The form of optimal value can be obtained immediately from (5.11) and Proposition 4.0.2.

The central role of having the existence and form of Lagrange multiplier $h_\pi^*$ becomes evident through Theorems 5.2.3 and 5.2.4. While the first one guarantees the existence of optimal stationary policy dependent on the existence of $h_\pi^*$ in dual problem corresponding to a candidate $\pi$, the second one characterize the form of the optimal policy based on the form of $h_\pi^*$. Moreover, Theorem 5.2.3 also provides a necessary condition in the form of Equation (5.12) (or equally 5.10), for $h_\pi^*$ to satisfy corresponding to a candidate ergodic state distribution $\pi$.

One may be tempted to consider (5.12) as a fixed point equation to be solved for $h_\pi$, or having a converse to Theorem 5.2.3, which is a common approach in similar perfect information situations.

However, the most we can get from Equation (5.12), is to have optimality among the family of policies which keep the expected value of a function of the form $h$ constant over time. In general, this does not even mean that the candidate distribution is invariant with respect to the controlled kernel. Note that in order to prove that a policy $\Phi$ induces a kernel $Q_\Phi$ for which $\pi$ is invariant, a necessary and sufficient condition is to show that this policy keeps the expected value of every $h \in C_b(\mathsf{X})$ constant over time (or the supremum, as we see in primal function).

However, one can show that if (5.12) holds for some $(\pi, h_\pi, \lambda_\pi)$ and the resulting control policy $\Phi$ can keep the candidate distribution $\pi$ invariant with respect to the controlled kernel, then one may claim that the policy actually attains the infimum for that candidate distribution.

**Theorem 5.2.5** *Suppose there exist $h_\pi \in L^1(\pi)$, $\lambda_\pi < \infty$ and a stochastic kernel $\Phi^* \in \mathcal{K}_\pi(R)$ such that*

$$\langle \pi, h_\pi \rangle + \lambda_\pi = D_\pi(R; c + Qh_\pi) = \langle \pi \otimes \Phi^*, c + Qh_\pi \rangle.$$

*Then $\Phi^* \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ achieves the infimum in (3.5) which is equal to $\lambda_\pi$.*

**Proof** For the proof of this theorem, refer to section A.5

To complete the computation of the optimal steady-state value $J^*(R)$ defined in (5.3), we need to consider all candidate invariant distributions $\pi \in \mathcal{P}(\mathsf{X})$ for which $\mathcal{K}_\pi(R)$ is nonempty, and then choose among them any $\pi$ that attains the smallest value of $J^*_\pi(R)$ (assuming this value is finite).

On the other hand, if $J^*_\pi(R) < \infty$ for *some* $\pi$, then Theorem 5.2.3 ensures that there exists a *suboptimal* control law satisfying the information constraint.

### 5.2.1 Some implications (recovery of perfect information case)

Using Theorem 5.2.4, we see that $J^*_\pi(R)$ is equal to the value of the following optimization problem:

$$\begin{aligned}
&\text{maximize } \lambda \\
&\text{s.t. } s\left\langle \pi, \log \frac{1}{\int_{\mathsf{U}} e^{-\frac{1}{s}[c(x,u)+Qh(x,u)]}\nu(\mathrm{d}u)} - \frac{h}{s} \right\rangle \geq \lambda + sR, \\
&\forall \nu \in \mathcal{P}(\mathsf{U}) \quad \lambda \geq 0,\ s \geq 0,\ h \in C_b(\mathsf{X}).
\end{aligned}$$

Let us examine what happens as we relax the information constraint, i.e., let $R \to \infty$. From the fact that the DRF is convex and nondecreasing in $R$, and from the characterization (5.14) of $-s$ as the slope of a tangent to the graph of the DRF at $R$, this is equivalent to letting $s$ approach 0 (with the convention that $sR \to 0$ even as $R \to \infty$). Let us recall *Laplace's principle* [32]: for any $\nu \in \mathcal{P}(\mathsf{U})$ and any measurable function $F : \mathsf{U} \to \mathbb{R}$ such that $e^{-F} \in L^1(\nu)$, we have

$$-\lim_{s\downarrow 0} s\log\int_{\mathsf{U}} e^{-\frac{1}{s}F(u)}\nu(\mathrm{d}u) = \operatorname*{ess\,inf}_{u\in\mathsf{U}} F(u),$$

where the essential infimum is defined w.r.t. $\nu$. In view of this, the limit of $J^*_\pi(R)$ as $R\to\infty$ is the value of

$$\begin{aligned}&\text{maximize } \lambda\\ &\qquad \text{s.t. } \left\langle \pi, \inf_{u\in\mathsf{U}}\left[c(x,u)+Qh(x,u)\right]-h\right\rangle \geq \lambda\\ &\qquad\qquad \lambda\geq 0,\ h\in C_b(\mathsf{X}).\end{aligned}$$

Performing now the minimization over $\pi\in\mathcal{P}(\mathsf{X})$ as well, we see that the limit of $J^*(R)$ as $R\to\infty$ is given by the value of the following problem:

$$\begin{aligned}&\text{maximize } \lambda\\ &\qquad \text{s.t. } \inf_{u\in\mathsf{U}}\left[c(x,u)+Qh(x,u)\right]-h \geq \lambda\\ &\qquad\qquad \lambda\geq 0,\ h\in C_b(\mathsf{X}).\end{aligned}$$

In particular, under Assumptions (**A.1**) and (**A.2**) of Theorem 5.1.1, there exists a function $h^\infty$ and constant $\lambda^\infty\geq 0$ that solve the Average Cost Optimality Equation (ACOE)

$$h^\infty(x)+\lambda^\infty = \inf_{u\in\mathsf{U}}\left[c(x,u)+Qh^\infty(x,u)\right],$$

and $J^*(R)\to\lambda^\infty$ as $R\to\infty$.

# CHAPTER 6

# INFORMATION-CONSTRAINED LQG PROBLEM

We now formulate the scalar LQG problem in the rational inattention regime. Consider the following linear time-invariant stochastic system:

$$X_{t+1} = aX_t + b\,U_t + W_t, \qquad t \geq 1, \tag{6.1}$$

where $a, b \neq 0$ are the system coefficients, $\{X_t\}_{t=1}^{\infty}$ is a real-valued state process, $\{U_t\}_{t=1}^{\infty}$ is a real-valued control process, and $\{W_t\}_{t=1}^{\infty}$ is a sequence of independent and identically distributed (i.i.d.) Gaussian random variables with mean 0 and variance $\sigma^2$. The initial state $X_1$ has some given distribution $\mu$. Here, $\mathsf{X} = \mathsf{U} = \mathbb{R}$, and the controlled transition kernel $Q \in \mathcal{M}(\mathsf{X}|\mathsf{X} \times \mathsf{U})$ corresponding to (6.1) is

$$Q(\mathrm{d}y|x, u) = \gamma(y; ax + bu, \sigma^2)\,\mathrm{d}y,$$

where

$$\gamma(y; \mathsf{m}, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y - \mathsf{m})^2}{2\sigma^2}\right)$$

is the probability density of a Gaussian distribution with mean $\mathsf{m}$ and variance $\sigma^2$, and $\mathrm{d}y$ is the Lebesgue measure.

We focus on the quadratic performance objective

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T} pX_t^2 + qU_t^2\right]$$

with $p, q > 0$. Following the formalism of Section 5, we seek a pair consisting of an invariant distribution $\pi \in \mathcal{P}(\mathsf{X})$ and an MRS control law $\Phi \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ to attain the steady-state value with $c(x, u) = px^2 + qu^2$ under the information constraint $I(\pi, \Phi) \leq R$.

## 6.1 Main result and some implications

We now state the main result of this thesis, which gives an upper bound on the information-constrained average cost in the LQG problem of Section 6:

**Theorem 6.1.1** *Suppose that the system* (6.1) *is open-loop stable, i.e.,* $a^2 < 1$. *Fix an information constraint* $R > 0$. *Let* $m_1 = m_1(R)$ *be the unique positive root of the* information-constrained discrete algebraic Riccati equation (IC-DARE)

$$p + m(a^2 - 1) + \frac{(mab)^2}{q + mb^2}(e^{-2R} - 1) = 0, \tag{6.2}$$

*and let* $m_2$ *be the unique positive root of the standard DARE*

$$p + m(a^2 - 1) - \frac{(mab)^2}{q + mb^2} = 0. \tag{6.3}$$

*Define the control gains* $k_1 = k_1(R)$ *and* $k_2$ *by*

$$k_i = -\frac{m_i ab}{q + m_i b^2} \tag{6.4}$$

*and steady-state variances* $\sigma_1^2 = \sigma_1^2(R)$ *and* $\sigma_2^2 = \sigma_2^2(R)$ *by*

$$\sigma_i^2 = \frac{\sigma^2}{1 - \left[e^{-2R}a^2 + (1 - e^{-2R})\left(a + bk_i\right)^2\right]}. \tag{6.5}$$

*Then*

$$J^*(R) \le \min\Big(m_1\sigma^2, m_2\sigma^2 + (q + m_2 b^2)k_2^2\sigma_2^2 e^{-2R}\Big). \tag{6.6}$$

*Also, let* $\Phi_1$ *and* $\Phi_2$ *be two MRS control laws with Gaussian conditional densities*

$$\begin{aligned} \varphi_i(u|x) &= \frac{\mathrm{d}\Phi_i(u|x)}{\mathrm{d}u} \\ &= \gamma\left(u; (1 - e^{-2R})k_i x, (1 - e^{-2R})e^{-2R}k_i\sigma_i^2\right), \end{aligned} \tag{6.7}$$

*and let* $\pi_i = N(0, \sigma_i^2)$ *for* $i = 1, 2$. *Then the first term on the right-hand side* *of* (6.6) *is achieved by* $\Phi_1$*, the second term is achieved by* $\Phi_2$*, and*

$$\Phi_i \in \mathcal{K}_{\pi_i}(R), \qquad i = 1, 2.$$

*Moreover, in each case the information constraint is met with equality:* $I(\pi_i, \Phi_i) = R$*,* $i = 1, 2$*.*

Before we proceed with the proof of Theorem 6.1.1, we pause to examine a few consequences. First of all, the controllers $\Phi_1$ and $\Phi_2$ coincide and attain global optimality in both the no-information ($R = 0$) and the perfect-information ($R = +\infty$) cases. Indeed, when $R = 0$, the quadratic IC-DARE (6.2) reduces to the linear Lyapunov equation [17]

$$p + m(a^2 - 1) = 0,$$

so the first term on the right-hand side of (6.6) is

$$m_1(0)\sigma^2 = \frac{p\sigma^2}{1 - a^2}.$$

On the other hand, using Eqs. (6.3) and (6.4), we can show that the second term is equal to the first term, so from (6.6)

$$J^*(0) \le \frac{p\sigma^2}{1 - a^2}. \tag{6.8}$$

Since this is the minimal average cost in the open-loop case, we have equality in (6.8). Also, the controllers $\Phi_1$ and $\Phi_2$ are both realized by the deterministic open-loop law $U_t \equiv 0$ for all $t$, as expected. Finally, the steady-state variance is

$$\sigma_1^2(0) = \sigma_2^2(0) = \frac{\sigma^2}{1 - a^2},$$

and $\pi_1 = \pi_2 = N(0, \sigma^2/(1 - a^2))$, which is the unique invariant distribution of the system (6.1) with $U_t \equiv 0$ for all $t$ (recall the stability assumption $a^2 < 1$).

On the other hand, in the limit $R \to \infty$ the IC-DARE (6.2) reduces to the usual DARE (6.3). Hence, $m_1(\infty) = m_2$, and both terms on the right-hand side of (6.6) are equal to $m_2\sigma^2$. This gives

$$J^*(\infty) \le m_2\sigma^2. \tag{6.9}$$

Since this is the minimal average cost attainable in the scalar LQG control problem with perfect information, we have equality in (6.9), as expected. The controllers $\Phi_1$ and $\Phi_2$ are again both deterministic and have the usual linear structure $U_t = k_2 X_t$ for all $t$. The steady-state variance is

$$\sigma_1^2(\infty) = \sigma_2^2(\infty) = \frac{\sigma^2}{1-(a+bk_2)^2},$$

which is the steady-state variance induced by the optimal controller in the standard LQG problem.

In the presence of a nontrivial information constraint ($0 < R < \infty$), the two control laws $\Phi_1$ and $\Phi_2$ are no longer the same. However, they are both *stochastic* and have the form

$$U_t = k_i \left[ (1-e^{-2R})X_t + e^{-R}\sqrt{1-e^{-2R}}V_t^{(i)} \right], \tag{6.10}$$

where $\{V_t^{(i)}\}_{t=1}^\infty$ is a sequence of i.i.d. $N(0,\sigma_i^2)$ random variables independent of $\{W_t\}_{t=1}^\infty$ and $X_1$. The corresponding closed-loop system is

$$X_{t+1} = \left[a + \left(1-e^{-2R}\right)bk_i\right] X_t + Z_t^{(i)}, \tag{6.11}$$

where $\{Z_t^{(i)}\}_{t=1}^\infty$ is a sequence of i.i.d. Gaussian random variables with mean 0 and variance

$$\bar{\sigma}_i^2 = e^{-2R}(1-e^{-2R})\left(bk_i\right)^2 \sigma_i^2 + \sigma^2.$$

Theorem 6.1.1 implies that, for each $i = 1, 2$, this system is stable and has the invariant distribution $\pi_i = N(0, \sigma_i^2)$. Moreover, this invariant distribution is unique, and the closed-loop transition kernels $K_{\Phi_i}$, $i = 1, 2$, are ergodic. We also note that the two controllers in (6.10) can be realized as a cascade consisting of an additive white Gaussian noise (AWGN) channel and a linear gain:

$$\begin{aligned} U_t &= k_i \widehat{X}_t^{(i)} \\ \widehat{X}_t^{(i)} &= (1 - e^{-2R}) X_t + e^{-R}\sqrt{1-e^{-2R}} V_t^{(i)}. \end{aligned}$$

We can view the stochastic mapping from $X_t$ to $\widehat{X}_t^{(i)}$ as a noisy *sensor* or *state observation channel* that adds just enough noise to the state to satisfy the information constraint in the steady state, while introducing a minimum amount of distortion. The difference between the two control laws $\Phi_1$ and $\Phi_2$ is due to the fact that, for $0 < R < \infty$, $k_1(R) \neq k_2$ and $\sigma_1^2(R) \neq \sigma_2^2(R)$. Note also that the deterministic (linear gain) part of $\Phi_2$ is exactly the same as in the standard LQG problem with perfect information, with or without noise. In particular, the gain $k_2$ is *independent* of the information constraint $R$. Hence, $\Phi_2$ as a *certainty-equivalent* control law which treats the output $\widehat{X}_t^{(2)}$ of the AWGN channel as the best representation of the state $X_t$ given the information constraint. A control law with this structure was proposed by Sims [1] on heuristic grounds for the information-constrained LQG problem with discounted cost. On the other hand, for $\Phi_1$ both the noise variance $\sigma_1^2$ in the channel $X_t \to \widehat{X}_t^{(1)}$ and the gain $k_1$ depend on the information constraint $R$. Numerical simulations show that $\Phi_1$ attains smaller steady-state cost for all sufficiently small values of $R$ (see Figure 6.1), whereas $\Phi_2$ outperforms $\Phi_1$ when $R$ is large. As shown above, the two controllers are exactly the same (and optimal) in the no-information ($R \to 0$) and perfect-information ($R \to \infty$) regimes.

Finally, we comment on the unstable case ($a^2 > 1$). A simple sufficient condition for the existence of an information-constrained controller that results in a stable closed-loop system is

$$R > \frac{1}{2} \log \frac{a^2 - (a + bk_2)^2}{1 - (a + bk_2)^2}, \tag{6.12}$$

where $k_2$ is the control gain defined in (6.4). Indeed, if $R$ satisfies (6.12), then the steady-state variance $\sigma_2^2$ is well-defined, so the closed-loop system (6.11) with $i = 2$ is stable.

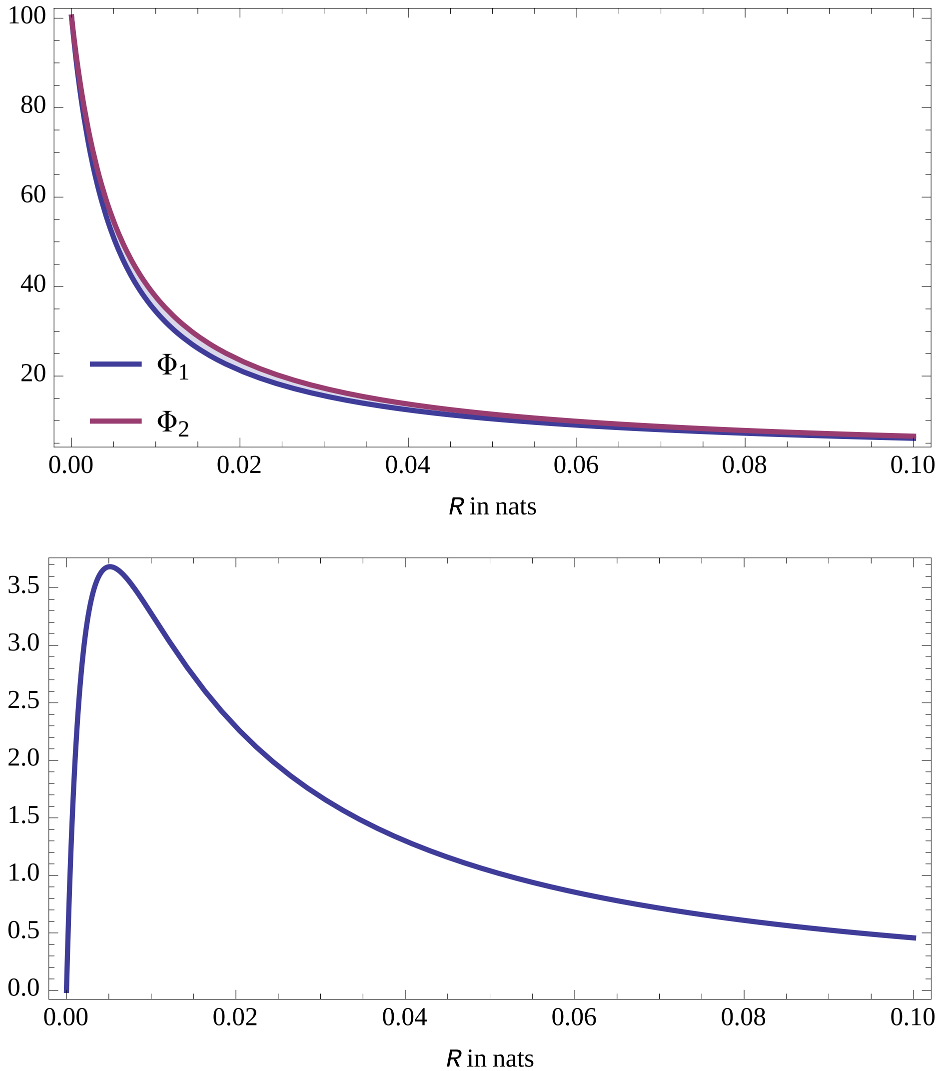


Figure 6.1: Comparison of $\Phi_1$ and $\Phi_2$ at low information rates (top: steady-state values, bottom: difference of steady-state values of $\Phi_2$ and $\Phi_1$). System parameters: $a = 0.995, b = 1, \sigma^2 = 1$, cost parameters: $p = q = 1$.

# CHAPTER 7

# CONCLUSION

The main contributions of this thesis are to pose a general *rational inattention model* for optimal control with information constraints, and to reveal structure for the associated optimal control equations. We also find a tight upper bound on the optimal steady-state value attainable in the scalar LQG control problem subject to a mutual information constraint. We have shown that there are two distinct control policies with different performances in the presence of a nontrivial information constraint, which reduce to optimal deterministic control laws in the two extreme cases of no information and perfect information.

A direction for future work can include the construction of reinforcement learning algorithms that make use of the general structure of the optimal solution in this thesis to obtain approximately optimal policies based on online measurements of the system to be controlled. This is especially important since most empirically verified utility functions in finance have more complicated forms.

# APPENDIX PROOFS

## A.1 Proof of the Principle of Irrelevant Information

Let $M(\mathrm{d}\upsilon|\psi)$ be the conditional distribution of $\Upsilon$ given $\Psi$, let $\Lambda(\mathrm{d}\psi|\theta,\xi)$ be the conditional distribution of $\Psi$ given $(\theta,\xi)$, and disintegrate the joint distribution of $\Theta,\Xi,\Psi,\Upsilon$ as

$$P(\mathrm{d}\theta,\mathrm{d}\xi,\mathrm{d}\psi,\mathrm{d}\upsilon) = P(\mathrm{d}\theta)P(\mathrm{d}\xi|\theta)\Lambda(\mathrm{d}\psi|\theta,\xi)M(\mathrm{d}\upsilon|\psi).$$

If we define $\Lambda'(\mathrm{d}\psi|\theta)$ by

$$\Lambda'(\cdot|\theta) = \int W(\cdot|\theta,\xi)P(\mathrm{d}\xi|\theta)$$

and let the tuple $(\Theta',\Xi',\Psi',\Upsilon')$ have the joint distribution

$$P'(\mathrm{d}\theta,\mathrm{d}\xi,\mathrm{d}\psi,\mathrm{d}\upsilon) = P(\mathrm{d}\theta)P(\mathrm{d}\xi|\theta)\Lambda'(\mathrm{d}\psi|\theta)M(\mathrm{d}\upsilon|\psi),$$

then it is easy to see that it has all of the desired properties.

## A.2 Proof of the Two-Stage Lemma

Note that $Z_1$ only depends on $X_1$, and that only the second-stage expected cost is affected by the choice of $W_2$. We can therefore apply the Principle of Irrelevant Information to $\Theta = X_2$, $\Xi = (X_1, Z_1, U_1)$, $\Psi = Z_2$ and $\Upsilon = U_2$. Because both the expected cost $\mathbb{E}[c(X_t, U_t)]$ and the mutual information $I(X_t; Z_t)$ depend only on the corresponding bivariate marginals, the lemma is proved.

## A.3 Proof of the Three-Stage Lemma

Again, $Z_1$ only depends on $X_1$, and only the second- and the third-stage expected costs are affected by the choice of $W_2$. By the law of iterated expectation, we have

$$\mathbb{E}[c(X_3, U_3)] = \mathbb{E}[\mathbb{E}[c(X_3, U_3)|X_2, U_2]] = \mathbb{E}[h(X_2, U_2)],$$

where $h(X_2, U_2) \triangleq \mathbb{E}[c(X_3, U_3)|X_2, U_2]$. Note that the functional form of $h$ does not depend on the choice of $W_2$, since for any fixed realizations $X_2 = x_2$ and $U_2 = u_2$ we have, by hypothesis,

$$\begin{aligned} h(x_2, u_2) &= \int c(x_3, u_3) P(\mathrm{d}x_3, \mathrm{d}u_3 | x_2, u_2) \\ &= \int c(x_3, u_3) Q(\mathrm{d}x_3 | x_2, u_2) W_3(\mathrm{d}z_3 | x_3) \Phi_3(\mathrm{d}u_3 | \,\mathrm{d}z_3). \end{aligned}$$

Therefore, applying the Principle of Irrelevant Information to $\Theta = X_2$, $\Xi = (X_1, Z_1, U_1)$, $\Psi = Z_2$, and $\Upsilon = U_2$,

$$\begin{aligned} \mathbb{E}[c(X_2', U_2') + c(X_3', U_3')] &= \mathbb{E}[c(X_2', U_2') + h(X_2', U_2')] \\ &= \mathbb{E}[c(X_2, U_2) + h(X_2, U_2)] \\ &= \mathbb{E}[c(X_2, U_2) + c(X_3, U_3)], \end{aligned}$$

where the variables $(X_t', Z_t', U_t')$ are obtained from the original ones by replacing $W_2(\mathrm{d}z_2|x^2, z_1, u_1)$ by $W_2'(\mathrm{d}z_2|x_2)$.

## A.4 The Gaussian distortion-rate function

Given a Borel probability measure $\mu$ on the real line, we denote by $D_\mu(R)$ its distortion-rate function w.r.t. the squared-error distortion $d(x, x') = (x - x')^2$:

$$D_\mu(R) \triangleq \inf_{\substack{K \in \mathcal{M}(\mathbb{R}|\mathbb{R}):\\ I(\mu, K) \le R}} \int_{\mathbb{R}\times\mathbb{R}} (x - x')^2 \mu(\mathrm{d}x) K(\mathrm{d}x'|x) \tag{7.1}$$

(where the mutual information is measured in nats). Let $\mu = N(0, \sigma^2)$. Then we have the following [16]:

- $D_\mu(R) = \sigma^2 e^{-2R}$ .

- The optimal kernel $K^*$ that achieves the infimum in (7.1) has the form

$$K^*(\mathrm{d}x'|x) \\ = \gamma\left(x'; (1-e^{-2R})x, (1-e^{-2R})e^{-2R}\sigma^2\right)\mathrm{d}x' \tag{7.2}$$

  and achieves the information constraint with equality: $I(\mu, K^*) = R$.

- $K^*$ can be realized as a stochastic linear system

$$X' = (1-e^{-2R})X + e^{-R}\sqrt{1-e^{-2R}}V, \tag{7.3}$$

  where $V \sim N(0, \sigma^2)$ is independent of $X$.

## A.5 Proof of Theorem 6.1.1

We want to show that, for $i = 1, 2$, the pair $(h_i, \lambda_i)$ with

$$\begin{aligned} h_1(x) &= m_1 x^2, \quad \lambda_1 = m_1\sigma^2 \\ h_2(x) &= m_2 x^2, \quad \lambda_2 = m_2\sigma^2 + (q + m_2 b^2)k_2^2\sigma_2^2 e^{-2R} \end{aligned}$$

solves the information-constrained ACOE (5.12) for $\pi_i$, i.e.,

$$\langle \pi_i, h_i\rangle + \lambda_i = D_{\pi_i}(R; c + Qh_i), \tag{7.4}$$

and that the MRS control law $\Phi_i$ achieves the value of the distortion-rate function in (7.4) and belongs to the set $\mathcal{K}_{\pi_i}(R)$. Then the desired results will follow.

### A.5.1 Existence, uniqueness, and closed-loop stability

In preparation for the proof, we first demonstrate that $m_1 = m_1(R)$ indeed exists and is positive, and that the steady-state variances $\sigma_1^2$ and $\sigma_2^2$ are finite and positive. This will imply that the closed-loop system (6.11) is stable and ergodic with the unique invariant distribution $\pi_i$.

**Lemma A.5.1** *For all nonzero $a, b$ and all $p, q, R > 0$, Eq. (6.2) has a unique positive root $m_1 = m_1(R)$.*

**Remark** Uniqueness and positivity of $m_2$ follow from well-known results on the standard LQG problem.

**Proof** Consider the function

$$F(m) \triangleq p + ma^2 + \frac{(mab)^2}{q + mb^2}(e^{-2R} - 1).$$

We have

$$\begin{aligned} F'(m) &= a^2 + \frac{(ab)^2(e^{-2R} - 1)\,(2q + mb^2)\,m}{(q + mb^2)^2} \\ F''(m) &= \frac{2a^2b^6(e^{-2R} - 1)}{(q + mb^2)^3}. \end{aligned}$$

whence it follows that $F$ is strictly increasing and concave for $m > -q/b^2$. Therefore, the fixed-point equation $F(m) = m$ has a unique positive root $m_1(R)$. (See the proof of Proposition 4.1 in [18] for a similar argument.)

**Lemma A.5.2** *For all $a, b \neq 0$ with $a^2 < 1$ and $p, q, R > 0$,*

$$e^{-2R}a^2 + (1 - e^{-2R})(a + bk_i)^2 \in (0, 1), \quad i = 1, 2. \tag{7.5}$$

*Consequently, the steady-state variance $\sigma_i^2 = \sigma_i^2(R)$ defined in (6.5) is finite and positive.*

**Proof** We write

$$\begin{aligned} &e^{-2R}a^2 + (1 - e^{-2R})(a + bk_i)^2 \\ &= e^{-2R}a^2 + (1 - e^{-2R})\left[a\left(1 - \frac{m_i b^2}{q + m_i b^2}\right)\right]^2 \\ &\le a^2, \end{aligned}$$

where the second step uses (6.4) and the last step follows from the fact that $q > 0$ and $m_i > 0$ (cf. Lemma A.5.1). By the assumption of open-loop stability ($a^2 < 1$), we get (7.5).

### A.5.2 A quadratic ansatz for the relative value function

Let $h(x) = mx^2$ for an arbitrary $m > 0$. Then

$$\begin{aligned} Qh(x,u) &= \int_{\mathsf{X}} h(y) Q(\mathrm{d}y|x,u) \\ &= m(ax+bu)^2 + m\sigma^2, \end{aligned}$$

and

$$\begin{aligned} &c(x,u) + Qh(x,u) \\ &= m\sigma^2 + px^2 + qu^2 + m(ax+bu)^2 \\ &= m\sigma^2 + (q+mb^2)u^2 + 2mabux + (p+ma^2)x^2. \end{aligned}$$

Let us complete the squares by letting $\tilde{x} = -\dfrac{mab}{q+mb^2}x$:

$$\begin{aligned} &c(x,u) + Qh(x,u) \\ &= m\sigma^2 + (q+mb^2)(u-\tilde{x})^2 + \left(p + ma^2 - \frac{m^2(ab)^2}{q+mb^2}\right)x^2. \end{aligned}$$

Therefore, for any $\pi \in \mathcal{P}(\mathsf{X})$ and any $\Phi \in \mathcal{M}(\mathsf{U}|\mathsf{X})$, such that $\pi$ and $\pi\Phi$ have finite second moments, we have

$$\begin{aligned} &\langle \pi \otimes \Phi, c + Qh - h \rangle \\ &= m\sigma^2 + \left(p + m(a^2-1) - \frac{(mab)^2}{q+mb^2}\right)\int_{\mathsf{X}} x^2 \pi(\mathrm{d}x) \\ &\qquad + (q+mb^2)\int_{\mathsf{X}\times\mathsf{U}} (u-\tilde{x})^2 \pi(\mathrm{d}x)\Phi(\mathrm{d}u|x). \end{aligned}$$

### A.5.3 Reduction to a static Gaussian rate-distortion problem

Now we consider the Gaussian case $\pi = N(0, \upsilon)$ with an arbitrary $\upsilon > 0$. Then for any $\Phi \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ we have

$$\begin{aligned}
&\langle \pi \otimes \Phi, c + Qh - h \rangle \\
&= m\sigma^2 + \left( p + m(a^2 - 1) - \frac{(mab)^2}{q + mb^2} \right) \upsilon \\
&\qquad + (q + mb^2) \int_{\mathsf{X}\times\mathsf{U}} (u - \tilde{x})^2 \pi(\mathrm{d}x) \Phi(\mathrm{d}u|x).
\end{aligned}$$

We need to minimize the above over all $\Phi \in \mathcal{I}_\pi(R)$.

If $X$ is a random variable with distribution $\pi = N(0, \upsilon)$, then its scaled version

$$\tilde{X} = -\frac{mab}{q + mb^2} X \equiv kX \tag{7.6}$$

has distribution $\tilde{\pi} = N(0, \tilde{\upsilon})$ with $\tilde{\upsilon} = k^2 \upsilon$. Since the transformation $X \mapsto \tilde{X}$ is one-to-one and mutual information is invariant under one-to-one transformations [3], we can write

$$\begin{aligned}
&D_\pi(R; c + Qh) - \langle \pi, h \rangle \\
&= \inf_{\Phi \in \mathcal{I}_\pi(R)} \langle \pi \otimes \Phi, c + Qh - h \rangle & (7.7) \\
&= m\sigma^2 + \left( p + m(a^2 - 1) - \frac{(mab)^2}{q + mb^2} \right) \upsilon \\
&\quad + (q + mb^2) \inf_{\tilde{\Phi} \in \mathcal{I}_{\tilde{\pi}}(R)} \int_{\mathsf{X}\times\mathsf{U}} (u - \tilde{x})^2 \tilde{\pi}(\mathrm{d}\tilde{x}) \tilde{\Phi}(\mathrm{d}u|\tilde{x}). & (7.8)
\end{aligned}$$

We recognize the infimum in (7.8) as the DRF for the Gaussian distribution $\tilde{\pi}$ w.r.t. the squared-error distortion $d(\tilde{x}, u) = (\tilde{x} - u)^2$. (For the reader's convenience, the Appendix contains a summary of standard results on the

Gaussian DRF.) Exploiting this fact, we can write

$$
\begin{aligned}
&D_\pi(R; c+Qh) - \langle \pi, h\rangle \\
&= m\sigma^2 + \left(p + m(a^2-1) - \frac{(mab)^2}{q+mb^2}\right)\upsilon + (q+mb^2)\tilde{\upsilon}e^{-2R} \\
&= m\sigma^2 + \left(p + m(a^2-1) + \frac{(mab)^2}{q+mb^2}(e^{-2R}-1)\right)\upsilon && (7.9) \\
&= m\sigma^2 + \left(p + m(a^2-1) - \frac{(mab)^2}{q+mb^2}\right)\upsilon \\
&\qquad + (q+mb^2)k^2\upsilon e^{-2R}, && (7.10)
\end{aligned}
$$

where Eqs. (7.9) and (7.10) are obtained by collecting appropriate terms and using the definition of $k$ from (7.6). We can now state the following result:

**Lemma A.5.3** *Let $\pi_i = N(0,\sigma_i^2)$, $i = 1,2$. Then the pair $(h_i, \lambda_i)$ solves the information-constrained ACOE (7.4). Moreover, for each $i$ the controller $\Phi_i$ defined in (6.7) achieves the DRF in (7.4) and belongs to the set $\mathcal{K}_{\pi_i}(R)$.*

**Proof** If we let $m = m_1$, then the second term in (7.9) is identically zero for any $\upsilon$. Similarly, if we let $m = m_2$, then the second term in (7.10) is zero for any $\upsilon$. In each case, the choice $\upsilon = \sigma_i^2$ gives (7.4).

From the results on the Gaussian DRF (see Appendix), we know that, for a given $\upsilon > 0$, the infimum in (7.8) is achieved by the kernel

$$K_i^*(\mathrm{d}u|\tilde{x}) = \gamma\left(u; (1-e^{-2R})\tilde{x}, e^{-2R}(1-e^{-2R})\tilde{\upsilon}\right)\mathrm{d}u.$$

Setting $\upsilon = \sigma_i^2$ for $i = 1,2$ and using the fact that $\tilde{x} = k_i x$ and $\tilde{\upsilon} = k_i^2\sigma_i^2$, we see that the infimum over $\Phi$ in (7.7) in each case is achieved by the composition of the deterministic mapping

$$\tilde{x} = k_i x = -\frac{m_i ab}{q + m_i b^2} x \tag{7.11}$$

with $K_i^*$. It is easy to see that this composition is precisely the stochastic control law $\Phi_i$ defined in (6.7). Since the map (7.11) is one-to-one, we have

$$I(\pi_i, \Phi_i) = I(\tilde{\pi}_i, K_i^*) = R.$$

Therefore, $\Phi_i \in \mathcal{I}_{\pi_i}(R)$.

It remains to show that $\Phi_i \in \mathcal{K}_{\pi_i}$, i.e., that $\pi_i$ is an invariant distribution of $Q_{\Phi_i}$. This follows immediately from the fact that $Q_{\Phi_i}$ is realized as

$$Y = (a + bk_i e^{-2R})X + bk_i e^{-R}\sqrt{1 - e^{-2R}}V^{(i)} + W,$$

where $V^{(i)} \sim N(0, \sigma_i^2)$ and $W \sim N(0, \sigma^2)$ are independent of one another and of $X$ [cf. (7.3)]. If $X \sim \pi_i$, then the variance of the output $Y$ is equal to

$$\begin{aligned}
&(a + bk_i e^{-2R})^2 \sigma_i^2 + (bk_i)^2 e^{-2R}(1 - e^{-2R})\sigma_i^2 + \sigma^2 \\
&= \left[e^{-2R}a^2 + (1 - e^{-2R})\,(a + bk_i)^2\right] \sigma_i^2 + \sigma^2 \\
&= \sigma_i^2,
\end{aligned}$$

where the last line follows from (6.5). This completes the proof of the lemma.

Putting together Lemmas A.5.1–A.5.3, we obtain Theorem 6.1.1.

# REFERENCES


[1] C. Sims, “Implications of rational inattention,” *Journal of Monetary Economics*, vol. 50, no. 3, pp. 665–690, 2003.

[2] C. Sims, “Rational inattention: Beyond the linear-quadratic case,” *The American Economic Review*, vol. 96, no. 2, pp. 158–163, 2006.

[3] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*. Holden-Day, 1964.

[4] F. Matejka and A. McKay, “Rational Inattention to discrete choices: A new foundation for the multinomial logit model,” 2011, working paper.

[5] F. Matejka and A. McKay, “Simple market equilibria with rationally inattentive consumers,” *The American Economic Review*, vol. 102, no. 3, pp. 24–29, 2012.

[6] L. Peng, “Learning with information capacity constraints,” *Journal of Financial and Quantitative Analysis*, vol. 40, no. 2, pp. 307–329, 2005.

[7] G. Moscarini, “Limited information capacity as a source of inertia,” *Journal of Economic Dynamics and Control*, vol. 28, no. 10, pp. 2003–2035, September 2004.

[8] Y. Luo, “Consumption dynamics under information processing constraints,” *Review of Economic Dynamics*, vol. 11, no. 2, pp. 366–385, April 2008.

[9] M. Woodford, “Information-constrained state-dependent pricing,” *Journal of Monetary Economics*, vol. 56(S), pp. 100–124, October 2009.

[10] P. Varaiya and J. Walrand, “Causal coding and control for Markov chains,” *Systems and Control Letters*, vol. 3, pp. 189–192, 1983.

[11] R. Bansal and T. Başar, “Simultaneous design of measurement and control strategies for stochastic systems with feedback,” *Automatica*, vol. 25, no. 5, pp. 679–694, 1989.

[12] S. Tatikonda and S. Mitter, “Control over noisy channels,” *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1196–2001, July 2004.

[13] S. Tatikonda, A. Sahai, and S. Mitter, "Stochastic linear control over a communication channel," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1549–1561, September 2004.

[14] C.-K. Ko, X. Gao, S. Prajna, and L. J. Schulman, "On scalar LQG control with communication cost," in *Proc. 44th IEEE CDC and ECC*, 2005, pp. 2805–2810.

[15] S. Yüksel and T. Linder, "Optimization and convergence of observation channels in stochastic control," *SIAM Journal on Control and Optimization*, vol. 50, no. 2, pp. 864–887, 2012.

[16] T. Berger, *Rate Distortion Theory, A Mathematical Basis for Data Compression.* Prentice Hall, 1971.

[17] P. E. Caines, *Linear Stochastic Systems.* Wiley, 1988.

[18] D. P. Bertsekas, *Dynamic Programming and Optimal Control.* Belmont, MA: Athena Scientific, 2000, vol. 1.

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

[20] S. P. Meyn and G. Mathew, "Shannon meets Bellman: Feature based Markovian models for detection and optimization," in *Proc. 47th IEEE CDC*, 2008, pp. 5558–5564.

[21] R. L. Stratonovich, "On value of information," *Izvestiya of USSR Academy of Sciences, Technical Cybernetics*, vol. 5, pp. 3–12, 1965.

[22] R. L. Stratonovich, *Information Theory.* Moscow, USSR: Sovetskoe Radio, 1975.

[23] I. Csiszár, "On an extremum problem of information theory," *Studia Scientiarum Mathematicarum Hungarica*, vol. 9, pp. 57–71, 1974.

[24] P. R. Halmos and L. J. Savage, "Application of the Radon–Nikodym theorem to the theory of sufficient statistics," *Annals of Mathematical Statistics*, vol. 20, pp. 225–241, 1949.

[25] G. Picci, "Some connections between the theory of sufficient statistics and the identifiability problem," *SIAM Journal on Applied Mathematics*, vol. 33, no. 3, pp. 383–398, 1977.

[26] Y. Bar-Shalom and E. Tse, "Dual effect, certainty equivalence, and separation in stochastic control," *IEEE Transactions on Automatic Control*, vol. 19, pp. 494–500, 1974.

[27] V. S. Borkar, "Convex analytic methods in Markov decision processes," in *Handbook of Markov Decision Processes*, E. Feinberg and A. Shwartz, Eds. Boston, MA: Kluwer, 2001.

[28] S. P. Meyn, *Control Techniques for Complex Networks.* Cambridge Univ. Press, 2008.

[29] A. Manne, "Linear programming and sequential decisions," *Management Science*, vol. 6, pp. 257–267, 1960.

[30] O. Hernández-Lerma and J. B. Lasserre, *Markov Chains and Invariant Probabilities.* Birkhäuser, 2003.

[31] M. Sion, "On general minimax theorems," *Pacific Journal of Mathematics*, vol. 8, pp. 171–176, 1958.

[32] P. Dupuis and R. S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations.* New York: Wiley, 1997.