Rational Inattention in Scalar LQG Control

Ehsan Shafieepoorfard and Maxim Raginsky

Abstract— **Motivated in part by the "rational inattention" framework of information-constrained decision-making by economic agents, we have recently introduced a general model for average-cost optimal control of Markov processes subject to mutual information constraints [1]. The optimal information-constrained control problem reduces to an infinite-dimensional convex program and admits a decomposition based on the Bellman error, which is the object of study in approximate dynamic programming.**

In this paper, we apply our 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 exhibit 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.

I. Introduction

The framework of "rational inattention," introduced into mathematical economics by Christopher Sims [2], [3], aims to model decision-making by agents who maximize expected utility (or minimize expected cost) given available information (hence "rational"), but are capable of handling only a limited amount of information (hence "inattention"). The main idea behind rational inattention is that such agents should design not only the policy that maps available information to actions, but also the *observation channel* that provides information about the state of the system of interest subject to the information constraint. Quantitatively, this constraint is stated in terms of an upper bound on the *mutual information* in the sense of Shannon [4] between the state of the system and the observation available to the agent.

Following the initial publications of Sims [2], [3], researchers have examined rational-inattention (or information-constrained) variants of many standard economic decision-making problems, both static and dynamic — see, e.g., [5]–[10]. 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. 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 [11]–[16]).

In an earlier paper [1], we have initiated the development of a general theory for optimal control subject to mutual information constraints. We focused on the average-cost optimal control problem for Markov processes and showed 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 object of study in approximate dynamic programming [17], [18]. This decomposition reveals a fundamental connection between information-constrained controller design and *rate-distortion theory* [19], a branch of information theory that deals with optimal compression of data subject to information constraints. (See [1] and Section II of this paper for more details.)

In this paper, we use the theoretical methodology developed in [1] to analyze the classic linear-quadratic-Gaussian (LQG) control problem [20], [21] in the rational inattention framework. Various information- or communication-constrained versions of the LQG problem have been studied in the literature (see, e.g., [2], [12], [14], [15]). In particular, Sims [2] 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 [2] 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.

Our main contribution is an explicit construction of rationally inattentive control laws for the LQG problem from first principles, using the convex-analytic approach we have developed in [1]. 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 [2] 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

E. Shafieepoorfard and M. Raginsky are with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana–Champaign, Urbana, IL 61801, USA; {shafiee1,maxim}@illinois.edu.

Research supported in part by the NSF under CAREER award no. CCF-1254041 and grant no. ECCS-1135598.

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.

To keep things simple, we focus on the scalar LQG problem and leave the general vector case for future work.

The remainder of the paper is structured as follows. Section II gives a concise summary of the results of our earlier paper [1]. The LQG problem with mutual information constraint is then introduced in Section III. Section IV contains our main result (Theorem 2) and discusses its consequences. The proof of Theorem 2 is given in Section V, followed by concluding remarks in Section VI. Background material on the Gaussian distortion-rate function, which is needed in the proof, is given in the Appendix.

II. Preliminaries and background

To keep the presentation self-contained, we give a brief summary of the results of our earlier paper [1].

A. Some definitions 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 σ-algebras. If X is such a space, then $\mathscr{B}(\mathsf{X})$ will denote its Borel σ-algebra, and $\mathscr{P}(\mathsf{X})$ will denote the space of all probability measures on $(\mathsf{X}, \mathscr{B}(\mathsf{X}))$. We use bilinear form notation for expectations:

$$\langle \mu, f \rangle \triangleq \int_{\mathsf{X}} f(x)\mu(\mathrm{d}x), \qquad \forall f \in L^1(\mu).$$

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

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

The *relative entropy* (or *information divergence*) between any two $\mu, \nu \in \mathscr{P}(\mathsf{X})$ [4] 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 $\mu \in \mathscr{P}(\mathsf{X})$ and any Markov kernel $K \in \mathscr{M}(\mathsf{Y}|\mathsf{X})$, we can define a probability measure $\mu \otimes K$ on the product space $(\mathsf{X} \times \mathsf{Y}, \mathscr{B}(\mathsf{X}) \otimes \mathscr{B}(\mathsf{Y}))$ via its action on the rectangles $A \times B$, $A \in \mathscr{B}(\mathsf{X}), B \in \mathscr{B}(\mathsf{Y})$:

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

Note that $\mu \otimes K(\mathsf{X} \times B) = \mu K(B)$ for all $B \in \mathscr{B}(\mathsf{X})$. The *Shannon mutual information* [4] in the pair (μ, K) is

$$I(\mu, K) \triangleq D(\mu \otimes K \| \mu \otimes \mu K),$$

where, for any $\mu \in \mathscr{P}(\mathsf{X})$ and $\nu \in \mathscr{P}(\mathsf{Y})$, $\mu \otimes \nu$ denotes the product measure defined via $(\mu \otimes \nu)(A \times B) \triangleq \mu(A)\nu(B)$ for all $A \in \mathscr{B}(\mathsf{X}), B \in \mathscr{B}(\mathsf{Y})$. In this paper, we use natural logarithms, so mutual information is measured in *nats*.

We will also need some notions from *rate-distortion theory* [19], which is a branch of information theory that deals with optimal compression of data subject to information constraints. Given a probability measure $\mu \in \mathscr{P}(\mathsf{X})$ and a measurable distortion function $d : \mathsf{X} \times \mathsf{Y} \to \mathbb{R}^+$, the *Shannon distortion-rate function* (DRF) of μ w.r.t. d is defined as

$$D_\mu(R) \triangleq \inf_{K \in \mathscr{I}_\pi(R)} \langle \mu \otimes K, d \rangle,$$

where

$$\mathscr{I}_\mu(R) \triangleq \{ K \in \mathscr{M}(\mathsf{Y}|\mathsf{X}) : I(\mu, K) \le R \}$$

is the set of all Markov kernels with X-valued input and Y-valued output, such that when the input has distribution μ, the resulting mutual information is no more than R nats.

B. System model

Consider a time-invariant controlled stochastic system with state space X and control (or action) space U, initial state distribution $\mu \in \mathscr{P}(\mathsf{X})$, and controlled Markov transition kernel $Q \in \mathscr{M}(\mathsf{X}|\mathsf{X} \times \mathsf{U})$. A *Markov randomized stationary* (MRS) *control law* is specified by a Markov kernel $\Phi \in \mathscr{M}(\mathsf{U}|\mathsf{X})$. Given Φ, the evolution of the system is described by the X-valued state process $\{X_t\}_{t=1}^\infty$ and the U-valued control process $\{U_t\}_{t=1}^\infty$. These processes are defined on a common probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and have the causal ordering

$$X_1, U_1, \dots, X_t, U_t, \dots,$$

where, $\mathbb{P}$-almost surely,

- $\mathbb{P}(X_1 \in A) = \mu(A)$ for all $A \in \mathscr{B}(\mathsf{X})$
- $\mathbb{P}(U_t \in B | X^t, U^{t-1}) = \Phi(B|X_t)$ for all $t = 1, 2, \dots$ and all $B \in \mathscr{B}(\mathsf{U})$
- $\mathbb{P}(X_{t+1} \in C | X^t, U^t) = Q(C|X_t, U_t)$ for all $t = 1, 2 \dots$ and all $C \in \mathscr{B}(\mathsf{X})$

This specification ensures that, for each t, the next state X_{t+1} is conditionally independent of X^{t-1}, 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-1}, U^{t-1} given X_t. In other words, at each time t the controller takes as input only the most recent state X_t. (The restriction of the optimization domain to such memoryless control laws is not always optimal, but it can be justified from first principles for a wide class of control architectures [1, Sec. III].)

C. Information-constrained control problem

Given a measurable state-action cost function $c : \mathsf{X} \times \mathsf{U} \to \mathbb{R}^+$, the objective is to minimize the long-term average cost

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

over all MRS control laws $\Phi \in \mathcal{M}(\mathsf{U}|\mathsf{X})$ satisfying the *information constraint*

$$\limsup_{t\to\infty} I(\mu_t, \Phi) \le R \tag{2}$$

for a given $R \ge 0$, where $\mu_t \in \mathscr{P}(\mathsf{X})$ is the state distribution at time t. When $R < +\infty$, this constraint ensures that the state-to-control transformation $X_t \to U_t$ must factor through a noisy observation channel with information capacity of no more than R nats per use, i.e., any realization of the control law must be a Markov chain

$$X_t \xrightarrow{K_t} Z_t \longrightarrow U_t,$$

where the observation Z_t takes values in some space Z, and $I(\mu_t, K_t) \le R$. Let $J_\mu(\Phi)$ denote the value of the objective (1) attained by a particular controller Φ, where $\mu = \mu_1$ is the initial state distribution. Thus, we seek an admissible control law that would minimize $J_\mu(\Phi)$ under the constraint (2).

D. A convex-analytic formulation

As we had shown in [1], the problem of finding an optimal information-constrained control law is best approached through the convex-analytic framework for Markov decision processes (see [17], [22]–[25]). Any MRS control law Φ induces a Markov kernel $Q_\Phi \in \mathcal{M}(\mathsf{X}|\mathsf{X})$ via

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

We say that Φ is *stable* if:

1) There exists a probability measure $\pi_\Phi \in \mathscr{P}(\mathsf{X})$ which is invariant under Q_Φ, 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 control laws. In the absence of any information constraint and under mild regularity conditions (which are satisfied in the LQG setting), it can be shown [23]–[25] that the optimal steady-state value of the average-cost control problem is

$$J^* = \inf_{\mu\in\mathscr{P}(\mathsf{X})} \inf_{\Phi\in\mathcal{M}(\mathsf{U}|\mathsf{X})} J_\mu(\Phi) = \inf_{\Phi\in\mathcal{K}} \langle \Gamma_\Phi, c\rangle. \tag{3}$$

If $\Phi^* \in \mathcal{K}$ achieves the infimum on the rightmost side of (3) and if the Markov kernel Q_{Φ^*} is ergodic, then the state distributions μ_t induced by Φ^* converge weakly to π_{Φ^*} regardless of the initial condition $\mu_1 = \mu$.

For the information-constrained problem, it is convenient to decompose the infimum over $\Phi \in \mathcal{K}$ in (3) by first fixing the candidate invariant distribution $\pi \in \mathscr{P}(\mathsf{X})$. For any $\pi \in \mathscr{P}(\mathsf{X})$, define the sets

$$\mathcal{K}_\pi \triangleq \left\{\Phi \in \mathcal{K} : \pi = \pi_\Phi\right\}$$
$$\mathscr{I}_\pi(R) \triangleq \left\{\Phi \in \mathcal{M}(\mathsf{U}|\mathsf{X}) : I(\pi,\Phi) \le R\right\}$$
$$\mathcal{K}_\pi(R) \triangleq \mathcal{K}_\pi \cap \mathscr{I}_\pi(R).$$

Then the optimal steady-state value of the information-constrained average-cost control problem is

$$J^*(R) \triangleq \inf_{\pi\in\mathscr{P}(\mathsf{X})} \inf_{\Phi\in\mathcal{K}_\pi(R)} \langle \pi\otimes\Phi, c\rangle. \tag{4}$$

We can summarize the results of [1] as follows:

Theorem 1. *For any $\pi \in \mathscr{P}(\mathsf{X})$, let*

$$J^*_\pi(R) \triangleq \inf_{\Phi\in\mathcal{K}_\pi(R)} \langle \pi\otimes\Phi, c\rangle. \tag{5}$$

Then

$$J^*_\pi(R) = \inf_{\Phi\in\mathscr{I}_\pi(R)} \sup_{h\in L^1(\pi)} \langle \pi\otimes\Phi, c+Qh-h\rangle \tag{6a}$$
$$= \sup_{h\in L^1(\pi)} \inf_{\Phi\in\mathscr{I}_\pi(R)} \langle \pi\otimes\Phi, c+Qh-h\rangle \tag{6b}$$

Suppose the infimum in (6a)–(6b) *is achieved by some $\Phi^* \in \mathcal{K}_\pi(R)$, and $J^*_\pi(R) < \infty$. Suppose also that there exist some $h \in L^1(\pi)$ and $\lambda \in \mathbb{R}^+$, such that*

$$\langle \pi, h\rangle + \lambda = D_\pi(R; c+Qh), \tag{7}$$

where

$$D_\pi(R; c+Qh) \triangleq \inf_{\Phi\in\mathscr{I}_\pi(R)} \langle \pi\otimes\Phi, c+Qh\rangle \tag{8}$$

is the DRF of π w.r.t. the distortion function $c+Qh$. Then Φ^* achieves the infimum in (8)*, and*

$$J^*_\pi(R) = J_\pi(\Phi^*) = \lambda.$$

Some remarks are in order. The function h in (6a)–(6b) 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 [17, Ch. 9].

If we let $\eta = \langle \pi\otimes\Phi, c\rangle$, then the function $c+Qh-h-\eta$ is the *Bellman error* associated with h. This object is used in approximate dynamic programming to quantify the deviation of a control law from optimality in terms of the error in the Bellman equation, also known as the *Average Cost Optimality Equation* (ACOE) [17], [18]. Moreover, we can interpret (7) as an *information-constrained* ACOE, and the standard ACOE can be recovered in the limit $R \to \infty$ [1]. When a nontrivial information constraint is present ($R < \infty$), the optimal steady-state value $J^*_\pi(R)$ is the optimal value of a *single-stage* (static) control problem under the same information constraint but with the cost function related to the Bellman error.

III. Information-constrained LQG problem

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

$$X_{t+1} = aX_t + bU_t + W_t, \qquad t \ge 1 \tag{9}$$

where $a, b \ne 0$ 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 σ^2. The initial state X_1 has some given distribution μ. Here, $\mathsf{X} = \mathsf{U} = \mathbb{R}$, and the controlled transition kernel $Q \in \mathscr{M}(\mathsf{X}|\mathsf{X} \times \mathsf{U})$ corresponding to (9) 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 m and variance σ^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 II-D, we seek a pair consisting of an invariant distribution $\pi \in \mathscr{P}(\mathsf{X})$ and an MRS control law $\Phi \in \mathscr{M}(\mathsf{U}|\mathsf{X})$ to attain the steady-state value (5) with $c(x,u) = px^2 + qu^2$ under the information constraint $I(\pi,\Phi) \le R$.

IV. Main result and some implications

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

Theorem 2. *Suppose that the system* (9) *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)

$$p + m(a^2 - 1) + \frac{(mab)^2}{q + mb^2}(e^{-2R} - 1) = 0, \tag{10}$$

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

$$p + m(a^2 - 1) - \frac{(mab)^2}{q + mb^2} = 0 \tag{11}$$

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

$$k_i = -\frac{m_i ab}{q + m_i b^2} \tag{12}$$

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

$$\sigma_i^2 = \frac{\sigma^2}{1 - \left[e^{-2R}a^2 + (1 - e^{-2R})(a + bk_i)^2\right]}. \tag{13}$$

Then

$$J^*(R) \le \min\Big(m_1\sigma^2, m_2\sigma^2 + (q + m_2 b^2)k_2^2\sigma_2^2 e^{-2R}\Big). \tag{14}$$

Also, let Φ_1 *and* Φ_2 *be two MRS control laws with Gaussian conditional densities*

$$\begin{aligned}\varphi_i(u|x) &= \frac{\mathrm{d}\Phi_i(u|x)}{\mathrm{d}u} \\ &= \gamma\left(u; (1 - e^{-2R})k_i x, (1 - e^{-2R})e^{-2R}k_i\sigma_i^2\right),\end{aligned} \tag{15}$$

and let $\pi_i = N(0, \sigma_i^2)$ *for* $i = 1, 2$. *Then the first term on the right-hand side of* (14) *is achieved by* Φ_1, *the second term is achieved by* Φ_2, *and*

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

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 2, we pause to examine a few consequences. First of all, the controllers Φ_1 and Φ_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 (10) reduces to the linear Lyapunov equation [20]

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

so the first term on the right-hand side of (14) is

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

On the other hand, using Eqs. (11) and (12), we can show that the second term is equal to the first term, so from (14)

$$J^*(0) \le \frac{p\sigma^2}{1 - a^2}. \tag{16}$$

Since this is the minimal average cost in the open-loop case, we have equality in (16). Also, the controllers Φ_1 and Φ_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 (9) 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 (10) reduces to the usual DARE (11). Hence, $m_1(\infty) = m_2$, and both terms on the right-hand side of (14) are equal to $m_2\sigma^2$. This gives

$$J^*(\infty) \le m_2\sigma^2. \tag{17}$$

Since this is the minimal average cost attainable in the scalar LQG control problem with perfect information, we have equality in (17), as expected. The controllers Φ_1 and Φ_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 Φ_1 and Φ_2 are no longer the same. However, they are both *stochastic* and have the form

$$U_t = k_i\left[(1 - e^{-2R})X_t + e^{-R}\sqrt{1 - e^{-2R}}\,V_t^{(i)}\right], \tag{18}$$

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

$$X_{t+1} = \left[a + \left(1 - e^{-2R}\right)bk_i\right]X_t + Z_t^{(i)}, \tag{19}$$

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

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

Theorem 2 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_{Φ_i}, $i = 1,2$, are ergodic. We also note that the two controllers in (18) can be realized as a cascade consisting of an additive white Gaussian noise (AWGN) channel and a linear gain:

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

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 Φ_1 and Φ_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 Φ_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, Φ_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 [2] on heuristic grounds for the information-constrained LQG problem with discounted cost. On the other hand, for Φ_1 both the noise variance σ_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 Φ_1 attains smaller steady-state cost for all sufficiently small values of R (see Figure 1), whereas Φ_2 outperforms Φ_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

$$R > \frac{1}{2}\log\frac{a^2-(a+bk_2)^2}{1-(a+bk_2)^2}, \tag{20}$$

where k_2 is the control gain defined in (12). Indeed, if R satisfies (20), then the steady-state variance σ_2^2 is well-defined, so the closed-loop system (19) with $i = 2$ is stable.

V. Proof of Theorem 2

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

$$h_1(x) = m_1x^2, \quad \lambda_1 = m_1\sigma^2$$
$$h_2(x) = m_2x^2, \quad \lambda_2 = m_2\sigma^2 + (q+m_2b^2)k_2^2\sigma_2^2e^{-2R}$$

solves the information-constrained ACOE (7) for π_i, i.e.,

$$\langle \pi_i, h_i\rangle + \lambda_i = D_{\pi_i}(R; c + Qh_i), \tag{21}$$

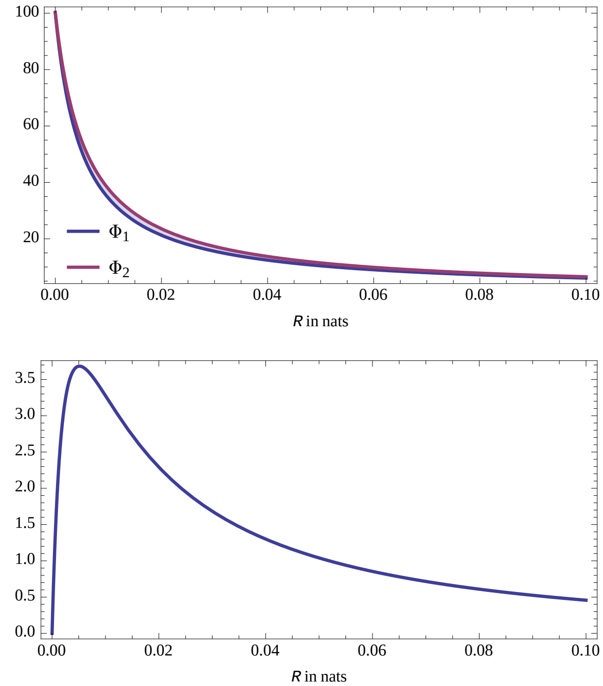

Fig. 1. Comparison of Φ_1 and Φ_2 at low information rates (top: steady-state values, bottom: difference of steady-state values of Φ_2 and Φ_1). System parameters: $a = 0.995, b = 1, \sigma^2 = 1$, cost parameters: $p = q = 1$.

and that the MRS control law Φ_i achieves the value of the distortion-rate function in (21) and belongs to the set $\mathscr{K}_{\pi_i}(R)$. Then the desired results will follow from Theorem 1. The proof is based on three lemmas (numbered 1–3 below), where Lemmas 1 and 2 show that the quantities listed in the statement of Theorem 2 are well-defined (thus ensuring closed-loop stability), while Lemma 3 shows that the proposed control laws satisfy the conditions of Theorem 1.

A. 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 σ_1^2 and σ_2^2 are finite and positive. This will imply that the closed-loop system (19) is stable and ergodic with the unique invariant distribution π_i.

Lemma 1. *For all nonzero a,b and all $p,q,R > 0$, Eq.* (10) *has a unique positive root $m_1 = m_1(R)$.*

Remark 1. *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

$$F'(m) = a^2 + \frac{(ab)^2(e^{-2R}-1)\left(2q+mb^2\right)m}{(q+mb^2)^2}$$

$$F''(m) = \frac{2a^2b^6(e^{-2R}-1)}{(q+mb^2)^3}$$

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 4.4.1 in [21] for a similar argument.) □

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

$$e^{-2R}a^2 + (1 - e^{-2R})(a + bk_i)^2 \in (0,1), \quad i = 1,2. \tag{22}$$

Consequently, the steady-state variance $\sigma_i^2 = \sigma_i^2(R)$ defined in (13) *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 (12) and the last step follows from the fact that $q > 0$ and $m_i > 0$ (cf. Lemma 1). By the assumption of open-loop stability ($a^2 < 1$), we get (22). □

B. 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

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

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

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

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

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

C. Reduction to a static Gaussian rate-distortion problem

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

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

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

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

$$\tilde{X} = -\frac{mab}{q + mb^2}X \equiv kX \tag{23}$$

has distribution $\tilde{\pi} = N(0, \tilde{v})$ with $\tilde{v} = k^2 v$. Since the transformation $X \mapsto \tilde{X}$ is one-to-one and mutual information is invariant under one-to-one transformations [4], we can write

$$\begin{aligned} &D_\pi(R; c + Qh) - \langle \pi, h \rangle \\ &= \inf_{\Phi \in \mathscr{I}_\pi(R)} \langle \pi \otimes \Phi, c + Qh - h \rangle \qquad (24) \\ &= m\sigma^2 + \left(p + m(a^2 - 1) - \frac{(mab)^2}{q + mb^2}\right)v \\ &\quad + (q + mb^2)\inf_{\tilde{\Phi} \in \mathscr{I}_{\tilde{\pi}}(R)} \int_{\mathsf{X}\times\mathsf{U}} (u - \tilde{x})^2 \tilde{\pi}(\mathrm{d}\tilde{x})\tilde{\Phi}(\mathrm{d}u|\tilde{x}). \qquad (25) \end{aligned}$$

We recognize the infimum in (25) 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

$$\begin{aligned} &D_\pi(R; c + Qh) - \langle \pi, h \rangle \\ &= m\sigma^2 + \left(p + m(a^2 - 1) - \frac{(mab)^2}{q + mb^2}\right)v + (q + mb^2)\tilde{v}e^{-2R} \\ &= m\sigma^2 + \left(p + m(a^2 - 1) + \frac{(mab)^2}{q + mb^2}(e^{-2R} - 1)\right)v \qquad (26) \\ &= m\sigma^2 + \left(p + m(a^2 - 1) - \frac{(mab)^2}{q + mb^2}\right)v \\ &\qquad + (q + mb^2)k^2 v e^{-2R}, \qquad (27) \end{aligned}$$

where Eqs. (26) and (27) are obtained by collecting appropriate terms and using the definition of k from (23). We can now state the following result:

Lemma 3. *Let $\pi_i = N(0, \sigma_i^2)$, $i = 1,2$. Then the pair (h_i, λ_i) solves the information-constrained ACOE* (21). *Moreover, for each i the controller Φ_i defined in* (15) *achieves the DRF in* (21) *and belongs to the set $\mathscr{K}_{\pi_i}(R)$.*

Proof. If we let $m = m_1$, then the second term in (26) is identically zero for any v. Similarly, if we let $m = m_2$, then the second term in (27) is zero for any v. In each case, the choice $v = \sigma_i^2$ gives (21).

From the results on the Gaussian DRF (see Appendix), we know that, for a given $v > 0$, the infimum in (25) is achieved by the kernel

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

Setting $v = \sigma_i^2$ for $i = 1,2$ and using the fact that $\tilde{x} = k_i x$ and $\tilde{v} = k_i^2 \sigma_i^2$, we see that the infimum over Φ in (24) in each case is achieved by the composition of the deterministic mapping

$$\tilde{x} = k_i x = -\frac{m_i ab}{q + m_i b^2}x \tag{28}$$

with K_i^*. It is easy to see that this composition is precisely the stochastic control law Φ_i defined in (15). Since the map

(28) is one-to-one, we have

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

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

It remains to show that $\Phi_i \in \mathscr{K}_{\pi_i}$, i.e., that π_i is an invariant distribution of Q_{Φ_i}. This follows immediately from the fact that Q_{Φ_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. (A.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 (13). This completes the proof of the lemma. □

Putting together Lemmas 1–3 and using Theorem 1, we obtain Theorem 2.

VI. Conclusions

The main contribution of this paper is 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 that have different performance in the presence of a nontrivial information constraint, but reduce to optimal deterministic control laws in the two extreme cases of no information and perfect information. Future work will include an extension to the vector LQG problem and a derivation of necessary conditions on the value of the information constraint to guarantee stabilizability.

Appendix

The Gaussian distortion-rate function

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

$$D_\mu(R) \triangleq \inf_{\substack{K \in \mathscr{M}(\mathbb{R}|\mathbb{R}):\\ I(\mu,K)\le R}} \int_{\mathbb{R}\times\mathbb{R}} (x-x')^2 \mu(\mathrm{d}x) K(\mathrm{d}x'|x) \qquad \text{(A.1)}$$

(where the mutual information is measured in nats). Let $\mu = N(0,\sigma^2)$. Then we have the following [19]:

- $D_\mu(R) = \sigma^2 e^{-2R}$
- The optimal kernel K^* that achieves the infimum in (A.1) has the form

  $$\begin{aligned}&K^*(\mathrm{d}x'|x)\\ &= \gamma\left(x'; (1-e^{-2R})x, (1-e^{-2R})e^{-2R}\sigma^2\right)\mathrm{d}x' \qquad \text{(A.2)}\end{aligned}$$

  and achieves the information constraint with equality: $I(\mu, K^*) = R$.
- K^* can be realized as a stochastic linear system

  $$X' = (1-e^{-2R})X + e^{-R}\sqrt{1-e^{-2R}}V, \qquad \text{(A.3)}$$

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

References

[1] E. Shafieepoorfard, M. Raginsky, and S. P. Meyn, "Rational inattention in controlled Markov processes," in *Proc. Amer. Control Conf.*, 2013, to appear.
[2] C. Sims, "Implications of rational inattention," *Journal of Monetary Economics*, vol. 50, no. 3, pp. 665–690, 2003.
[3] ——, "Rational inattention: Beyond the linear-quadratic case," *The American Economic Review*, vol. 96, no. 2, pp. 158–163, 2006.
[4] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*. Holden-Day, 1964.
[5] 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.
[6] L. Peng, "Learning with information capacity constraints," *Journal of Financial and Quantitative Analysis*, vol. 40, no. 2, pp. 307–329, 2005.
[7] Y. Luo, "Consumption dynamics under information processing constraints," *Review of Economic Dynamics*, vol. 11, no. 2, pp. 366–385, April 2008.
[8] M. Woodford, "Information-constrained state-dependent pricing," *Journal of Monetary Economics*, vol. 56(S), pp. 100–124, October 2009.
[9] F. Matejka and A. McKay, "Rational Inattention to discrete choices: A new foundation for the multinomial logit model," 2011, working paper.
[10] ——, "Simple market equilibria with rationally inattentive consumers," *The American Economic Review*, vol. 102, no. 3, pp. 24–29, 2012.
[11] P. Varaiya and J. Walrand, "Causal coding and control for Markov chains," *Systems and Control Letters*, vol. 3, pp. 189–192, 1983.
[12] 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.
[13] S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1196–2001, July 2004.
[14] 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.
[15] 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.
[16] 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.
[17] S. P. Meyn, *Control Techniques for Complex Networks*. Cambridge Univ. Press, 2008.
[18] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 4th ed. Belmont, MA: Athena Scientific, 2012, vol. 2.
[19] T. Berger, *Rate Distortion Theory, A Mathematical Basis for Data Compression*. Prentice Hall, 1971.
[20] P. E. Caines, *Linear Stochastic Systems*. Wiley, 1988.
[21] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Belmont, MA: Athena Scientific, 2005, vol. 1.
[22] A. Manne, "Linear programming and sequential decisions," *Management Science*, vol. 6, pp. 257–267, 1960.
[23] V. S. Borkar, "A convex analytic approach to Markov decision processes," *Probabillity Theory and Related Fields*, vol. 78, no. 583-602, 1988.
[24] O. Hernández-Lerma and J. B. Lasserre, "Linear programming and average optimality of Markov control processes on Borel spaces — unbounded costs," *SIAM Journal on Control and Optimization*, vol. 32, no. 2, pp. 480–500, March 1994.
[25] 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.