Three-Stage Separation Theorem for Information-Frugal LQG Control

Takashi Tanaka[1] Peyman Mohajerin Esfahani[2] Sanjoy K. Mitter[1]

***Abstract*— We propose a framework of LQG optimal control in which Massey's directed information from the state sequence to the control sequence incurs additional cost. The information-oriented cost in this study is motivated by a broad range of applications in which communication costs, privacy constraints, and bounded rationality of the decision-maker are present. Remarkably, we show that the most "information-frugal" LQG control policy in our framework can be realized by an attractively simple three-stage architecture comprising (i) a linear sensor with additive Gaussian noise, (ii) a Kalman filter, and (iii) a certainty equivalence controller. This result can be viewed as an integration of two previously known separation theorems: the filter-controller separation theorem in the standard LQG control theory, and the sensor-filter separation theorem that arises in zero-delay rate-distortion theory for Gauss-Markov sources. A tractable computational algorithm based on semidefinite programming is also available to synthesize an optimal policy.**

I. Introduction

In this paper, we propose a framework and methodology to identify the minimal information for real-time decision-making with acceptable accuracy. This is a fundamental question that has been raised throughout science and engineering. On one hand, this question is important because realistic decision-makers (either humans or digital computers) have bounded data-processing abilities (e.g., neuroscience [19], robotics [20], theory of bounded rationality [33], [27], networked control theory [26], [1], [14], [25], [42]); on the other hand, thorough understanding of the problem leads us to novel socio-engineering technologies (e.g., optimal privacy mechanisms [28], [9], [30]). A key feature of the algorithms needed in these contexts is a carefully designed data-selection mechanism that intentionally discards less important data from all available information to mitigate the exogenous information-oriented costs. However, this perspective is seldom discussed in control literature.

Although the interplay between control and information has been extensively studied in the aforementioned *networked control theory* literature, our angle of attack in this paper is different from many of the existing approaches. As of today, the majority of networked control literature is centered around the "control over communication channel" problem where a channel model is given as part of the problem set-up. Examples include noiseless digital channels (quantizers) [8], [6], [12], [15], noisy discrete memoryless

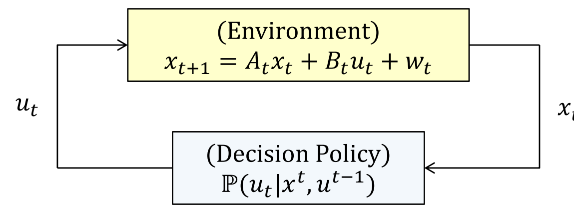

Fig. 1. "Information-frugal" LQG control problem.

channels [2], Gaussian channels [40], [5], [42], packet-dropping channels [13], and fading channels [11]. Although such considerations are motivated by realistic communication models, the results tend to be case-specific and are not suitable for identifying the minimal information for control in general contexts. In contrast, our programme in this paper is to carve out the fundamental trade-off between the best achievable control performance and the required data-rate directly, without assuming any specific communication models. In this sense, our results can be compared with [10], [22], [31]. However, our concern is different from [29], where the main purpose of the study was to characterize invariant properties of communication channels as far as control is concerned.

In order to quantify the minimal information needed to achieve desired control performance, we consider a general linear-Gaussian control system in Figure 1 and how Massey's directed information from the state sequence $\mathbf{x}_t$ of the plant to the control sequence $\mathbf{u}_t$ can be minimized while achieving desired LQG control performances. We refer to this framework as *information-frugal LQG control*, whose precise description is available in Section II. The relevance of directed information is previously discussed in [38], [10], [32], and we will revisit its interpretation again in the same section.

As the first part of the main result, we show that the optimal decision policy for the information-frugal LQG control in Figure 1 is realized by an attractive "sensor-filter-controller" separation structure. (Figure 4 depicts this architecture.) More precisely, we show that an optimal decision at every time step can be made by first observing $\mathbf{x}_t$ through a carefully designed MIMO linear sensor mechanism, which acquires "just enough" information for control purposes. Sensor outputs are processed by a causal and recursive estimator (Kalman filter) and then a certainty equivalence controller produces $\mathbf{u}_t$. It is remarkable that this simple three-stage policy outperforms all other (Borel-measurable) policies. The result in this part can be viewed as an integration of the previously known filter-controller separation principle

[1]TT and SM are with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139 USA. {ttanaka, mitter}@mit.edu.

[2]PME is with the automatic Control Laboratory, ETH Zurich, 8092 Zurich, Switzerland mohajerin@control.ee.ethz.ch.

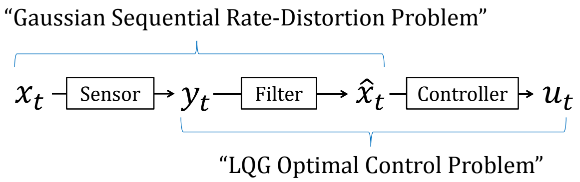

Fig. 2. Channel-filter-controller separation principle: integration of the channel-filter and filter-controller separation principles.

in the LQG control theory [41] and the recent channel-filter separation principle in the Gaussian sequential rate-distortion theory [38]; Figure 2 illustrates these principles.

As the second part of the main result, we show that the optimal policy can be synthesized by an efficient numerical algorithm. Namely, besides the optimal architecture of the control policy, another key result is that the aforementioned MIMO sensor can be constructed by means of semidefinite programming (SDP). It turns out that the use of SDP is crucial here, since we are not aware of a simple analytical expression of the optimal signal-to-noise ratio (SNR) for the MIMO sensor. The proposed SDP-based synthesis allows us to obtain a systematic framework representing the trade-off function between the controller performance and the minimal data-rate for general multi-variable and time-varying control systems. In this paper, we do not discuss the operational achievability (i.e., how to design source quantizers, encoders and decoders). However, it is expected that the structural results in this paper will provide useful insights for designing practical source coding schemes, which is important future work.

The rest of the paper is organized as follows. The main problem is formally introduced in Section II, where we also discuss its interpretations. Main results of this paper are summarized in Section III. Section IV is devoted to deriving the main theorem. Finally, we conclude in Section V and discuss potential directions as future work. To save space and improve the readability of the paper, some technical proofs are provided in the appendix and an online report available at [36].

Notation

Calligraphic symbols such as $\mathcal{X} \subseteq \mathbb{R}^n$ are used to denote subsets of Euclidean spaces. Borel σ-algebra on $\mathcal{X}$ with respect to the usual topology on $\mathbb{R}^n$ is denoted by $\mathcal{B}_\mathcal{X}$. An $(\mathcal{X}, \mathcal{B}_\mathcal{X})$-valued random variable is denoted by a bold symbol $\mathbf{x}$. We use a notation $\mathbf{x}^t \triangleq (\mathbf{x}_1, \cdots, \mathbf{x}_t)$. The probability measure induced by a random variable $\mathbf{x}$ is denoted by $\mathbb{P}_\mathbf{x}$ or $\mathbb{P}(\mathrm{d}x)$. A Gaussian random variable $\mathbf{x}$ with mean μ and covariance Σ is denoted by $\mathbf{x} \sim \mathcal{N}(\mu, \Sigma)$. The relative entropy of $\mathbb{Q}$ from $\mathbb{P}$ is a non-negative quantity defined by

$$D(\mathbb{P}\|\mathbb{Q}) \triangleq \begin{cases} \int \log_2 \frac{\mathbb{P}(\mathrm{d}x)}{\mathbb{Q}(\mathrm{d}x)} \mathbb{P}(\mathrm{d}x) & \text{if } \mathbb{P} \ll \mathbb{Q} \\ +\infty & \text{otherwise} \end{cases}$$

where $\mathbb{P} \ll \mathbb{Q}$ means that $\mathbb{P}$ is absolutely continuous with respect to $\mathbb{Q}$, and $\frac{\mathbb{P}(\mathrm{d}x)}{\mathbb{Q}(\mathrm{d}x)}$ denotes the Radon-Nikodym derivative. We say that a probability measure $\mathbb{P}$ on $\mathbb{R}^n$ admits a density if $\mathbb{P} \ll \mathbb{L}$, where $\mathbb{L}$ is the Lebesgue measure on $\mathbb{R}^n$. We denote by $\text{supp}(\mathbb{P}) \subseteq \mathbb{R}^n$ the support of $\mathbb{P}$ defined as the smallest closed set whose complement has measure 0. If $\mathbf{x}$ and $\mathbf{y}$ have a joint probability distribution $\mathbb{P}(\mathrm{d}x, \mathrm{d}y)$ with marginals $\mathbb{P}(\mathrm{d}x)$ and $\mathbb{P}(\mathrm{d}y)$, then the mutual information between $\mathbf{x}$ and $\mathbf{y}$ is defined by $I(\mathbf{x}; \mathbf{y}) \triangleq D(\mathbb{P}(\mathrm{d}x, \mathrm{d}y)\|\mathbb{P}(\mathrm{d}x)\mathbb{P}(\mathrm{d}y))$, where $\mathbb{P}(\mathrm{d}x)\mathbb{P}(\mathrm{d}y)$ is the product measure. A stochastic kernel $\mathbb{P}(\mathrm{d}x|z)$ is a probability measure in the first argument and a (Borel) measurable function in the second argument. The conditional mutual information is defined by $I(\mathbf{x}; \mathbf{y}|\mathbf{z}) \triangleq \int D(\mathbb{P}(\mathrm{d}x, \mathrm{d}y|z)\|\mathbb{P}(\mathrm{d}x|z)\mathbb{P}(\mathrm{d}y|z))\mathbb{P}(\mathrm{d}z)$. The entropy of a discrete random variable $\mathbf{x}$ with probability mass function $\mathbb{P}(x_i)$ is defined by $H(\mathbf{x}) \triangleq -\sum_i \mathbb{P}(x_i) \log_2 \mathbb{P}(x_i)$.

II. Information-Frugal LQG Control Problem

A. Problem description

Consider the linear time-varying stochastic difference equation

$$\mathbf{x}_{t+1} = A_t\mathbf{x}_t + B_t\mathbf{u}_t + \mathbf{w}_t, \quad t = 1, \cdots, T, \tag{1}$$

where $\mathbf{x}_t$ is an $\mathbb{R}^n$-valued stochastic process describing the current state of the system, and $\mathbf{u}_t$ is the control process synthesized by the decision maker. We assume that the process $\mathbf{x}_t$ starts from an initial state with the known distribution $\mathbf{x}_1 \sim \mathcal{N}(0, P_{1|0})$ for some $P_{1|0} \succ 0$, which is independent from the process noises $\mathbf{w}_t \sim \mathcal{N}(0, W_t)$ with $W_t \succ 0$.

The main objective in this study is to synthesize a decision policy that, informally speaking, "consumes" the least amount of information of the state process (measured in *bits*) among all policies achieving a desired control performance. In this context, the control policy is described via a sequence of (Borel measurable) stochastic kernels denoted by $\gamma \triangleq \prod_{t=1}^T \mathbb{P}(\mathrm{d}u_t|x^t, u^{t-1})$; see Figure 1 for a visual interaction between the state and control processes. A thorough treatment of stochastic kernels can be found in [3]. Let Γ denote the space of such policies.

The performance of the policy $\gamma \in \Gamma$ is quantified based on two criteria:

(i) the LQG control cost

$$J_\gamma(\mathbf{x}^{T+1}, \mathbf{u}^T) \triangleq \sum\nolimits_{t=1}^T \mathbb{E}\left(\|\mathbf{x}_{t+1}\|_{Q_t}^2 + \|\mathbf{u}_t\|_{R_t}^2\right); \tag{2}$$

(ii) and the *directed information*

$$I_\gamma(\mathbf{x}^T \to \mathbf{u}^T) \triangleq \sum\nolimits_{t=1}^T I_\gamma(\mathbf{x}^t; \mathbf{u}_t|\mathbf{u}^{t-1}). \tag{3}$$

The directed information quantity in (3) is defined using the conditional mutual information evaluated with respect to the joint probability measure induced by the policy γ. For the sake of notational simplicity we may use "I" instead of "I_γ" when the underlying policy γ is clear from the context.

The main problem studied in this paper is defined as

$$\min_{\gamma \in \Gamma} \quad I_\gamma(\mathbf{x}^T \to \mathbf{u}^T) \tag{4a}$$

$$\text{s.t.} \quad J_\gamma(\mathbf{x}^{T+1}, \mathbf{u}^T) \leq D, \tag{4b}$$

where $D > 0$ is a given parameter reflecting the desired upper bound for the LQG control cost. We denote the optimal value of the problem (4) by $R(D)$. In this paper we provide an optimization-based characterization, more specifically an SDP reformulation, of the problem (4) which is amenable to existing computational tools for numerical purposes.

B. Interpretation

Before proceeding with the main result of this paper, in this subsection we elaborate further details concerning the problem (4). The notion of directed information is first introduced by Massey [23] based on Marko's earlier work [21]. The directed information is an important concept in the context of feedback capacity of communication channels and has already been studied in the literature [17], [39].

Directed information (4a) reflects *only* the information flow *from* the state process $\mathbf{x}_t$ *to* the control process $\mathbf{u}_t$, while the standard mutual information also captures the information flow from control to state processes. This feature makes the notion of directed information particularly interesting in the context of feedback control design.

As depicted in Figure 1, one can observe that the interaction between $\mathbf{x}_t$ and $\mathbf{u}_t$ is bidirectional. Namely, there is an information flow from $\mathbf{x}_t$ to $\mathbf{u}_t$ through the decision policy, and another flow from $\mathbf{u}_t$ to $\mathbf{x}_t$ through the environment. This bilateral information flow between the state and control processes leads to a decomposition of mutual information between these processes into two directed information terms as

$$I(\mathbf{x}^T;\mathbf{u}^T) = I(\mathbf{x}^T \to \mathbf{u}^T) + I(\mathbf{u}_+^{T-1} \to \mathbf{x}^T),$$

where the sequence $\mathbf{u}_+^{T-1} = (0, \mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{T-1})$ denotes an index-shifted version of $\mathbf{u}^T$. The above equality is also called *conservation of information* [24].

Given the above interpretation, it is clear that the natural quantity in the design of the feedback policy with minimal information consumption is indeed the directed information, as it reflects the information processing cost for a decision-maker in most of real applications. In this regard, in the remainder of this subsection we discuss an application of the problem (4) to networked control systems.

Consider a feedback control system in Figure 3, where the sensor data must be transmitted to the controller in the form of binary sequence of length r_t. Let us denote this sequence by $\{\mathbf{z}_t\}_{t=1}^T \subset \{0,1\}^{r_t}$. We assume that the "sensor + encoder" block is modeled by a stochastic kernel $\mathbb{P}(\mathrm{d}z_t|x^t, z^{t-1}, u^{t-1})$, while the "decoder + controller" block is modeled by another stochastic kernel $\mathbb{P}(\mathrm{d}u_t|z^t, u^{t-1})$. Notice that the composition of these stochastic kernels uniquely characterizes a policy $\gamma \triangleq \prod_{t=1}^T \mathbb{P}(\mathrm{d}u_t|x^t, u^{t-1})$. In this setting a fundamental question in accordance with the performance of the control design is as follows.

Question 1 (Transmission rate)**.** *Suppose that a feedback control architecture in Figure 3 is required to meet the cost constraint $J_\gamma \leq D$. What is the fundamental lower bound of the total number of bits $\sum_{t=1}^T r_t$ that must be transmitted?*

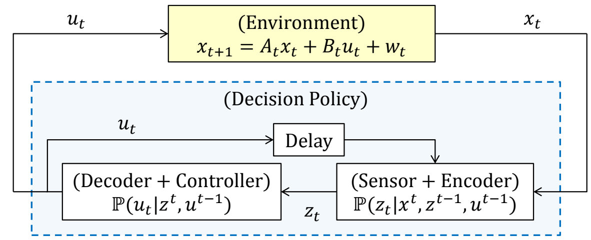

Fig. 3. An example of the internal architecture of the decision policy in Figure 1.

In order to address Question 1, we continue with a lemma, describing a "data-processing inequality" for directed information. The lemma is expressed in a rather general setting in which the process $\mathbf{z}_t$ is not necessarily a binary value, i.e., $\mathbf{z}_t \in \mathbb{R}^{r_t}$.

Lemma 1 (Feedback Data-Processing Inequality)**.** *Consider a control system (1) with a decision policy $\prod_{t=1}^T \mathbb{P}(\mathrm{d}u_t|x^t, u^{t-1})$. Assume that $\mathbb{P}(\mathrm{d}u_t|x^t, u^{t-1})$ can be realized as a composition of stochastic kernels $\mathbb{P}(\mathrm{d}z_t|x^t, z^{t-1}, u^{t-1})$ and $\mathbb{P}(\mathrm{d}u_t|z^t, u^{t-1})$, where $\mathbf{z}_t$ is an $\mathbb{R}^{r_t}$-valued random variable. Then, we have an inequality*

$$I(\mathbf{x}^T \to \mathbf{u}^T) \leq I(\mathbf{x}^T \to \mathbf{z}^T \| \mathbf{u}_+^{T-1}),$$

where the right hand side is a short-hand notation for $\sum_{t=1}^T I(\mathbf{x}^T; \mathbf{z}_t|\mathbf{z}^{t-1}, \mathbf{u}^{t-1})$, written with Kramer's causal conditioning [18].

Proof: Let us highlight that the version of the directed data processing inequality in [38] is not equivalent to Lemma 1, as the source $\mathbf{x}_t$ in Lemma 1 is affected by feedback. See Appendix for the detailed proof.

Note that when the stochastic process $\mathbf{z}_{t=1}^T$ takes binary values, we then have the standard inequality $H(\mathbf{z}_t) \leq \log_2 |\mathbf{z}_t| = r_t$, where $H(\cdot)$ is the entropy function. Using the assertion of Lemma 1 in the first place we have the following chain of inequalities:

$$\begin{aligned}
& I(\mathbf{x}^T \to \mathbf{u}^T) && \text{(5a)}\\
\leq & \sum_{t=1}^T I(\mathbf{x}^t; \mathbf{z}_t|\mathbf{z}^{t-1}, \mathbf{u}^{t-1}) && \text{(5b)}\\
= & \sum_{t=1}^T \left(H(\mathbf{z}_t|\mathbf{z}^{t-1}, \mathbf{u}^{t-1}) - H(\mathbf{z}_t|\mathbf{x}^t, \mathbf{z}^{t-1}, \mathbf{u}^{t-1}) \right) && \text{(5c)}\\
\leq & \sum_{t=1}^T H(\mathbf{z}_t|\mathbf{z}^{t-1}, \mathbf{u}^{t-1}) && \text{(5d)}\\
\leq & \sum_{t=1}^T H(\mathbf{z}_t) && \text{(5e)}\\
\leq & \sum_{t=1}^T r_t. && \text{(5f)}
\end{aligned}$$

Note that Lemma 1 is used in the first step. The inequalities (5) show that the directed information indeed suggests a lower bound for the minimum number of required bits to be transmitted in order to ensure the desired level of performance cost. By contrast, the standard mutual information $I(\mathbf{x}^T;\mathbf{u}^T)$ in general fails to provide an insight to Question 1. In a similar context, an observation on the relationship between directed information and the operational

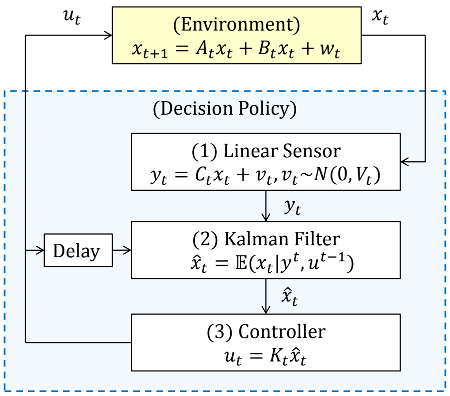

Fig. 4. Channel-filter-controller separation architecture for "information-frugal" LQG control.

source-coding rate is also discussed in [32, Theorem 4.1]; see references therein for further details.

Despite the above discussion, it should be noted that there is an important caveat concerning the achievability of the lower bound offered by the directed information in (5). That is, the bound is generally not operationally achievable even asymptotically (i.e., $T \to \infty$). This is mainly due to the fact that the standard rate-distortion function requires to consider arbitrarily large block-lengths [7, Theorem 10.2.1], which leads to arbitrarily large delays in our setting, and thus is not acceptable. It is, however, shown that the conservatisms of this lower bound (5) is bounded by a small constant [32]. In this light, we consider the problem (4) as the main focus of this paper rather than Question 1.

III. Main Result

This section includes the main message of the paper, presenting an optimal solution to the problem (4) comprising three data-processing components. Figure 4 visually illustrates these components which will be detailed in the rest of this section.

The control architecture consists of three components:

(i) A linear sensor mechanism $\mathbf{y}_t = C_t \mathbf{x}_t + \mathbf{v}_t$ with mutually independent additive Gaussian noise $\mathbf{v}_t \sim \mathcal{N}(0, V_t)$ with $V_t \succ 0$;

(ii) The Kalman filter computing $\hat{\mathbf{x}}_t = \mathbb{E}(\mathbf{x}_t | \mathbf{y}^t, \mathbf{u}^{t-1})$;

(iii) And the certainty equivalence controller $\mathbf{u}_t = K_t \hat{\mathbf{x}}_t$.

The parameters of these steps (i.e., matrices C_t and noise process $\mathbf{v}_t$ in component (i), the filter in (ii), and the controller matrix gain K_t in (iii)) can be synthesized in a sequential tractable computational procedure. That is, the preceding step provides input for the following one. The procedure involves three steps, starting with controller design in (iii), followed by the linear sensor design in (i), and finally the Kalman filter design in (ii):

- **Step 1.** (Controller design) Compute a backward Riccati recursion.

$$S_t = \begin{cases} Q_t & \text{if } t = T \\ Q_t + \Phi_{t+1} & \text{if } t = 1, \cdots, T-1 \end{cases} \tag{6a}$$

$$\Phi_t = A_t^\top (S_t - S_t B_t (B_t^\top S_t B_t + R_t)^{-1} B_t^\top S_t) A_t \tag{6b}$$

$$K_t = -(B_t^\top S_t B_t + R_t)^{-1} B_t^\top S_t A_t \tag{6c}$$

$$\Theta_t = K_t^\top (B_t^\top S_t B_t + R_t) K_t \tag{6d}$$

The optimal feedback control gains $\{K_t\}_{t=1}^T$ are obtained in (6c), while the positive semidefinite matrices Θ_t are computed as an input to the next step.

- **Step 2.** (Linear sensor design) Solve the max-det problem

$$\min_{\{P_{t|t}, \Pi_t\}_{t=1}^T} \quad \frac{1}{2} \sum_{t=1}^T \log\det \Pi_t^{-1} + c_1 \tag{7a}$$

$$\text{s.t.} \quad \sum_{t=1}^T \text{Tr}(\Theta_t P_{t|t}) + c_2 \le D \tag{7b}$$

$$\Pi_t \succ 0, \tag{7c}$$

$$P_{1|1} \preceq P_{1|0}, P_{T|T} = \Pi_T, \tag{7d}$$

$$P_{t+1|t+1} \preceq A_t P_{t|t} A_t^\top + W_t, \tag{7e}$$

$$\begin{bmatrix} P_{t|t} - \Pi_t & P_{t|t} A_t^\top \\ A_t P_{t|t} & A_t P_{t|t} A_t^\top + W_t \end{bmatrix} \succeq 0. \tag{7f}$$

The constraint (7c) is defined for every $t = 1, \cdots, T$, while (7e) and (7f) are for every $t = 1, \cdots, T-1$. The following computation provides the input for the last step. Set the constants c_1 and c_2 as

$$c_1 = \tfrac{1}{2} \log\det P_{1|0} + \tfrac{1}{2} \sum_{t=1}^T \log\det W_t$$

$$c_2 = \text{Tr}(N_1 P_{1|0}) + \sum_{t=1}^T \text{Tr}(W_t S_t).$$

Let $r_t = \text{rank}(P_{t|t}^{-1} - P_{t|t-1}^{-1})$ for $t = 1, \cdots, T$, where

$$P_{t|t-1} \triangleq A_{t-1} P_{t-1|t-1} A_{t-1}^\top + W_{t-1}, \quad t = 2, \cdots, T.$$

Apply the singular value decomposition to find matrices $C_t \in \mathbb{R}^{r_t \times n_t}$ and $V_t \in \mathbb{S}_{++}^{r_t}$ such that the matrix-valued signal-to-noise ratio (SNR) satisfies

$$\mathsf{SNR}_t \triangleq C_t^\top V_t^{-1} C_t = P_{t|t}^{-1} - P_{t|t-1}^{-1} \tag{8}$$

for $t = 1, \cdots, T$. In case of $r_t = 0$, C_t and V_t are considered to be null (zero dimensional) matrices.

- **Step 3.** (Filter design) Determine the Kalman gains by

$$L_t = P_{t|t-1} C_t^\top (C_t P_{t|t-1} C_t^\top + V_t)^{-1}. \tag{9}$$

Construct the Kalman filter by

$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1} + L_t (\mathbf{y}_t - C_t \hat{\mathbf{x}}_{t|t-1}) \tag{10a}$$

$$\hat{\mathbf{x}}_{t+1|t} = A_t \hat{\mathbf{x}}_t + B_t \mathbf{u}_t. \tag{10b}$$

If $r_t = 0$, L_t is a null matrix and (10a) is simply replaced by $\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1}$.

The following theorem formally explains the output of the above constructive procedure and bridges the result to the problem (4).

Theorem 1 (Information-Fugal LQG Controller)**.** *An optimal policy for the problem* (4) *exists if and only if the max-det problem* (7) *is feasible, and the optimal value of the program* (4) *coincides with the optimal value of* (7)*. If the optimal* *value of* (4) *is finite, an optimal policy can be realized by an interconnection of a linear sensor, Kalman filter, and a certainty equivalence controller as shown in Figure 4. Moreover, each of these components can be constructed by an SDP-based algorithm summarized in Steps 1-3.*

Remark 1 (Three-stage Separation Principle)**.** *The assertion of Theorem 1 is indeed an integration of the previously known filter-controller separation principle in the LQG control theory [41] and sensor-filter separation principle in the Gaussian sequential rate-distortion theory [38] (cf. Figure 2). We also emphasize that our separation principles enjoys a tractable computational characterization using an SDP algorithm.*

In the literature of classical LQG control theory, a linear sensor mechanism $\mathbf{y}_t = C_t\mathbf{x}_t + \mathbf{v}_t$ is traditionally considered to be a part of the given model. In this view, it is well-known from the filter-controller separation principle that the optimal LQG controller is a composition of the Kalman filter and the certainty equivalence controller [41]. On the other hand, for (uncontrolled) linear dynamical systems (i.e., when the controller gain K_t is fixed) the design problem is translated into the minimization of the directed information $I(\mathbf{x}^T \to \hat{\mathbf{x}}^T)$ over the reconstruction policies $\prod_{t=1}^{T} \mathbb{P}(\mathrm{d}\hat{x}_t|x^t, \hat{x}^{t-1})$. This problem turns out to be the Gaussian sequential rate-distortion problem [37], where a sensor-filter separation principle is known [38] (see also [35]). However, to the best of our knowledge, the three-stage sensor-filter-controller separation principle in Theorem 1 is derived for the first time in this paper for a control problem of the form (4).

IV. Derivation of Main Result

In this section we sketch the main ideas to establish the assertion of Theorem 1, and refer the interested reader to [36] for further details.

Let us define subsets $\Gamma_1 \subset \Gamma_2 \subset \Gamma$ of the policy space as follows.

- Subset $\Gamma_1 \subset \Gamma$ A policy $\gamma \in \Gamma_1$ is a sequence of stochastic kernels $\gamma = \prod_{t=1}^{T} \mathbb{P}(\mathrm{d}u_t|x^t, u^{t-1})$ where $\mathbb{P}(\mathrm{d}u_t|x^t, u^{t-1})$ can be written as a composition of the following terms:
  - (i) A stochastic kernel defined as $\mathbb{P}(\mathrm{d}y_t|x_t) = \mathcal{N}(C_t x_t, V_t)$ with some nonnegative integer r_t, a matrix $C_t \in \mathbb{R}^{r_t \times n_t}$ and $V_t \succ 0$. This kernel can be simply realized through a linear sensing mechanism with mutually independent additive Gaussian noise

$$\mathbf{y}_t = C_t\mathbf{x}_t + \mathbf{v}_t, \quad \mathbf{v}_t \sim \mathcal{N}(0, V_t); \tag{11}$$

  - (ii) A linear map $\mathbf{u}_t = l_t(\mathbf{y}^t)$.
- Subset $\Gamma_2 \subset \Gamma$ A policy $\gamma \in \Gamma_2$ is a sequence of stochastic kernels $\gamma = \prod_{t=1}^{T} \mathbb{P}(\mathrm{d}u_t|x^t, u^{t-1})$ where $\mathbb{P}(\mathrm{d}u_t|x^t, u^{t-1}) = \mathcal{N}(M_t\mathbf{x}_t + N_t\mathbf{u}^{t-1}, G_t)$ with some matrices M_t, N_t, and $G_t \succeq 0$. This kernel can be realized through

$$\mathbf{u}_t = M_t\mathbf{x}_t + N_t\mathbf{u}^{t-1} + \mathbf{g}_t, \quad \mathbf{g}_t \sim \mathcal{N}(0, G_t). \tag{12}$$

The outline of the proof of Theorem 1 can be described through a chain of inequalities as follows:

$$\inf_{\gamma\in\Gamma: J_\gamma \le D} I(\mathbf{x}^T \to \mathbf{u}^T) \tag{13a}$$
$$\ge \inf_{\gamma\in\Gamma: J_\gamma \le D} \sum_{t=1}^{T} I(\mathbf{x}_t; \mathbf{u}_t|\mathbf{u}^{t-1}) \tag{13b}$$
$$\ge \inf_{\gamma\in\Gamma_2: J_\gamma \le D} \sum_{t=1}^{T} I(\mathbf{x}_t; \mathbf{u}_t|\mathbf{u}^{t-1}) \tag{13c}$$
$$\ge \inf_{\gamma\in\Gamma_1: J_\gamma \le D} \sum_{t=1}^{T} I(\mathbf{x}_t; \mathbf{y}_t|\mathbf{y}^{t-1}) \tag{13d}$$
$$\ge \inf_{\gamma\in\Gamma_1: J_\gamma \le D} I(\mathbf{x}^T \to \mathbf{u}^T) \tag{13e}$$

Notice that since $\Gamma_1 \subset \Gamma$, we clearly have (13a) $\le$ (13e). Therefore, showing the above inequalities proves that all quantities in (13) are equal. This is indeed an insightful observation indicating that the restriction of the class of optimal policies to the subset Γ_1 does not deteriorate the performance.

The first inequality (13b) follows directly from the definition of the directed information. As such, the proof of inequality (13e) reveals that an optimal solution to (13d), if exists, is also an optimal solution to (13e). This implies that it suffices to find an optimal solution to a simplified problem (13d). It is remarkable that the class of control policies Γ_1 by construction enjoys the separation structure as detailed in (i) and (ii). We further show that, by invoking standard LQG optimal control theory, the optimal linear map $\mathbf{u}_t = l_t(\mathbf{y}^t)$ for (13d) can be written as $\mathbf{u}_t = K_t\mathbb{E}(\mathbf{x}_t|\mathbf{y}^t, \mathbf{u}^{t-1})$ where $\mathbb{E}(\mathbf{x}_t|\mathbf{y}^t, \mathbf{u}^{t-1})$ is computed by the Kalman filter. This observation establishes the sensor-filter-controller separation principle. With this structural understanding, we show that problem (13d) can be reformulated as an optimization problem in terms of $\mathsf{SNR}_t \triangleq C_t^\top V_t^{-1} C_t$, which is further converted to an SDP problem.

V. Conclusion and Future Directions

This paper considered an optimal control problem in which the directed information from the state variables to the control actions is minimized subject to the requirement that the control policy achieves a desired level of LQG control performance. It was shown that an optimal control policy admits a novel three-stage separation structure comprising (i) an additive Gaussian channel, (ii) Kalman filter, and (iii) certainty equivalence controller. We also proposed a tractable numerical algorithm to synthesize the optimal policy with the three-stage architecture.

The problem setting in this paper is concerned with a finite horizon performance cost for a linear time-varying dynamical system. An interesting future direction is to investigate the implication of the main result of the paper, Theorem 1, in other special cases including time invariant dynamics and/or an infinite horizon performance cost. Besides, throughout this study we assume that we have access to full state of the system, potentially subject to some measurement noise. A natural extension is when we have only partial observation of the state of the system, i.e., the decision-maker is only allowed to measure a process $\mathbf{y}_t \triangleq h(\mathbf{x}_t)$ for a given output function h. The key idea for this extension is the *innovations approach* [16], which is a standard technique since [4]. Other relevant discussions can also be found in [34] in the context of zero-delay rate-distortion theory for partially observable Gauss-Markov processes.

Appendix

In this appendix we show the data-processing inequality for directed information, and in particular provide the proof of Lemma 1. First, let us recall that Kramer's *causally conditioned directed information* [18] is defined by

$$I(\mathbf{x}^T \to \mathbf{y}^T \| \mathbf{z}^T) \triangleq \sum_{t=1}^{T} I(\mathbf{x}^T; \mathbf{y}_t | \mathbf{y}^{t-1}, \mathbf{z}^t).$$

Notice that the following chain of equalities hold for every $t = 1, \cdots, T$.

$$\begin{aligned} & I(\mathbf{x}^t; \mathbf{z}_t | \mathbf{z}^{t-1}, \mathbf{u}^{t-1}) - I(\mathbf{x}^t; \mathbf{u}_t | \mathbf{u}^{t-1}) \\ =& I(\mathbf{x}^t; \mathbf{z}_t, \mathbf{u}_t | \mathbf{z}^{t-1}, \mathbf{u}^{t-1}) - I(\mathbf{x}^t; \mathbf{u}_t | \mathbf{u}^{t-1}) && \text{(14a)} \\ =& I(\mathbf{x}^t; \mathbf{z}^t | \mathbf{u}^t) - I(\mathbf{x}^t; \mathbf{z}^{t-1} | \mathbf{u}^{t-1}) && \text{(14b)} \\ =& I(\mathbf{x}^t; \mathbf{z}^t | \mathbf{u}^t) - I(\mathbf{x}^{t-1}; \mathbf{z}^{t-1} | \mathbf{u}^{t-1}) \\ & \qquad - I(\mathbf{x}^{t-1}; \mathbf{z}^{t-1} | \mathbf{x}^{t-1}, \mathbf{u}^{t-1}) && \text{(14c)} \\ =& I(\mathbf{x}^t; \mathbf{z}^t | \mathbf{u}^t) - I(\mathbf{x}^{t-1}; \mathbf{z}^{t-1} | \mathbf{u}^{t-1}). && \text{(14d)} \end{aligned}$$

When $t = 1$, the above identity is understood to mean $I(\mathbf{x}_1; \mathbf{z}_1) - I(\mathbf{x}_1; \mathbf{u}_1) = I(\mathbf{x}_1; \mathbf{z}_1 | \mathbf{u}_1)$ which clearly holds as $\mathbf{x}_1$–$\mathbf{z}_1$–$\mathbf{u}_1$ form a Markov chain. Equation (14a) holds because $I(\mathbf{x}^t; \mathbf{z}_t, \mathbf{u}_t | \mathbf{z}^{t-1}, \mathbf{u}^{t-1}) = I(\mathbf{x}^t; \mathbf{z}_t | \mathbf{z}^{t-1}, \mathbf{u}^{t-1}) + I(\mathbf{x}^t; \mathbf{u}_t | \mathbf{z}^t, \mathbf{u}^{t-1})$ and the second term is zero since $\mathbf{x}^t$–$(\mathbf{z}^t, \mathbf{u}^{t-1})$–$\mathbf{u}_t$ form a Markov chain. Equation (14b) is obtained by applying the chain rule for mutual information in two different ways:

$$\begin{aligned} & I(\mathbf{x}^t; \mathbf{z}^t, \mathbf{u}_t | \mathbf{u}^{t-1}) \\ &= I(\mathbf{x}^t; \mathbf{z}^{t-1} | \mathbf{u}^{t-1}) + I(\mathbf{x}^t; \mathbf{z}_t, \mathbf{u}_t | \mathbf{z}^{t-1}, \mathbf{u}^{t-1}) \\ &= I(\mathbf{x}^t; \mathbf{u}_t | \mathbf{u}^{t-1}) + I(\mathbf{x}^t; \mathbf{z}^t | \mathbf{u}^t). \end{aligned}$$

The chain rule is applied again in step (14c). Finally, (14d) follows as $\mathbf{z}^{t-1}$–$(\mathbf{x}^{t-1}, \mathbf{u}^{t-1})$–$\mathbf{x}_t$ form a Markov chain.

Now, the desired inequality can be verified by computing the right hand side minus the left hand side as

$$\begin{aligned} & \sum_{t=1}^{T} \left[I(\mathbf{x}^t; \mathbf{z}_t | \mathbf{z}^{t-1}, \mathbf{u}^{t-1}) - I(\mathbf{x}^t; \mathbf{u}_t | \mathbf{u}^{t-1}) \right] \\ =& \sum_{t=1}^{T} \left[I(\mathbf{x}^t; \mathbf{z}^t | \mathbf{u}^t) - I(\mathbf{x}^{t-1}; \mathbf{z}^{t-1} | \mathbf{u}^{t-1}) \right] && \text{(15a)} \\ =& I(\mathbf{x}^T; \mathbf{z}^T | \mathbf{u}^T) \geq 0. && \text{(15b)} \end{aligned}$$

In step (15a), the identity (14) is used. The telescoping sum (15a) cancels all but the final term (15b); this concludes Lemma 1 assertion.

References

[1] J. Baillieul and P. J. Antsaklis, *Control and communication challenges in networked real-time systems*, Proceedings of the IEEE, 95 (2007), pp. 9–28.

[2] L. Bao, M. Skoglund, and K. H. Johansson, *Iterative encoder-controller design for feedback control over noisy channels*, IEEE Transactions on Automatic Control, 56 (2011), pp. 265–278.

[3] D. P. Bertsekas and S. E. Shreve, *Stochastic optimal control: The discrete time case*, vol. 139, Academic Press New York, 1978.

[4] H. W. Bode and C. E. Shannon, *A simplified derivation of linear least square smoothing and prediction theory*, Proceedings of the IRE, 38 (1950), pp. 417–425.

[5] J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg, *Feedback stabilization over signal-to-noise ratio constrained channels*, IEEE Transactions on Automatic Control, 52 (2007), pp. 1391–1403.

[6] R. W. Brockett and D. Liberzon, *Quantized feedback stabilization of linear systems*, IEEE Transactions on Automatic Control, 45 (2000), pp. 1279–1289.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley-Interscience, New York, NY, USA, 1991.

[8] D. F. Delchamps, *Stabilizing a linear system with quantized state feedback*, IEEE Transactions on Automatic Control, 35 (1990), pp. 916–924.

[9] C. Dwork, *Differential privacy*, in Automata, languages and programming, Springer, 2006, pp. 1–12.

[10] N. Elia, *When bode meets shannon: Control-oriented feedback communication schemes*, IEEE Transactions on Automatic Control, 49 (2004), pp. 1477–1488.

[11] ———, *Remote stabilization over fading channels*, Systems & Control Letters, 54 (2005), pp. 237–249.

[12] N. Elia and S. K. Mitter, *Stabilization of linear systems with limited information*, IEEE Transactions on Automatic Control, 46 (2001), pp. 1384–1400.

[13] V. Gupta, B. Hassibi, and R. M. Murray, *Optimal lqg control across packet-dropping links*, Systems & Control Letters, 56 (2007), pp. 439–446.

[14] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, *A survey of recent results in networked control systems*, PROCEEDINGS-IEEE, 95 (2007), p. 138.

[15] H. Ishii and B. A. Francis, *Quadratic stabilization of sampled-data systems with quantization*, Automatica, 39 (2003), pp. 1793–1800.

[16] T. Kailath, *An innovations approach to least-squares estimation–Part I: Linear filtering in additive white noise*, IEEE Transactions on Automatic Control, 13 (1968), pp. 646–655.

[17] Y.-H. Kim, *Feedback capacity of stationary gaussian channels*, IEEE Transactions on Information Theory, 56 (2010), pp. 57–85.

[18] G. Kramer, *Capacity results for the discrete memoryless network*, IEEE Transactions on Information Theory, 49 (2003), pp. 4–21.

[19] S. B. Laughlin, *Energy as a constraint on the coding and processing of sensory information*, Current opinion in neurobiology, 11 (2001), pp. 475–480.

[20] S. M. LaValle, *Sensing and filtering: A fresh perspective based on preimages and information spaces*, Now Publishers, 2012.

[21] H. Marko, *The bidirectional communication theory–a generalization of information theory*, IEEE Transactions on Communications, 21 (1973), pp. 1345–1351.

[22] N. C. Martins and M. A. Dahleh, *Feedback control in the presence of noisy channels: "Bode-like" fundamental limitations of performance*, IEEE Transactions on Automatic Control, 53 (2008), pp. 1604–1615.

[23] J. Massey, *Causality, feedback and directed information*, in International Symposium on Information Theory and its Applications (ISITA-90), 1990, pp. 303–305.

[24] J. L. Massey and P. C. Massey, *Conservation of mutual and directed information*, in International Symposium on Information Theory, IEEE, 2005, pp. 157–158.

[25] A. S. Matveev and A. V. Savkin, *Estimation and control over communication networks*, Springer Science & Business Media, 2009.

[26] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, *Feedback control under data rate constraints: An overview*, Proceedings of the IEEE, 95 (2007), pp. 108–137.

[27] P. A. Ortega and D. A. Braun, *Thermodynamics as a theory of decision-making with information-processing costs*, in Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 469, 2013, p. 20120683.

[28] I. S. Reed, *Information theory and privacy in data banks*, in Proceedings of the June 4-8, 1973, national computer conference and exposition, 1973, pp. 581–587.

[29] A. Sahai and S. Mitter, *The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link Part I: Scalar systems*, IEEE Transactions on Information Theory, 52 (2006), pp. 3369–3395.

[30] L. SANKAR, S. R. RAJAGOPALAN, AND H. V. POOR, *Utility-privacy tradeoffs in databases: An information-theoretic approach*, IEEE Transactions on Information Forensics and Security, 8 (2013), pp. 838–852.

[31] E. SHAFIEEPOORFARD AND M. RAGINSKY, *Rational inattention in scalar LQG control*, in 52nd IEEE Conference on Decision and Control, 2013, pp. 5733–5739.

[32] E. SILVA, M. S. DERPICH, J. OSTERGAARD, ET AL., *A framework for control system design subject to average data-rate constraints*, IEEE Transactions on Automatic Control, 56 (2011), pp. 1886–1899.

[33] C. A. SIMS, *Implications of rational inattention*, Journal of monetary Economics, 50 (2003), pp. 665–690.

[34] T. TANAKA, *Zero-delay rate-distortion optimization for partially observable Gauss-Markov processes*, (2015). 52nd IEEE Conference on Decision and Control.

[35] T. TANAKA, K.-K. K. KIM, P. A. PARRILO, AND S. K. MITTER, *Semidefinite programming approach to Gaussian sequential rate-distortion trade-offs*, (2014). Preprint available at http://arxiv.org/abs/1411.7632.

[36] T. TANAKA, P. MOHAJERIN ESFAHANI, AND S. MITTER, *LQG control with minimal information: Three-stage separation principle and SDP-based solution synthesis*, (2015). Available at http://control.ee.ethz.ch/~peymanm/drafts/LQG.pdf.

[37] T. TANAKA AND H. SANDBERG, *SDP-based joint sensor and controller design for information-regularized optimal LQG control*, arXiv:1503.01848, (2015). 52nd IEEE Conference on Decision and Control (To appear). Preprint available at http://arxiv.org/abs/1503.01848.

[38] S. TATIKONDA, *Control under communication constraints*, PhD thesis, Massachusetts Institute of Technology, (2000).

[39] S. TATIKONDA AND S. MITTER, *The capacity of channels with feedback*, IEEE Transactions on Information Theory, 55 (2009), pp. 323–349.

[40] S. TATIKONDA, A. SAHAI, AND S. MITTER, *Stochastic linear control over a communication channel*, IEEE Transactions on Automatic Control, 49 (2004), pp. 1549–1561.

[41] H. S. WITSENHAUSEN, *Separation of estimation and control for discrete time systems*, Proceedings of the IEEE, 59 (1971), pp. 1557–1566.

[42] S. YÜKSEL AND T. BAŞAR, *Stochastic networked control systems*, vol. 10 of Systems & Control Foundations & Applications, Springer, New York, NY, 2013.