LQG Control with Minimum Directed Information: Semidefinite Programming Approach

Takashi Tanaka[1] Peyman Mohajerin Esfahani[2] Sanjoy K. Mitter[3]

***Abstract*—We consider a discrete-time Linear-Quadratic-Gaussian (LQG) control problem in which Massey's directed information from the observed output of the plant to the control input is minimized while required control performance is attainable. This problem arises in several different contexts, including joint encoder and controller design for data-rate minimization in networked control systems. We show that the optimal control law is a Linear-Gaussian randomized policy. We also identify the state space realization of the optimal policy, which can be synthesized by an efficient algorithm based on semidefinite programming. Our structural result indicates that the filter-controller separation principle from the LQG control theory, and the sensor-filter separation principle from the zero-delay rate-distortion theory for Gauss-Markov sources hold simultaneously in the considered problem. A connection to the data-rate theorem for mean-square stability by Nair & Evans is also established.**

***Index Terms*—Control over communications; Kalman filtering; LMIs; Stochastic optimal control; Communication Networks**

I. Introduction

There is a fundamental trade-off between the best achievable control performance and the data-rate at which plant information is fed back to the controller. Studies of such a trade-off hinge upon analytical tools developed at the interface between traditional feedback control theory and Shannon's information theory. Although the interface field has been significantly expanded by the surged research activities on *networked control systems (NCS)* over the last two decades [1]–[5], many important questions concerning the rate-performance trade-off studies are yet to be answered.

A central research topic in the NCS literature has been the stabilizability of a linear dynamical system using a rate-constrained feedback [6]–[9]. The critical data-rate below which stability cannot be attained by any feedback law has been extensively studied in various NCS setups. As pointed out by [10], many results including [6]–[9] share the same conclusion that this critical data-rate is characterized by an intrinsic property of the open-loop system known as topological entropy, which is determined by the unstable open-loop poles. This result holds irrespective of different definitions of the "data-rate" considered in these papers. For instance, in [9] the data-rate is defined as the log-cardinality of channel alphabet, while in [8], it is the frequency of the use of noiseless binary channel.

[1]TT is with the Department of Aerospace Engineering and Engineering Mechanics at the University of Texas at Austin, USA. ttanaka@utexas.edu; [2]PME is with the Delft Center for Systems and Control at the Delft University of Technology, Netherlands. P.MohajerinEsfahani@tudelft.nl; [3] SM is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, USA. mitter@mit.edu.

As a natural next step, the rate-performance trade-offs are of great interest from both theoretical and practical perspectives. The trade-off between Linear-Quadratic-Gaussian (LQG) performance and the required data-rate has attracted attention in the literature [11]–[24]. Generalized interpretations of the classical Bode's integral also provide fundamental performance limitations of closed-loop systems in the information-theoretic terms [25]–[28]. However, the rate-performance trade-off analysis introduces additional challenges that were not present through the lens of the stability analysis. First, it is largely unknown whether different definitions of the data-rate considered in the literature listed above lead to different conclusions. This issue is less visible in the stability analysis, since the critical data-rate for stability turns out to be invariant across several different definitions of the data-rate [6]–[9]. Second, for many operationally meaningful definitions of the data-rate considered in the literature, computation of the rate-performance trade-off function involves intractable optimization problems (e.g., dynamic programming [21] and iterative algorithm [18]), and trade-off achieving controller/encoder policies are difficult to obtain. This is not only inconvenient in practice, but also makes theoretical analyses difficult.

In this paper, we study the information-theoretic requirements for LQG control using the notion of *directed information* [29]–[31]. In particular, we define the rate-performance trade-off function as the minimal directed information from the observed output of the plant to the control input, optimized over the space of causal decision policies that achieve the desired level of LQG control performance. Among many possible definitions of the "data-rate" as mentioned earlier, we focus on directed information for the following reasons.

First, directed information (or related quantity known as *transfer entropy*) is a widely used *causality measure* in science and engineering [32]–[34]. Applications include communication theory (e.g., the analysis of channels with feedback), portfolio theory, neuroscience, social science, macroeconomics, statistical mechanics, and potentially more. Since it is natural to measure the "data-rate" in networked control systems by a causality measure from the observation to action, directed information is a natural option.

Second, it is recently reported by Silva et al. [22]–[24] that directed information has an important operational meaning in a practical NCS setup. Starting from an LQG control problem over a noiseless binary channel with prefix-free codewords, they show that the directed information obtained by solving the aforementioned optimization problem provides a tight lower bound for the minimum data-rate (defined operationally) required to achieve the desired level of control performance.

A. Contributions of this paper

The central question in this paper is the characterization of the most "data-frugal" LQG controller that minimizes directed information of interest among all decision policies achieving a given LQG control performance. In this paper, we make the following contributions.

(i) In a general setting including MIMO, time-varying, and partially observable plants, we identify the structure of an optimal decision policy in a state space model.

(ii) Based on the above structural result, we further develop a tractable optimization-based framework to synthesize the optimal decision policy.

(iii) In the stationary setting with MIMO plants, we show how our proposed computational framework, as a special case, recovers the existing data-rate theorem for mean-square stability.

Concerning (i), we start with general time-varying, MIMO, and fully observable plants. We emphasize that the optimal decision policy in this context involves two important tasks: (1) the *sensing* task, indicating which state information of the plant should be dynamically measured with what precision, and (2) the *control* task, synthesizing an appropriate control action given available sensing information. To this end, we first show that the optimal policy that minimizes directed information from the state to the control sequences under the LQG control performance constraint is linear. In this vein, we illustrate that the optimal policy can be realized by a three-stage architecture comprising linear sensor with additive Gaussian noise, Kalman filter, and certainty equivalence controller (Theorem 1). We then show how this result can be extended to partially observed plants (Theorem 3).

Regarding (ii), we provide a semidefinite programming (SDP) framework characterizing the optimal policy proposed in step (i) (Sections IV and VII). As a result, we obtain a computationally accessible form of the considered rate-performance trade-off functions.

Finally, as highlighted in (iii), we analyze the horizontal asymptote of the considered rate-performance trade-off function for MIMO time-invariant plants (Theorem 2), which coincides with the critical data-rate identified by Nair and Evans [9] (Corollary 1).

B. Organization of this paper

The rest of this paper is organized as follows. After some notational remarks, the problem considered in this paper is formally introduced in Section II, and its operational interpretation is provided in Section III. Main results are summarized in Section IV, where connections to the existing results are also explained in detail. Section V contains a simple numerical example, and the derivation of the main results is presented in Section VI. The results are extended to partially observable plants in Section VII. We conclude in Section VIII.

C. Notational remarks

Throughout this paper, random variables are denoted by lower case bold symbols such as $\mathbf{x}$. Calligraphic symbols such as $\mathcal{X}$ are used to denote sets, and $x \in \mathcal{X}$ is an element. We denote by x^t a sequence $x_1, x_2, ..., x_t$, and $\mathbf{x}^t$ and $\mathcal{X}^t$ are understood similarly. All random variables in this paper are Euclidean valued, and is measurable with respect to the usual topology. A probability distribution of $\mathbf{x}$ is demoted by $\mathbb{P}_{\mathbf{x}}$. A Gaussian distribution with mean μ and covariance Σ is denoted by $\mathcal{N}(\mu, \Sigma)$. The relative entropy of $\mathbb{Q}$ from $\mathbb{P}$ is a non-negative quantity defined by

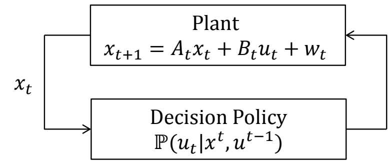

Fig. 1. LQG control of fully observable plant with minimum directed information.

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

where $\mathbb{P} \ll \mathbb{Q}$ means that $\mathbb{P}$ is absolutely continuous with respect to $\mathbb{Q}$, and $\frac{\mathrm{d}\mathbb{P}(x)}{\mathrm{d}\mathbb{Q}(x)}$ denotes the Radon-Nikodym derivative. The mutual information between $\mathbf{x}$ and $\mathbf{y}$ is defined by $I(\mathbf{x};\mathbf{y}) \triangleq D(\mathbb{P}_{\mathbf{x},\mathbf{y}} \| \mathbb{P}_{\mathbf{x}} \otimes \mathbb{P}_{\mathbf{y}})$, where $\mathbb{P}_{\mathbf{x},\mathbf{y}}$ and $\mathbb{P}_{\mathbf{x}} \otimes \mathbb{P}_{\mathbf{y}}$ are joint and product probability measures respectively. 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. Problem Formulation

Consider a linear time-varying stochastic plant

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

where $\mathbf{x}_t$ is an $\mathbb{R}^n$-valued state of the plant, and $\mathbf{u}_t$ is the control input. We assume that initial state $\mathbf{x}_1 \sim \mathcal{N}(0, P_{1|0}), P_{1|0} \succ 0$ and noise process $\mathbf{w}_t \sim \mathcal{N}(0, W_t)$, $W_t \succ 0, t = 1, ..., T$ are mutually independent.

The design objective is to synthesize a decision policy that "consumes" the least amount of information among all policies achieving the required LQG control performance (Figure 1). Specifically, let Γ be the space of decision policies, i.e., the space of sequences of Borel measurable stochastic kernels [35]

$$\mathbb{P}(u^T \| x^T) \triangleq \{\mathbb{P}(u_t | x^t, u^{t-1})\}_{t=1,...,T}.$$

A decision policy $\gamma \in \Gamma$ is evaluated by two criteria:

(i) the LQG control cost

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

(ii) and *directed information*

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

The right hand side of (2) and (3) are evaluated with respect to the joint probability measure induced by the state space model (1) and a decision policy γ. In what follows, we often write

(2) and (3) as J_γ and I_γ to indicate their dependency on γ. The main problem studied in this paper is formulated as

$$\mathsf{DI}_T(D) \triangleq \min_{\gamma\in\Gamma} \quad I_\gamma(\mathbf{x}^T \to \mathbf{u}^T) \tag{4a}$$
$$\text{s.t.} \quad J_\gamma(\mathbf{x}^{T+1}, \mathbf{u}^T) \leq D, \tag{4b}$$

where $D > 0$ is the desired LQG control performance.

Directed information (3) can be interpreted as the information flow from the state random variable $\mathbf{x}_t$ to the control random variable $\mathbf{u}_t$. The following equality called *conservation of information* [36] shows a connection between directed information and the standard mutual information:

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

Here, the sequence $\mathbf{u}_+^{T-1} = (0, \mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{T-1})$ denotes an index-shifted version of $\mathbf{u}^T$. Intuitively, this equality shows that the standard mutual information can be written as a sum of two directed information terms corresponding to feedback (through decision policy) and feedforward (through plant) information flows. Thus (4) is interpreted as the minimum information that must "flow" through the decision policy to achieve the LQG control performance D.

We also consider time-invariant and infinite-horizon LQG control problems. Consider a time-invariant plant

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + \mathbf{w}_t, \quad t\in\mathbb{N} \tag{5}$$

with $\mathbf{w}_t \sim \mathcal{N}(0, W)$, and assume $Q_t = Q$ and $R_t = R$ for $t \in \mathbb{N}$. We also assume (A, B) is stabilizable, (A, Q) is detectable, and $R \succ 0$. Let Γ be the space of Borel-measurable stochastic kernels $\mathbb{P}(u^\infty \| x^\infty)$. The problem of interest is

$$\mathsf{DI}(D) \triangleq \min_{\gamma\in\Gamma} \quad \limsup_{T\to\infty} \tfrac{1}{T} I_\gamma(\mathbf{x}^T \to \mathbf{u}^T) \tag{6a}$$
$$\text{s.t.} \quad \limsup_{T\to\infty} \tfrac{1}{T} J_\gamma(\mathbf{x}^{T+1}, \mathbf{u}^T) \leq D. \tag{6b}$$

More general problem formulations with partially observable plants will be discussed in Section VII.

III. Operational Meaning

In this section, we revisit a networked LQG control problem considered in [22]–[24]. Here we consider time-invariant MIMO plants while [22]–[24] focus on SISO plants. For simplicity, we consider fully observable plants only. Consider a feedback control system in Figure 2, where the state information is encoded by the "sensor + encoder" block and is transmitted to the controller over a noiseless binary channel. For each $t = 1, ..., T$, let $\mathcal{A}_t \subset \{0, 1, 00, 01, 10, 11, 000, \cdots\}$ be a set of uniquely decodable variable-length codewords [37, Ch.5]. Assume that codewords are generated by a causal policy

$$\mathbb{P}(a^\infty \| x^\infty) \triangleq \{\mathbb{P}(a_t|x^t, a^{t-1})\}_{t=1,2,\dots}.$$

The "decoder + controller" block interprets codewords and computes control input according to a causal policy

$$\mathbb{P}(u^\infty \| a^\infty) \triangleq \{\mathbb{P}(u_t|a^t, u^{t-1})\}_{t=1,2,\dots}.$$

The length of a codeword $\mathbf{a}_t \in \mathcal{A}_t$ is denoted by a random variable l_t. Let Γ' be the space of triplets

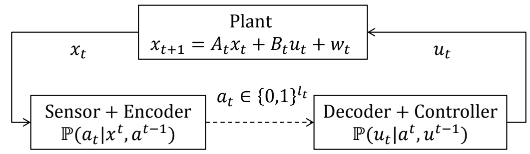

Fig. 2. LQG control over noiseless binary channel.

$\{\mathbb{P}(a^\infty \| x^\infty), \mathcal{A}^\infty, \mathbb{P}(u^\infty \| a^\infty)\}$. Introduce a quadratic control cost

$$J(\mathbf{x}^{T+1}, \mathbf{u}^T) \triangleq \sum\nolimits_{t=1}^{T} \mathbb{E}\left(\|\mathbf{x}_{t+1}\|_Q^2 + \|\mathbf{u}_t\|_R^2\right)$$

with $Q \succ 0$ and $R \succ 0$. We are interested in a design $\gamma' \in \Gamma'$ that minimizes data-rate among those attaining control cost smaller than D. Formally, the problem is formulated as

$$\mathsf{R}(D) \triangleq \min_{\gamma'\in\Gamma'} \quad \limsup_{T\to+\infty} \frac{1}{T}\sum\nolimits_{t=1}^{T} \mathbb{E}(l_t) \tag{7a}$$
$$\text{s.t.} \quad \limsup_{T\to+\infty} \frac{1}{T} J(\mathbf{x}^{T+1}, \mathbf{u}^T) \leq D. \tag{7b}$$

It is difficult to evaluate $\mathsf{R}(D)$ directly since (7) is a highly complex optimization problem. Nevertheless, Silva et al. [22] observed that $\mathsf{R}(D)$ is closely related to $\mathsf{DI}(D)$ defined by (6). The following result is due to [38].

$$\mathsf{DI}(D) \leq \mathsf{R}(D) < \mathsf{DI}(D) + \frac{r}{2}\log\frac{4\pi e}{12} + 1 \;\; \forall D > 0. \tag{8}$$

Here, r is an integer no greater than the state space dimension of the plant.[1] The following inequality plays an important role to prove (8).

Lemma 1: Consider a control system (1) with a decision policy $\gamma' \in \Gamma'$. Then, we have an inequality

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

where the right hand side is Kramer's notation [31] for causally conditioned directed information $\sum_{t=1}^T I(\mathbf{x}^t; \mathbf{a}_t | \mathbf{a}^{t-1}, \mathbf{u}^{t-1})$.

Proof: See Appendix A. ∎

Lemma 1 can be thought of as a generalization of the standard data-processing inequality. It is different from the directed data-processing inequality in [6, Lemma 4.8.1] since the source $\mathbf{x}_t$ is affected by feedback. See also [39] for relevant inequalities involving directed information.

Now, the first inequality in (8) can be directly verified as

$$I(\mathbf{x}^T \to \mathbf{u}^T) \tag{9a}$$
$$\leq \sum\nolimits_{t=1}^{T} I(\mathbf{x}^t; \mathbf{a}_t | \mathbf{a}^{t-1}, \mathbf{u}^{t-1}) \tag{9b}$$
$$= \sum\nolimits_{t=1}^{T} \left(H(\mathbf{a}_t|\mathbf{a}^{t-1}, \mathbf{u}^{t-1}) - H(\mathbf{a}_t|\mathbf{x}^t, \mathbf{a}^{t-1}, \mathbf{u}^{t-1})\right) \tag{9c}$$
$$\leq \sum\nolimits_{t=1}^{T} H(\mathbf{a}_t|\mathbf{a}^{t-1}, \mathbf{u}^{t-1}) \tag{9d}$$
$$\leq \sum\nolimits_{t=1}^{T} H(\mathbf{a}_t) \tag{9e}$$
$$\leq \sum\nolimits_{t=1}^{T} \mathbb{E}(l_t). \tag{9f}$$

[1] More precisely, r is the rank of the optimal signal-to-noise ratio matrix obtained by semidefinite programming, as will be clear in Section IV-B.

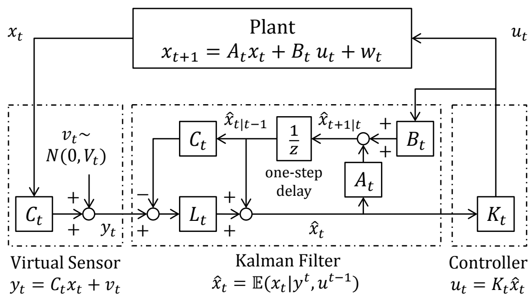

Fig. 3. Structure of optimal control policy for problem (4). Matrices C_t, V_t, L_t and K_t are determined by SDP-based algorithm in Section IV.

Lemma 1 is used in the first step. The last step follows from the fact that expected codeword length of of uniquely decodable codes is lower bounded by its entropy [37, Theorem 5.3.1].

Proving the second inequality in (8) requires a key technique proposed in [22] involving the construction of dithered uniform quantizer [40]. Detailed discussion is available in [38].

IV. Main Result

In this section we present the main results of this article. For the clarity of the presentation, this section is only devoted to a setting with full state measurements and shows how the main objective of control synthesis can be achieved by a three-step procedure. We shall later discuss in Section VII in regard to an extension to partial observable systems.

A. Time-varying plants

We show that the optimal solution to (4) can be realized by the following three data-processing components as shown in Figure 3.

1. A linear sensor mechanism

$$\mathbf{y}_t = C_t \mathbf{x}_t + \mathbf{v}_t, \quad \mathbf{v}_t \sim \mathcal{N}(0, V_t), \quad V_t \succ 0 \tag{10}$$

   where $\mathbf{v}_t, t = 1, ..., T$ are mutually independent.
2. The Kalman filter computing $\hat{\mathbf{x}}_t = \mathbb{E}(\mathbf{x}_t|\mathbf{y}^t, \mathbf{u}^{t-1})$.
3. The certainty equivalence controller $\mathbf{u}_t = K_t \hat{\mathbf{x}}_t$.

The role of the mechanism (10) is noteworthy. Recall that in the current problem setting in Figure 1, the state vector $\mathbf{x}_t$ is directly observable by the decision policy. The purpose of introducing an artificial mechanism (10) is to reduce data "consumed" by the decision policy while desired control performance is still attainable. Intuitively, the optimal mechanism (10) acquires just enough information from the state vector $\mathbf{x}_t$ for control purposes and discards less important information. Since the importance of information is a task-dependent notion, such a mechanism is designed jointly with other components in 2 and 3. The mechanism (10) may not be a physical sensor mechanism, but rather be a mere computational procedure. For this reason, we also call (10) a "virtual sensor." A virtual sensor can also be viewed as an instantaneous lossy data-compressor in the context of networked LQG control [22], [38]. As shown in [38], the knowledge of the optimal virtual sensor can be used to design a dithered uniform quantizer with desired performance.

We also claim that data-processing components in 1-3 can be synthesized by a tractable computational procedure based on SDP summarized below. The procedure is sequential, starting from controller design, followed by virtual sensor design and Kalman filter design.

- **Step 1** (Controller design) Determine feedback control gains K_t via the backward Riccati recursion:

$$S_t = \begin{cases} Q_t & \text{if } t = T \\ Q_t + \Phi_{t+1} & \text{if } t = 1, \cdots, T-1 \end{cases} \tag{11a}$$
$$\Phi_t = A_t^\top (S_t - S_t B_t (B_t^\top S_t B_t + R_t)^{-1} B_t^\top S_t) A_t \tag{11b}$$
$$K_t = -(B_t^\top S_t B_t + R_t)^{-1} B_t^\top S_t A_t \tag{11c}$$
$$\Theta_t = K_t^\top (B_t^\top S_t B_t + R_t) K_t. \tag{11d}$$

  Positive semidefinite matrices Θ_t will be used in Step 2.

- **Step 2** (Virtual sensor design) Let $\{P_{t|t}, \Pi_t\}_{t=1}^T$ be the optimal solution to a max-det problem:

$$\min_{\{P_{t|t}, \Pi_t\}_{t=1}^T} \quad \frac{1}{2} \sum_{t=1}^{T} \log\det \Pi_t^{-1} + c_1 \tag{12a}$$
$$\text{s.t.} \quad \sum_{t=1}^{T} \mathrm{Tr}(\Theta_t P_{t|t}) + c_2 \le D \tag{12b}$$
$$\Pi_t \succ 0, \tag{12c}$$
$$P_{1|1} \preceq P_{1|0}, P_{T|T} = \Pi_T, \tag{12d}$$
$$P_{t+1|t+1} \preceq A_t P_{t|t} A_t^\top + W_t, \tag{12e}$$
$$\begin{bmatrix} P_{t|t} - \Pi_t & P_{t|t} A_t^\top \\ A_t P_{t|t} & A_t P_{t|t} A_t^\top + W_t \end{bmatrix} \succeq 0. \tag{12f}$$

  The constraint (12c) is imposed for every $t = 1, \cdots, T$, while (12e) and (12f) are for every $t = 1, \cdots, T-1$. Constants c_1 and c_2 are given by

$$c_1 = \tfrac{1}{2} \log\det P_{1|0} + \tfrac{1}{2} \sum_{t=1}^{T-1} \log\det W_t$$
$$c_2 = \mathrm{Tr}(N_1 P_{1|0}) + \sum_{t=1}^{T} \mathrm{Tr}(W_t S_t).$$

  Define signal-to-noise ratio matrices $\{\mathsf{SNR}_t\}_{t=1}^T$ by

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

  and set $r_t = \mathrm{rank}(\mathsf{SNR}_t)$. Apply the singular value decomposition to find $C_t \in \mathbb{R}^{r_t \times n_t}$ and $V_t \in \mathbb{S}_{++}^{r_t}$ such that

$$\mathsf{SNR}_t = C_t^\top V_t^{-1} C_t, \quad t = 1, \cdots, T. \tag{13}$$

  If $r_t = 0$, C_t and V_t are null (zero dimensional) matrices.

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

$$L_t = P_{t|t-1} C_t^\top (C_t P_{t|t-1} C_t^\top + V_t)^{-1}. \tag{14}$$

  Construct a Kalman filter by

$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1} + L_t (\mathbf{y}_t - C_t \hat{\mathbf{x}}_{t|t-1}) \tag{15a}$$
$$\hat{\mathbf{x}}_{t+1|t} = A_t \hat{\mathbf{x}}_t + B_t \mathbf{u}_t. \tag{15b}$$

  If $r_t = 0$, L_t is a null matrix and (15a) becomes $\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1}$.

An optimization problem (12) plays a key role in the proposed synthesis. Intuitively, (12) "schedules" the optimal sequence of covariance matrices $\{P_{t|t}\}_{t=1}^T$ in such a way that there exists a virtual sensor mechanism to realize it and the required data-rate is minimized. The virtual sensor and the Kalman filter are designed later to realize the scheduled covariance.

Theorem 1: An optimal policy for the problem (4) exists if and only if the max-det problem (12) is feasible, and the optimal value of (4) coincides with the optimal value of (12). If the optimal value of (4) is finite, an optimal policy can be realized by a virtual sensor, Kalman filter, and a certainty equivalence controller as shown in Figure 3. Moreover, each of these components can be constructed by an SDP-based algorithm summarized in Steps 1-3.

Proof: See Section VI. ∎

Remark 1: If W_t is singular for some t, we suggest to factorize it as $W_t = F_t F_t^\top$ and use the following alternative max-det problem instead of (12):

$$\min_{\{P_{t|t},\Delta_t\}_{t=1}^T} \quad \frac{1}{2}\sum\nolimits_{t=1}^{T} \log\det \Delta_t^{-1} + c_1 \tag{16a}$$

$$\text{s.t.} \quad \sum\nolimits_{t=1}^{T} \mathrm{Tr}(\Theta_t P_{t|t}) + c_2 \leq D \tag{16b}$$

$$\Delta_t \succ 0, \tag{16c}$$

$$P_{1|1} \preceq P_{1|0}, P_{T|T} = \Delta_T, \tag{16d}$$

$$P_{t+1|t+1} \preceq A_t P_{t|t} A_t^\top + F_t F_t^\top, \tag{16e}$$

$$\begin{bmatrix} I-\Delta_t & F_t^\top \\ F_t & A_t P_{t|t} A_t^\top + F_t F_t^\top \end{bmatrix} \succeq 0. \tag{16f}$$

The constraint (16c) is imposed for every $t = 1, \cdots, T$, while (16e) and (16f) are for every $t = 1, \cdots, T-1$. Constants c_1 and c_2 are given by $c_1 = \frac{1}{2}\log\det P_{1|0} + \sum_{t=1}^{T-1} \log|\det A_t|$ and $c_2 = \mathrm{Tr}(N_1 P_{1|0}) + \sum_{t=1}^{T} \mathrm{Tr}(F_t^\top S_t F_t)$. This formulation requires that $A_t, t = 1, ..., T-1$ are non-singular matrices. Derivation is omitted for brevity.

B. Time-invariant plants

For time-invariant and infinite-horizon problems (5) and (6), it can be shown that there exists an optimal policy with the same three-stage structure as in Figure 4 in which all components are time-invariant. The optimal policy can be explicitly constructed by the following numerical procedure:

- **Step 1** (Controller design) Find the unique stabilizing solution to an algebraic Riccati equation
$$A^\top SA - S - A^\top SB(B^\top SB + R)^{-1}B^\top SA + Q = 0 \tag{17}$$
and determine the optimal feedback control gain by $K = -(B^\top SB+R)^{-1}B^\top SA$. Set $\Theta = K^\top(B^\top SB + R)K$.
- **Step 2** (Virtual sensor design) Choose P and Π as the solution to a max-det problem:
$$\min_{P,\Pi} \quad \frac{1}{2}\log\det \Pi^{-1} + \frac{1}{2}\log\det W \tag{18a}$$
$$\text{s.t.} \quad \mathrm{Tr}(\Theta P) + \mathrm{Tr}(WS) \leq D, \tag{18b}$$
$$\Pi \succ 0, \tag{18c}$$
$$P \preceq APA^\top + W, \tag{18d}$$
$$\begin{bmatrix} P-\Pi & PA^\top \\ AP & APA^\top + W \end{bmatrix} \succeq 0. \tag{18e}$$
Define $\tilde{P} \triangleq APA^\top + W$, $\mathsf{SNR} \triangleq P^{-1} - \tilde{P}^{-1}$ and set $r = \mathrm{rank}(\mathsf{SNR})$. Choose a virtual sensor $\mathbf{y}_t = C\mathbf{x}_t + \mathbf{v}_t$, $\mathbf{v}_t \sim \mathcal{N}(0, V)$ with matrices $C \in \mathbb{R}^{r\times n}$ and $V \in \mathbb{S}_{++}^r$ such that $C^\top V^{-1} C = \mathsf{SNR}$.
- **Step 3** (Filter design) Design a time-invariant Kalman filter
$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1} + L(\mathbf{z}_t - C\hat{\mathbf{x}}_{t|t-1})$$
$$\hat{\mathbf{x}}_{t+1|t} = A\hat{\mathbf{x}}_t + B\mathbf{u}_t$$
with $L = \tilde{P}C^\top(C\tilde{P}C^\top + V)^{-1}$.

Theorem 2: An optimal policy for (6) exists if and only if a max-det problem (18) is feasible, and the optimal value of (6) coincides with that of (18). Moreover, an optimal policy can be realized by a virtual sensor, Kalman filter, and a certainty equivalence controller as shown in Figure 4, all of which are time-invariant. Each of these components can be constructed by Steps 1-3.

Proof: See Appendix E. ∎

Theorem 2 shows a noteworthy fact that $\mathsf{DI}(D)$ defined by (6) admits a single-letter characterization, i.e., it can be evaluated by solving a finite-dimensional optimization problem (18).

C. Data-rate theorem for mean-square stabilization

Theorem 2 shows that $\mathsf{DI}(D)$ defined by (6) admits a semidefinite representation (18). By analyzing the structure of the optimization problem (18), one can obtain a closed-from expression of the quantity $\lim_{D\to+\infty}\mathsf{DI}(D)$. Notice that this quantity can be interpreted as the minimum data-rate (measured in directed information) required for mean-square stabilization. The next corollary shows a connection between our study in this paper and the data-rate theorem by Nair and Evans [9].

Corollary 1: Denote by $\sigma_+(A)$ the set of eigenvalues λ_i of A such that $|\lambda_i| \geq 1$ counted with multiplicity. Then,

$$\lim_{D\to+\infty} \mathsf{DI}(D) = \sum_{\lambda_i \in \sigma_+(A)} \log|\lambda_i|. \tag{19}$$

Proof: See Appendix F. ∎

Corollary 1 indicates that the minimal data-rate for mean-square stabilization does not depend on the noise property W. This result is consistent with the observation in [9]. However, as is clear from the semidefinite representation (18), minimal data-rate to achieve control performance $J_t \leq D$ depends on W when D is finite.

Corollary 1 has a further implication that there exists a quantized LQG control scheme implementable over a noiseless binary channel such that data-rate is arbitrarily close to (19) and the closed-loop systems in stabilized in the mean-square sense. See [41] for details.

Mean-square stabilizability of linear systems by quantized feedback with Markovian packet losses is considered in [42], where a necessary and sufficient condition in terms of nominal data-rate and packet dropping probability is obtained. Although directed information is not used in [42], it would be an interesting future work to compute $\lim_{T\to\infty}\frac{1}{T}I(X^T \to U^T)$ under the stabilization scheme proposed there and study how it is compared to the right hand side of (19).

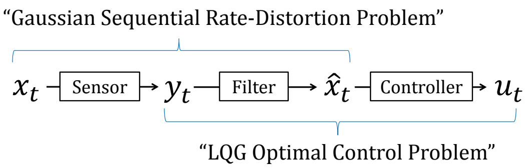

Fig. 4. Sensor-filter-controller separation principle: integration of the sensor-filter and filter-controller separation principles.

D. *Connections to existing results*

We first note that the "sensor-filter-controller" structure identified by Theorem 1 is not a simple consequence of the filter-controller separation principle in the standard LQG control theory [43]. Unlike the standard framework in which a sensor mechanism (10) is given *a priori*, in (4) we *design* a sensor mechanism jointly with other components. Intuitively, a sensor mechanism in our context plays a role to reduce information flow from $\mathbf{y}_t$ to $\mathbf{x}_t$. The proposed sensor design algorithm has already appeared in [44]. In this paper we strengthen the result by showing that the designed linear sensor turns out to be optimal among all nonlinear (Borel measurable) sensor mechanisms.

Information-theoretic fundamental limitations of feedback control are derived in [25]–[28] via the "Bode-like" integrals. However, the connection between [25]–[28] and our problem (4) is not straightforward, and the structural result shown in Figure 3 does not appear in [25]–[28]. Also, we note that our problem formulation (4) is different from networked LQG control problem over Gaussian channels [12], [14], [45] where a model of Gaussian channel is given *a priori*. In such problems, linearity of the optimal policy is already reported [4, Ch.10,11].

It should be noted that problem (4) is closely related to the sequential rate-distortion problem (also called zero-delay or non-anticipative rate-distortion problem) [6], [46], [47]. In the Gaussian sequential rate-distortion problem where the plant (1) is an uncontrolled system (i.e., $\mathbf{u}_t = 0$), it can be shown that the optimal policy can be realized by a two-stage "sensor-filter" structure [46]. However, the same result is not known for the case in which feedback controllers must be designed simultaneously. Relevant papers towards this direction include [47]–[49], where Csiszár's formulation of rate-distortion functions [50] is extended to the non-anticipative regime. In particular, [49] considers non-anticipative rate-distortion problems with feedback. In [51] and [52], LQG control problems with information-theoretic costs similar to (4) are considered. However, the optimization problem considered in these papers are not equivalent to (4), and the structural result shown in Figure 4 does not appear.

In a very recent paper [24, Lemma 3.1], it is independently reported that the optimal policy for (4) can be realized by an additive white Gaussian noise (AWGN) channel and linear filters. While this result is compatible to ours, it is noteworthy that the proof technique there is different from ours and is based on fundamental inequalities for directed information obtained in [39]. In comparison to [24], we additionally prove that the optimal control policy can be realized by a state space model with a three-stage structure (shown in Figure 3, 4), which appears to be a new observation to the best of our knowledge.

The SDP-based algorithms to solve (4), (6) and (38) are newly developed in this paper, using the techniques presented in [46] and [44]. Due to the lack of analytical expression of the optimal policy (especially for MIMO and time-varying plants), the use of optimization-based algorithms seems critical. In [53], an iterative water-filling algorithm is proposed for a highly relevant problem. In this paper, the main algorithmic tool is SDP, which allows us to generalize the results in [22]–[24] to MIMO and time-varying settings.

V. Example

In this section, we consider a simple numerical example to demonstrate the SDP-based control design presented in Section IV-B. Consider a time-invariant plant (5) with randomly generated matrices

$$A = \begin{bmatrix} 0.12 & 0.63 & -0.52 & 0.33 \\ 0.26 & -1.28 & 1.57 & 1.13 \\ -1.77 & -0.30 & 0.77 & 0.25 \\ -0.16 & 0.20 & -0.58 & 0.56 \end{bmatrix}, W = \begin{bmatrix} 4.94 & -0.10 & 1.29 & 0.35 \\ & 5.55 & 2.07 & 0.31 \\ & & 2.02 & 1.43 \\ \text{sym.} & & & 3.10 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.66 & -0.58 & 0.03 & -0.20 \\ 2.61 & -0.91 & 0.87 & -0.07 \\ -0.64 & -1.12 & -0.19 & 0.61 \\ 0.93 & 0.58 & -1.18 & -1.21 \end{bmatrix},$$

and the optimization problem (6) with $Q = I$ and $R = I$. By solving (18) with various D, we obtain the rate-performance trade-off curve shown in Figure 5 (top left). The vertical asymptote $D = \text{Tr}(WS)$ corresponds to the best achievable control performance when unrestricted amount of information about the state is available. This corresponds to the performance of the state-feedback linear-quadratic regulator (LQR). The horizontal asymptote $\sum_{\lambda_i \in \sigma_+(A)} \log|\lambda_i| = 1.169$ [bits/sample] is the minimum data-rate to achieve mean-square stability. Figure 5 (bottom left) shows the rank of SNR matrices as a function of D. Since SNR is computed numerically by an SDP solver with some finite numerical precision, rank(SNR) is obtained by truncating singular values smaller than 0.1% of the maximum singular value. Figure 5 (right) shows selected singular values at $D = 33, 40$ and 80. Observe the phase transition (rank dropping) phenomena. The optimal dimension of the sensor output changes as D changes.

Specifically, the minimum data-rate to achieve control performance $D = 33$ is found to be 6.133 [bits/sample]. The optimal sensor mechanism $\mathbf{y}_t = C\mathbf{x}_t + \mathbf{v}_t, \mathbf{v}_t \sim \mathcal{N}(0, V)$ to achieve this performance is given by

$$C = \begin{bmatrix} -0.864 & 0.258 & -0.205 & -0.382 \\ -0.469 & -0.329 & 0.662 & 0.483 \\ -0.130 & 0.332 & -0.502 & 0.780 \end{bmatrix}, V = \begin{bmatrix} 0.029 & 0 & 0 \\ 0 & 0.208 & 0 \\ 0 & 0 & 1.435 \end{bmatrix}.$$

If $D = 40$, required data-rate is 3.266 [bits/sample] and the optimal sensor is given by

$$C = \begin{bmatrix} -0.886 & 0.241 & -0.170 & -0.359 \\ -0.431 & -0.350 & 0.647 & 0.523 \end{bmatrix}, \quad V = \begin{bmatrix} 0.208 & 0 \\ 0 & 2.413 \end{bmatrix}.$$

Similarly, minimum data-rate to achieve $D = 80$ is 1.602 [bits/sample], and this is achieved by a sensor mechanism with

$$C = \begin{bmatrix} -0.876 & 0.271 & -0.169 & -0.362 \end{bmatrix}, \quad V = 1.775.$$

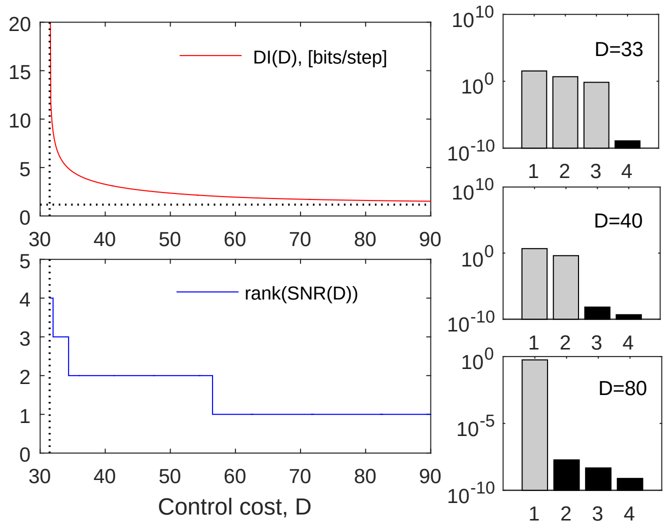

Fig. 5. (Top left) Data rate $\mathsf{DI}(D)$ [bits/step] required to achieve control performance D. (Bottom left) Rank of SNR(D), evaluated after truncating singular values smaller than 0.1% of the maximum singular value. (Right) Singular values of SNR(D) evaluated at $D = 33$, 40 and 80. Truncated singular values are shown in block bars. An SDP solver SDPT3 [54] with YALMIP [55] interface is used.

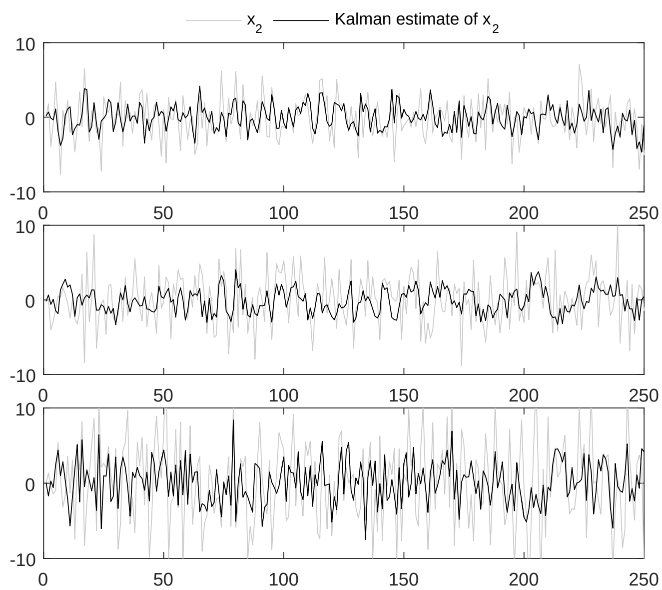

Fig. 6. Closed-loop performances of the controllers designed for $D = 33$ (top), $D = 40$ (middle), and $D = 80$ (bottom). Trajectories of the second component of the state vector and their Kalman estimates are shown.

Figure 6 shows the closed-loop responses of the state trajectories simulated in each scenario.

VI. Derivation of Main Result

This section is devoted to prove Theorem 1. We first define subsets Γ_0, Γ_1, and Γ_2 of the policy space Γ as follows.

Γ_0 : The space of policies with three-stage separation structure explained in Section IV.

Γ_1 : The space of linear sensors without memory followed by linear deterministic feedback control. Namely, a policy $\mathbb{P}(u^T \| x^T)$ in Γ_1 can be expressed as a composition of

$$\mathbf{y}_t = C_t \mathbf{x}_t + \mathbf{v}_t, \;\; \mathbf{v}_t \sim \mathcal{N}(0, V_t) \tag{20}$$

and $\mathbf{u}_t = l_t(\mathbf{y}^t)$, where $C_t \in \mathbb{R}^{r_t \times n_t}$, r_t is some nonnegative integer, $V_t \succ 0$, and $l_t(\cdot)$ is a linear map.

Γ_2 : The space of linear policies without state memory. Namely, a policy $\mathbb{P}(u^T \| x^T)$ in Γ_2 can be expressed as

$$\mathbf{u}_t = M_t \mathbf{x}_t + N_t \mathbf{u}^{t-1} + \mathbf{g}_t, \;\; \mathbf{g}_t \sim \mathcal{N}(0, G_t) \tag{21}$$

with some matrices M_t, N_t, and $G_t \succeq 0$.

A. Proof outline

To prove Theorem 1, we establish a chain of inequalities:

$$\inf_{\gamma \in \Gamma : J_\gamma \le D} I_\gamma(\mathbf{x}^T \to \mathbf{u}^T) \tag{22a}$$
$$\ge \inf_{\gamma \in \Gamma : J_\gamma \le D} \sum\nolimits_{t=1}^T I_\gamma(\mathbf{x}_t; \mathbf{u}_t | \mathbf{u}^{t-1}) \tag{22b}$$
$$\ge \inf_{\gamma \in \Gamma_2 : J_\gamma \le D} \sum\nolimits_{t=1}^T I_\gamma(\mathbf{x}_t; \mathbf{u}_t | \mathbf{u}^{t-1}) \tag{22c}$$
$$\ge \inf_{\gamma \in \Gamma_1 : J_\gamma \le D} \sum\nolimits_{t=1}^T I_\gamma(\mathbf{x}_t; \mathbf{y}_t | \mathbf{y}^{t-1}) \tag{22d}$$
$$\ge \inf_{\gamma \in \Gamma_0 : J_\gamma \le D} \sum\nolimits_{t=1}^T I_\gamma(\mathbf{x}_t; \mathbf{y}_t | \mathbf{y}^{t-1}) \tag{22e}$$
$$\ge \inf_{\gamma \in \Gamma_0 : J_\gamma \le D} I_\gamma(\mathbf{x}^T \to \mathbf{u}^T). \tag{22f}$$

Since $\Gamma_0 \subset \Gamma$, clearly (22a) $\le$ (22f). Thus, showing the above chain of inequalities proves that all quantities in (22) are equal. This observation implies that the search for an optimal solution to our main problem (4) can be restricted to the class Γ_0 without loss of performance. The first inequality (22b) is immediate from the definition of directed information. We prove inequalities (22c), (22d), (22e) and (22f) in subsequent subsections VI-B, VI-C, VI-D and VI-E. It will follow from the proof of inequality (22f) that an optimal solution to (22e), if exists, is also an optimal solution to (22f). In particular, this implies that an optimal solution to the original problem (22a), if exists, can be found by solving a simplified problem (22e). This observation establishes the sensor-filter-controller separation principle depicted in Figure 3.

Then, we focus on solving problem (22e) in Subsection VI-F. We show that problem (22e) 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.

B. Proof of inequality (22c)

We will show that for every $\gamma_{\mathbb{P}} = \{\mathbb{P}(u_t | x^t, u^{t-1})\}_{t=1}^T \in \Gamma$ that attains a finite objective value in (22b), there exists $\gamma_{\mathbb{Q}} = \{\mathbb{Q}(u_t | x^t, u^{t-1})\}_{t=1}^T \in \Gamma_2$ such that $J_{\mathbb{P}} = J_{\mathbb{Q}}$ and

$$\sum\nolimits_{t=1}^T I_{\mathbb{P}}(\mathbf{x}_t; \mathbf{u}_t | \mathbf{u}^{t-1}) \ge \sum\nolimits_{t=1}^T I_{\mathbb{Q}}(\mathbf{x}_t; \mathbf{u}_t | \mathbf{u}^{t-1})$$

where subscripts of I and J indicate probability measures on which these quantities are evaluated. Without loss of generality, we assume $\mathbb{P}(x^{T+1}, u^T)$ has zero-mean. Otherwise, we can consider an alternative policy $\gamma_{\tilde{\mathbb{P}}} = \{\tilde{\mathbb{P}}(u_t | x^t, u^{t-1})\}_{t=1}^T$, where

$$\tilde{\mathbb{P}}(u_t | x^t, u^{t-1}) \triangleq \mathbb{P}(u_t + \mathbb{E}_{\mathbb{P}}(\mathbf{u}_t) | x^t + \mathbb{E}_{\mathbb{P}}(\mathbf{x}^t), u^{t-1} + \mathbb{E}_{\mathbb{P}}(\mathbf{u}^{t-1}))$$

which generates a zero-mean joint distribution $\tilde{\mathbb{P}}(x^{T+1}, u^T)$. We have $I_{\tilde{\mathbb{P}}} = I_{\mathbb{P}}$ in view of the translation invariance of mutual information, and $J_{\tilde{\mathbb{P}}} \leq J_{\mathbb{P}}$ due to the fact that the cost function is quadratic.

First, we consider a zero-mean, jointly Gaussian probability measure $\mathbb{G}(x^{T+1}, u^T)$ having the same covariance matrix as $\mathbb{P}(x^{T+1}, u^T)$.

Lemma 2: The following inequality holds whenever the left hand side is finite.

$$\sum\nolimits_{t=1}^{T} I_{\mathbb{P}}(\mathbf{x}_t; \mathbf{u}_t | \mathbf{u}^{t-1}) \geq \sum\nolimits_{t=1}^{T} I_{\mathbb{G}}(\mathbf{x}_t; \mathbf{u}_t | \mathbf{u}^{t-1}) \tag{23}$$

Proof: See Appendix C. ■

Next, we are going to construct a policy $\gamma_{\mathbb{Q}} = \{\mathbb{Q}(u_t|x^t, u^{t-1})\}_{t=1}^{T} \in \Gamma_2$ using a jointly Gaussian measure $\mathbb{G}(x^{T+1}, u^T)$. Let $E_t\mathbf{x}_t + F_t\mathbf{u}^{t-1}$ be the least mean-square error estimate of $\mathbf{u}_t$ given $(\mathbf{x}_t, \mathbf{u}^{t-1})$ in $\mathbb{G}(x^{T+1}, u^T)$, and let V_t be the resulting estimation error covariance matrix. Define a stochastic kernel $\mathbb{Q}(u_t|x_t, u^{t-1})$ by $\mathbb{Q}(u_t|x_t, u^{t-1}) = \mathcal{N}(E_t\mathbf{x}_t + F_t\mathbf{u}^{t-1}, V_t)$. By construction, $\mathbb{Q}(u_t|x_t, u^{t-1})$ satisfies[2]

$$\mathrm{d}\mathbb{G}(x_t, u^t) = \mathrm{d}\mathbb{Q}(u_t|x_t, u^{t-1})\mathrm{d}\mathbb{G}(x_t, u^{t-1}). \tag{24}$$

Define $\mathbb{Q}(x^{T+1}, u^T)$ recursively by

$$\mathrm{d}\mathbb{Q}(x^t, u^{t-1}) = \mathrm{d}\mathbb{P}(x_t|x_{t-1}, u_{t-1})\mathrm{d}\mathbb{Q}(x^{t-1}, u^{t-1}) \tag{25}$$

$$\mathrm{d}\mathbb{Q}(x^t, u^t) = \mathrm{d}\mathbb{Q}(u_t|x_t, u^{t-1})\mathrm{d}\mathbb{Q}(x^t, u^{t-1}) \tag{26}$$

where $\mathbb{P}(x_t|x_{t-1}, u_{t-1})$ is a stochastic kernel defined by (1). The following identity holds between two Gaussian measures $\mathbb{G}(x^{T+1}, u^T)$ and $\mathbb{Q}(x^{T+1}, u^T)$.

Lemma 3: $\mathbb{G}(x_{t+1}, u^t) = \mathbb{Q}(x_{t+1}, u^t) \;\; \forall t = 1, \cdots, T$.

Proof: See Appendix D. ■

We are now ready to prove (22c). First, replacing a policy $\gamma_{\mathbb{P}}$ with a new policy $\gamma_{\mathbb{Q}}$ does not change the LQG control cost.

$$\begin{aligned} J_{\gamma_{\mathbb{P}}} &= \int \left(\|x_{t+1}\|_{Q_t}^2 + \|u_t\|_{R_t}^2\right) \mathrm{d}\mathbb{P}(x_{t+1}, u^t) \\ &= \int \left(\|x_{t+1}\|_{Q_t}^2 + \|u_t\|_{R_t}^2\right) \mathrm{d}\mathbb{G}(x_{t+1}, u^t) && \text{(27a)} \\ &= \int \left(\|x_{t+1}\|_{Q_t}^2 + \|u_t\|_{R_t}^2\right) \mathrm{d}\mathbb{Q}(x_{t+1}, u^t) && \text{(27b)} \\ &= J_{\gamma_{\mathbb{Q}}}. \end{aligned}$$

Equality (27a) holds since $\mathbb{P}$ and $\mathbb{G}$ have the same second order moments. Step (27b) follows from Lemma 3. Second, replacing $\gamma_{\mathbb{P}}$ with $\gamma_{\mathbb{Q}}$ does not increase the information cost.

$$\begin{aligned} \sum\nolimits_{t=1}^{T} I_{\mathbb{P}}(\mathbf{x}_t; \mathbf{u}_t | \mathbf{u}^{t-1}) &\geq \sum\nolimits_{t=1}^{T} I_{\mathbb{G}}(\mathbf{x}_t; \mathbf{u}_t | \mathbf{u}^{t-1}) && \text{(28a)} \\ &= \sum\nolimits_{t=1}^{T} I_{\mathbb{Q}}(\mathbf{x}_t; \mathbf{u}_t | \mathbf{u}^{t-1}). && \text{(28b)} \end{aligned}$$

The inequality (28a) is due to Lemma 2. In (28b), $I_{\mathbb{G}}(\mathbf{x}_t; \mathbf{u}_t | \mathbf{u}^{t-1}) = I_{\mathbb{Q}}(\mathbf{x}_t; \mathbf{u}_t | \mathbf{u}^{t-1})$ holds for every $t = 1, \cdots, T$ because of Lemma 3.

[2]Equation $\mathrm{d}\mathbb{P}(x, y) = \mathrm{d}\mathbb{P}(y|x)\mathrm{d}\mathbb{P}(x)$ is a short-hand notation for $\mathbb{P}(B_X \times B_Y) = \int_{B_X} \mathbb{P}(B_Y|x)\mathrm{d}\mathbb{P}(x) \; \forall B_X \in \mathcal{B}_{\mathcal{X}}, B_Y \in \mathcal{B}_{\mathcal{Y}}$.

C. *Proof of inequality (22d)*

Given a policy $\gamma_2 \in \Gamma_2$, we are going to construct a policy $\gamma_1 \in \Gamma_1$ such that $J_{\gamma_1} = J_{\gamma_2}$ and

$$I_{\gamma_2}(\mathbf{x}_t; \mathbf{u}_t | \mathbf{u}^{t-1}) = I_{\gamma_1}(\mathbf{x}_t; \mathbf{y}_t | \mathbf{y}^{t-1}) \tag{29}$$

for every $t = 1, \cdots, T$. Let $\gamma_2 \in \Gamma_2$ be given by

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

Define $\tilde{\mathbf{y}}_t \triangleq M_t\mathbf{x}_t + \mathbf{g}_t$. If we write $N_t\mathbf{u}^{t-1} = N_{t,t-1}\mathbf{u}_{t-1} + \cdots + N_{t,1}\mathbf{u}_1$, it can be seen that $\mathbf{u}^t$ and $\tilde{\mathbf{y}}^t$ are related by an invertible linear map

$$\begin{bmatrix} \tilde{\mathbf{y}}_1 \\ \vdots \\ \vdots \\ \tilde{\mathbf{y}}_t \end{bmatrix} = \begin{bmatrix} I & 0 & \cdots & 0 \\ -N_{2,1} & I & & \vdots \\ \vdots & & \ddots & 0 \\ -N_{t,1} & \cdots & -N_{t,t-1} & I \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \vdots \\ \mathbf{u}_t \end{bmatrix} \tag{30}$$

for every $t = 1, \cdots, T$. Hence,

$$\begin{aligned} I(\mathbf{x}_t; \mathbf{u}_t | \mathbf{u}^{t-1}) &= I(\mathbf{x}_t; \tilde{\mathbf{y}}_t + N_t\mathbf{u}^{t-1} | \tilde{\mathbf{y}}^{t-1}, \mathbf{u}^{t-1}) \\ &= I(\mathbf{x}_t; \tilde{\mathbf{y}}_t | \tilde{\mathbf{y}}^{t-1}). \end{aligned} \tag{31}$$

Let $G_t = E_t^\top V_t E_t$ be the (thin) singular value decomposition. Since we assume (31) is bounded, we must have

$$\mathrm{Im}(M_t) \subseteq \mathrm{Im}(G_t) = \mathrm{Im}(E_t^\top). \tag{32}$$

Otherwise, the component of $\mathbf{u}_t$ in $\mathrm{Im}(G_t)^\perp$ depends deterministically on $\mathbf{x}_t$ and (31) is unbounded. Now, define $\mathbf{y}_t \triangleq E_t\tilde{\mathbf{y}}_t = E_tM_t\mathbf{x}_t + E_t\mathbf{g}_t, \;\; \mathbf{g}_t \sim \mathcal{N}(0, G_t)$. Then, we have

$$\begin{aligned} E_t^\top\mathbf{y}_t &= E_t^\top E_tM_t\mathbf{x}_t + E_t^\top E_t\mathbf{g}_t, \;\; \mathbf{g}_t \sim \mathcal{N}(0, G_t) \\ &= M_t\mathbf{x}_t + \mathbf{g}_t = \tilde{\mathbf{y}}_t. \end{aligned}$$

In the second line, we used the facts that $E_t^\top E_tM_t = M_t$ and $E_t^\top E_t\mathbf{g}_t = \mathbf{g}_t$ under (32). Thus, we have $\mathbf{y}_t = E_t\tilde{\mathbf{y}}_t$ and $\tilde{\mathbf{y}}_t = E_t^\top\mathbf{y}_t$. This implies that $\mathbf{y}_t$ and $\tilde{\mathbf{y}}_t$ contain statistically equivalent information, and that

$$I(\mathbf{x}_t; \tilde{\mathbf{y}}_t | \tilde{\mathbf{y}}^{t-1}) = I(\mathbf{x}_t; \mathbf{y}_t | \mathbf{y}^{t-1}). \tag{33}$$

Also, since $\mathbf{u}_t$ depends linearly on $\tilde{\mathbf{y}}^t$ by (30), there exists a linear map l_t such that

$$\mathbf{u}_t = l_t(\mathbf{y}^t). \tag{34}$$

Setting $C_t \triangleq E_tM_t$, construct a policy $\gamma_1 \in \Gamma_1$ using $\mathbf{y}_t \triangleq E_t\tilde{\mathbf{y}}_t = C_t\mathbf{x}_t + \mathbf{v}_t$ with $\mathbf{v}_t \sim \mathcal{N}(0, V_t)$ and a linear map (34). Since joint distribution $\mathbb{P}(x^{T+1}, u^T)$ is the same under γ_1 and γ_2, we have $J_{\gamma_1} = J_{\gamma_2}$. From (31) and (33), we also have (29).

D. *Proof of inequality (22e)*

Notice that for every $\gamma \in \Gamma_1$, conditional mutual information can be written in terms of $P_{t|t} = \mathrm{Cov}(\mathbf{x}_t - \mathbb{E}(\mathbf{x}_t|\mathbf{y}^t, \mathbf{u}^{t-1}))$:

$$\begin{aligned} &I_\gamma(\mathbf{x}_t; \mathbf{y}_t | \mathbf{y}^{t-1}) \\ =&I_\gamma(\mathbf{x}_t; \mathbf{y}_t | \mathbf{y}^{t-1}, \mathbf{u}^{t-1}) \\ =&h(\mathbf{x}_t | \mathbf{y}^{t-1}, \mathbf{u}^{t-1}) - h(\mathbf{x}_t | \mathbf{y}^t, \mathbf{u}^{t-1}) \\ =&\tfrac{1}{2}\log\det(A_{t-1}P_{t-1|t-1}A_{t-1}^\top + W_{t-1}) - \tfrac{1}{2}\log\det P_{t|t}. \end{aligned} \tag{35}$$

Moreover, for every fixed sensor equation (20), covariance matrices are determined by the Kalman filtering formula

$$P_{t|t} = ((A_{t-1}P_{t-1|t-1}A_{t-1}^\top + W_{t-1})^{-1} + \mathsf{SNR}_t)^{-1}.$$

Hence, conditional mutual information (35) depends only on the choice of $\{\mathsf{SNR}_t\}_{t=1}^T$, and is independent of the choice of a linear map l_t. On the other hand, the LQG control cost J_γ depends on the choice of l_t. In particular, for every fixed linear sensor (20), it follows from the standard filter-controller separation principle in the LQG control theory that the optimal l_t that minimizes J_γ is a composition of a Kalman filter $\hat{\mathbf{x}}_t = \mathbb{E}(\mathbf{x}_t|\mathbf{y}^t, \mathbf{u}^{t-1})$ and a certainty equivalence controller $\mathbf{u}_t = K_t\hat{\mathbf{x}}_t$. This implies that an optimal solution γ can always be found in the class Γ_0, establishing the inequality in (22e).

For a fixed linear sensor (20), an explicit form of the Kalman filter and the certainty equivalence controller is given by Steps 1 and 3 in Section IV. Derivation is standard and hence is omitted. It is also possible to write J_γ explicitly as

$$J_\gamma = \text{Tr}(N_1 P_{1|0}) + \sum\nolimits_{t=1}^T \left(\text{Tr}(W_t S_t) + \text{Tr}(\Theta_t P_{t|t})\right). \quad (36)$$

Derivation of (36) is also straightforward, and can be found in [44, Lemma 1].

E. Proof of inequality (22f)

For every fixed $\gamma \in \Gamma_0$, by Lemma 1 we have

$$\begin{aligned} I_\gamma(\mathbf{x}^T \to \mathbf{u}^T) &\leq I_\gamma(\mathbf{x}^T \to \mathbf{y}^T \| \mathbf{u}_+^{T-1}) \\ &= \sum\nolimits_{t=1}^T I_\gamma(\mathbf{x}^t; \mathbf{y}_t | \mathbf{y}^{t-1}, \mathbf{u}^{t-1}) \\ &= \sum\nolimits_{t=1}^T I_\gamma(\mathbf{x}^t; \mathbf{y}_t | \mathbf{y}^{t-1}) \\ &= \sum\nolimits_{t=1}^T I_\gamma(\mathbf{x}_t; \mathbf{y}_t | \mathbf{y}^{t-1}) + I_\gamma(\mathbf{x}^{t-1}; \mathbf{y}_t | \mathbf{x}_t, \mathbf{y}^{t-1}) \\ &= \sum\nolimits_{t=1}^T I_\gamma(\mathbf{x}_t; \mathbf{y}_t | \mathbf{y}^{t-1}). \end{aligned}$$

The last equality holds since, by construction, $\mathbf{y}_t = C_t\mathbf{x}_t + \mathbf{v}_t$ is conditionally independent of $\mathbf{x}^{t-1}$ given $\mathbf{x}_t$.

F. SDP formulation of problem (22e)

Invoking (35) and (36) hold for every $\gamma \in \Gamma_0$, problem (22e) can be written as an optimization problem in terms of $\{P_{t|t}, \mathsf{SNR}_t\}_{t=1}^T$ as

$$\begin{aligned} \min \; & \sum_{t=2}^T \left(\frac{1}{2}\log\det(A_{t-1}P_{t-1|t-1}A_{t-1}^\top + W_t) - \frac{1}{2}\log\det P_{t|t}\right) \\ & + \frac{1}{2}\log\det P_{1|0} - \frac{1}{2}\log\det P_{1|1} \\ \text{s.t. } & \text{Tr}(N_1 P_{1|0}) + \sum\nolimits_{t=1}^T \left(\text{Tr}(W_t S_t) + \text{Tr}(\Theta_t P_{t|t})\right) \leq D, \\ & P_{1|1}^{-1} = P_{1|0}^{-1} + \mathsf{SNR}_1, \\ & P_{t|t}^{-1} = (A_{t-1}P_{t-1|t-1}A_{t-1}^\top + W_{t-1})^{-1} + \mathsf{SNR}_t, \; t = 2, ..., T \\ & \mathsf{SNR}_t \succeq 0, \quad t = 1, ..., T. \end{aligned}$$

This problem can be reformulated as a max-det problem as follows. First, variables $\{\mathsf{SNR}_t\}_{t=1}^T$ are eliminated from the problem by replacing the last three constraints with equivalent conditions

$$\begin{aligned} & 0 \prec P_{1|1} \preceq P_{1|0}, \\ & 0 \prec P_{t|t} \preceq A_{t-1}P_{t-1|t-1}A_{t-1}^\top + W_{t-1}, \; t = 2, ..., T. \end{aligned}$$

Second, the following equalities can be used for $t = 1, ..., T-1$ to rewrite the objective function:

$$\begin{aligned} & \frac{1}{2}\log\det(A_t P_{t|t} A_t^\top + W_t) - \frac{1}{2}\log\det P_{t|t} \\ = \; & \frac{1}{2}\log\det(P_{t|t}^{-1} + A_t^\top W_t^{-1} A_t) + \frac{1}{2}\log\det W_t & (37a) \\ = \; & \inf_{\Pi_t} \; \frac{1}{2}\log\det \Pi_t^{-1} + \frac{1}{2}\log\det W_t & (37b) \\ & \text{s.t.} \quad 0 \prec \Pi_t \preceq (P_{t|t}^{-1} + A_t^\top W_t^{-1} A_t)^{-1} \\ = \; & \inf_{\Pi_t} \; \frac{1}{2}\log\det \Pi_t^{-1} + \frac{1}{2}\log\det W_t & (37c) \\ & \text{s.t.} \quad \Pi_t \succ 0, \; \begin{bmatrix} P_{t|t} - \Pi_t & P_{t|t}A_t^\top \\ A_t P_{t|t} & A_t P_{t|t} A_t^\top + W_t \end{bmatrix} \succeq 0. \end{aligned}$$

In step (37a), we have used the matrix determinant theorem [56, Theorem 18.1.1]. An additional variable Π_t is introduced in step (37b). The constraint is rewritten using the matrix inversion lemma in (37c).

These two techniques allow us to formulate the above problem as a max-det problem (12). Thus, we have shown that Steps 1-3 in Section IV provide an optimal solution to problem (22d), which is also an optimal solution to the original problem (22a).

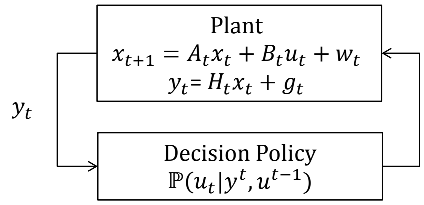

Fig. 7. LQG control of partially observable plant with minimum directed information.

VII. Extension to Partially Observable Plants

So far, our focus has been on a control system in Figure 1 in which the decision policy has an access to the state $\mathbf{x}_t$ of the plant. Often in practice, the state of the plant is only partially observable through a given physical sensor mechanism. We now consider an extension of the control synthesis to partially observable plants.

Consider a control system in Figure 7 where a state space model (1) and a sensor model $\mathbf{y}_t = H_t\mathbf{x}_t + \mathbf{g}_t$ are given. We assume that initial state $\mathbf{x}_1 \sim \mathcal{N}(0, P_{1|0})$, $P_{1|0} \succ 0$ and noise processes $\mathbf{w}_t \sim \mathcal{N}(0, W_t)$, $W_t \succ 0$, $\mathbf{g}_t \sim \mathcal{N}(0, G_t)$, $G_t \succeq 0$, $t = 1, ..., T$ are mutually independent. We also assume that H_t has full row rank for $t = 1, ..., T$. Consider the following problem:

$$\begin{aligned} \min_{\gamma \in \Gamma} \quad & I_\gamma(\mathbf{y}^T \to \mathbf{u}^T) & (38a) \\ \text{s.t.} \quad & J_\gamma(\mathbf{x}^{T+1}, \mathbf{u}^T) \leq D. & (38b) \end{aligned}$$

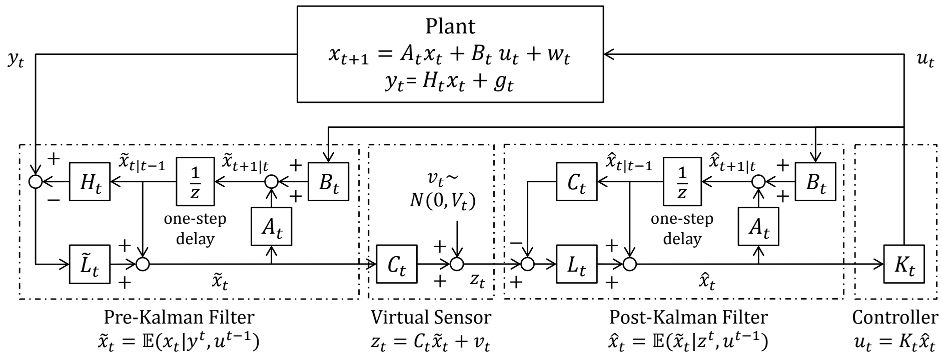

Fig. 8. Structure of optimal control policy for problem (38). Matrices $\tilde{L}_t$, C_t, V_t, L_t and K_t are determined by SDP-based algorithm in Section VII.

where Γ is the space of policies $\gamma = \mathbb{P}(u^T \| y^T)$. Relevant optimization problems to (38) are considered in [22]–[24] in the context of Section III. Based on the control synthesis developed so far for fully observable plants, it can be shown that the optimal control policy can be realized by the architecture shown in Figure 8. Moreover, as in the fully observable cases, the optimal control policy can be synthesized by an SDP-based algorithm.

Step 1. (Pre-Kalman filter design) Design a Kalman filter

$$\tilde{\mathbf{x}}_t = \tilde{\mathbf{x}}_{t|t-1} + \tilde{L}_t(\mathbf{y}_t - H_t \tilde{\mathbf{x}}_{t|t-1}) \tag{39a}$$

$$\tilde{\mathbf{x}}_{t+1|t} = A_t \tilde{\mathbf{x}}_t + B_t \mathbf{u}_t, \; \tilde{\mathbf{x}}_{1|0} = 0 \tag{39b}$$

where the Kalman gains $\{\tilde{L}_t\}_{t=1}^{T+1}$ are computed by

$$\begin{aligned}
\tilde{L}_t &= \tilde{P}_{t|t-1} H_t^\top (H_t \tilde{P}_{t|t-1} H_t^\top + G_t)^{-1} \\
\tilde{P}_{t|t} &= (I - \tilde{L}_t H_t) \tilde{P}_{t|t-1} \\
\tilde{P}_{t+1|t} &= A_t \tilde{P}_{t|t} A_t^\top + W_t.
\end{aligned}$$

Matrices $\Psi_t = \tilde{L}_{t+1}(H_{t+1} \tilde{P}_{t+1|t} H_{t+1}^\top + G_{t+1}) \tilde{L}_{t+1}^\top$ will be used in Step 3.

Step 2. (Controller design) Determine feedback control gains K_t via the backward Riccati recursion:

$$S_t = \begin{cases} Q_t & \text{if } t = T \\ Q_t + N_{t+1} & \text{if } t = 1, \cdots, T-1 \end{cases} \tag{40a}$$

$$M_t = B_t^\top S_t B_t + R_t \tag{40b}$$

$$N_t = A_t^\top (S_t - S_t B_t M_t^{-1} B_t^\top S_t) A_t \tag{40c}$$

$$K_t = -M_t^{-1} B_t^\top S_t A_t \tag{40d}$$

$$\Theta_t = K_t^\top M_t K_t \tag{40e}$$

Positive semidefinite matrices Θ_t will be used in Step 3.

Step 3. (Virtual sensor design) Solve a max-det problem with respect to $\{P_{t|t}, \Pi_t\}_{t=1}^T$:

$$\min \quad \frac{1}{2} \sum_{t=1}^{T} \log\det \Pi_t^{-1} + c_1 \tag{41a}$$

$$\text{s.t.} \quad \sum_{t=1}^{T} \operatorname{Tr}(\Theta_t P_{t|t}) + c_2 \le D \tag{41b}$$

$$\Pi_t \succ 0, \tag{41c}$$

$$P_{1|1} \preceq P_{1|0}, P_{T|T} = \Pi_T, \tag{41d}$$

$$P_{t+1|t+1} \preceq A_t P_{t|t} A_t^\top + \Psi_t, \tag{41e}$$

$$\begin{bmatrix} P_{t|t} - \Pi_t & P_{t|t} A_t^\top \\ A_t P_{t|t} & A_t P_{t|t} A_t^\top + \Psi_t \end{bmatrix} \succeq 0. \tag{41f}$$

The constraint (41c) is imposed for every $t = 1, \cdots, T$, while (41e) and (41f) are for every $t = 1, \cdots, T-1$. Constants c_1 and c_2 are given by

$$\begin{aligned}
c_1 &= \frac{1}{2} \log\det P_{1|0} + \frac{1}{2} \sum_{t=1}^{T-1} \log\det \Psi_t \\
c_2 &= \operatorname{Tr}(N_1 P_{1|0}) + \sum_{t=1}^{T} \operatorname{Tr}(\Psi_t S_t).
\end{aligned}$$

If Ψ_t is singular for some t, consider an alternative max-det problem suggested in Remark 1. Set $r_t = \operatorname{rank}(P_{t|t}^{-1} - P_{t|t-1}^{-1})$, where

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

Choose matrices $C_t \in \mathbb{R}^{r_t \times n_t}$ and $V_t \in \mathbb{S}_{++}^{r_t}$ so that

$$C_t^\top V_t^{-1} C_t = P_{t|t}^{-1} - P_{t|t-1}^{-1} \tag{42}$$

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 4. (Post-Kalman filter design) Design a Kalman filter

$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1} + \hat{L}_t(\mathbf{z}_t - C_t \hat{\mathbf{x}}_{t|t-1}) \tag{43a}$$

$$\hat{\mathbf{x}}_{t+1|t} = A_t \hat{\mathbf{x}}_t + B_t \mathbf{u}_t. \tag{43b}$$

where Kalman gains $\hat{L}_t$ are computed by

$$\hat{L}_t = P_{t|t-1} C_t^\top (C_t P_{t|t-1} C_t^\top + V_t)^{-1}. \tag{44}$$

If $r_t = 0$, L_t is a null matrix and (43a) is simply replaced by $\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1}$.

Theorem 3: An optimal policy for the problem (38) exists if and only if the max-det problem (41) is feasible, and the optimal value of (38) coincides with the optimal value of (41). If the optimal value of (38) is finite, an optimal policy can be realized by an interconnection of a pre-Kalman filter, a virtual sensor, post-Kalman filter, and a certainty equivalence controller as shown in Figure 8. Moreover, each of these components can be constructed by an SDP-based algorithm summarized in Steps 1-4 above.

Proof: See Appendix G. ∎

VIII. Conclusion

In this paper, we considered an optimal control problem in which directed information from the observed output of the plant to the control input is minimized subject to the constraint that the control policy achieves the desired LQG control performance. When the state of the plant is directly observable, the optimal control policy can be realized by a three-stage structure comprised of (1) linear sensor with additive Gaussian noise, (2) Kalman filter, and (3) certainty equivalence controller. An extension to partially observable plants was also discussed. In both cases, the optimal policy is synthesized by an efficient numerical algorithm based on SDP.

Appendix

A. Data-processing inequality for directed information

Lemma 1 is shown as follows. Notice that the following chain of equalities hold for every $t=1,\cdots,T$.

$$\begin{aligned}
&I(\mathbf{x}^t;\mathbf{a}_t|\mathbf{a}^{t-1},\mathbf{u}^{t-1})-I(\mathbf{x}^t;\mathbf{u}_t|\mathbf{u}^{t-1})\\
=&I(\mathbf{x}^t;\mathbf{a}_t,\mathbf{u}_t|\mathbf{a}^{t-1},\mathbf{u}^{t-1})-I(\mathbf{x}^t;\mathbf{u}_t|\mathbf{u}^{t-1}) && \text{(45a)}\\
=&I(\mathbf{x}^t;\mathbf{a}^t|\mathbf{u}^t)-I(\mathbf{x}^t;\mathbf{a}^{t-1}|\mathbf{u}^{t-1}) && \text{(45b)}\\
=&I(\mathbf{x}^t;\mathbf{a}^t|\mathbf{u}^t)-I(\mathbf{x}^{t-1};\mathbf{a}^{t-1}|\mathbf{u}^{t-1})\\
&\qquad -I(\mathbf{x}_t;\mathbf{a}^{t-1}|\mathbf{x}^{t-1},\mathbf{u}^{t-1}) && \text{(45c)}\\
=&I(\mathbf{x}^t;\mathbf{a}^t|\mathbf{u}^t)-I(\mathbf{x}^{t-1};\mathbf{a}^{t-1}|\mathbf{u}^{t-1}). && \text{(45d)}
\end{aligned}$$

When $t=1$, the above identity is understood to mean $I(\mathbf{x}_1;\mathbf{a}_1)-I(\mathbf{x}_1;\mathbf{u}_1)=I(\mathbf{x}_1;\mathbf{a}_1|\mathbf{u}_1)$ which clearly holds as $\mathbf{x}_1$–$\mathbf{a}_1$–$\mathbf{u}_1$ form a Markov chain. Equation (45a) holds because $I(\mathbf{x}^t;\mathbf{a}_t,\mathbf{u}_t|\mathbf{a}^{t-1},\mathbf{u}^{t-1})=I(\mathbf{x}^t;\mathbf{a}_t|\mathbf{a}^{t-1},\mathbf{u}^{t-1})+I(\mathbf{x}^t;\mathbf{u}_t|\mathbf{a}^t,\mathbf{u}^{t-1})$ and the second term is zero since $\mathbf{x}^t$–$(\mathbf{a}^t,\mathbf{u}^{t-1})$–$\mathbf{u}_t$ form a Markov chain. Equation (45b) is obtained by applying the chain rule for mutual information in two different ways:

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

The chain rule is applied again in step (45c). Finally, (45d) follows as $\mathbf{a}^{t-1}$–$(\mathbf{x}^{t-1},\mathbf{u}^{t-1})$–$\mathbf{x}_t$ form a Markov chain.

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

$$\begin{aligned}
&\sum\nolimits_{t=1}^{T}\left[I(\mathbf{x}^t;\mathbf{a}_t|\mathbf{a}^{t-1},\mathbf{u}^{t-1})-I(\mathbf{x}^t;\mathbf{u}_t|\mathbf{u}^{t-1})\right]\\
=&\sum\nolimits_{t=1}^{T}\left[I(\mathbf{x}^t;\mathbf{a}^t|\mathbf{u}^t)-I(\mathbf{x}^{t-1};\mathbf{a}^{t-1}|\mathbf{u}^{t-1})\right] && \text{(46a)}\\
=&I(\mathbf{x}^T;\mathbf{a}^T|\mathbf{u}^T)\geq 0. && \text{(46b)}
\end{aligned}$$

In step (46a), the identity (45) is used. The telescoping sum (46a) cancels all but the final term (46b).

B. Some basic lemmas for probability measures

Lemma 4: Let $\mathbb{P}_{\mathbf{x},\mathbf{y}}$ be a joint probability measure on $(\mathcal{X}\times\mathcal{Y},\mathcal{B}_{\mathcal{X}}\otimes\mathcal{B}_{\mathcal{Y}})$. Let $\mathbb{P}_{\mathbf{x}}$ and $\mathbb{P}_{\mathbf{y}}$ be the marginal probability measures, $\mathbb{P}_{\mathbf{x}}\otimes\mathbb{P}_{\mathbf{y}}$ be the product measure, and $\mathbb{P}_{\mathbf{x}|y}$ be a Borel measurable stochastic kernel such that

$$\mathbb{P}_{\mathbf{x},\mathbf{y}}(B_X\times B_Y)=\int_{B_Y}\mathbb{P}_{\mathbf{x}|y}(B_X|y)\mathbb{P}_{\mathbf{y}}(\mathrm{d}y) \qquad (47)$$

for every $B_X\in\mathcal{B}_{\mathcal{X}}$ and $B_Y\in\mathcal{B}_{\mathcal{Y}}$. If $\mathbb{P}_{\mathbf{x},\mathbf{y}}\ll\mathbb{P}_{\mathbf{x}}\otimes\mathbb{P}_{\mathbf{y}}$, then $\mathbb{P}_{\mathbf{x}|y}\ll\mathbb{P}_{\mathbf{x}},\mathbb{P}_{\mathbf{y}}-a.e.$, and

$$\frac{\mathrm{d}\mathbb{P}_{\mathbf{x},\mathbf{y}}}{\mathrm{d}(\mathbb{P}_{\mathbf{x}}\otimes\mathbb{P}_{\mathbf{y}})}=\frac{\mathrm{d}\mathbb{P}_{\mathbf{x}|y}}{\mathrm{d}\mathbb{P}_{\mathbf{x}}},\mathbb{P}_{\mathbf{y}}-a.e..$$

Proof: Suppose $\mathbb{P}_{\mathbf{x},\mathbf{y}}\ll\mathbb{P}_{\mathbf{x}}\otimes\mathbb{P}_{\mathbf{y}}$, and let $f(x,y)=\frac{\mathrm{d}\mathbb{P}_{\mathbf{x},\mathbf{y}}}{\mathrm{d}(\mathbb{P}_{\mathbf{x}}\otimes\mathbb{P}_{\mathbf{y}})}$ be the Radon-Nikodym derivative. For every $B_X\in\mathcal{B}_{\mathcal{X}}$ and $B_Y\in\mathcal{B}_{\mathcal{Y}}$, we have

$$\begin{aligned}
\mathbb{P}(B_X\times B_Y)&=\int_{B_X\times B_Y}f(x,y)\mathrm{d}(\mathbb{P}_{\mathbf{x}}\otimes\mathbb{P}_{\mathbf{y}})(x,y)\\
&=\int_{B_Y}\left[\int_{B_X}f(x,y)\mathrm{d}\mathbb{P}_{\mathbf{x}}(x)\right]\mathrm{d}\mathbb{P}_{\mathbf{y}}(y) \qquad (48)
\end{aligned}$$

The first line is by definition of the Radon-Nikodym derivative. The second equality holds due to the Fubini's theorem [57], since clearly $f\in L^1(\mathbb{P}_{\mathbf{x}}\otimes\mathbb{P}_{\mathbf{y}})$. Comparing (47) and (48), we have

$$\mathbb{P}_{\mathbf{x}|y}(B_X|y)=\int_{B_X}f(x,y)\mathrm{d}\mathbb{P}_{\mathbf{x}}(x),\mathbb{P}_{\mathbf{y}}-a.e.. \qquad (49)$$

It follows from (49) that $\mathbb{P}_{\mathbf{x}}(B_X)=0\Rightarrow\mathbb{P}_{\mathbf{x}|y}(B_X|y)=0$ holds $\mathbb{P}_{\mathbf{y}}-a.e.$. Also, (49) implies $f(x,y)=\frac{\mathrm{d}\mathbb{P}_{\mathbf{x}|y}}{\mathrm{d}\mathbb{P}_{\mathbf{x}}},\mathbb{P}_{\mathbf{y}}-a.e.$. ■

Lemma 5: Let $\mathbb{P}_{\mathbf{x},\mathbf{y},\mathbf{z}}$ be a zero-mean Borel probability measure on $\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}$, where $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ are Euclidean spaces. Suppose $\mathbb{P}_{\mathbf{x},\mathbf{y},\mathbf{z}}$ has a covariance matrix $\Sigma_{x,y,z}$, and there exists a matrix L such that $\mathbf{z}-L\mathbf{y}$ is independent of $\mathbf{x}$ and $\mathbf{y}$ on $\mathbb{P}_{\mathbf{x},\mathbf{y},\mathbf{z}}$. (This implies $\mathbf{x}$–$\mathbf{y}$–$\mathbf{z}$ form a Markov chain on $\mathbb{P}_{\mathbf{x},\mathbf{y},\mathbf{z}}$.) Let $\mathbb{G}_{\mathbf{x},\mathbf{y},\mathbf{z}}$ be a zero-mean, jointly Gaussian probability measure with the same covariance matrix $\Sigma_{x,y,z}$. Then $\mathbf{x}$–$\mathbf{y}$–$\mathbf{z}$ form a Markov chain on $\mathbb{G}_{\mathbf{x},\mathbf{y},\mathbf{z}}$.

Proof: See [58, Lemma 3.2]. ■

Lemma 6: Let $\mathbf{x}$ be an $(\mathbb{R}^n,\mathcal{B}_{\mathbb{R}^n})$-valued zero mean random variable with covariance $\Sigma_x\succeq 0$. Define an $(\mathbb{R}^m,\mathcal{B}_{\mathbb{R}^m})$-valued random variable $\mathbf{y}$ by $\mathbf{y}=A\mathbf{x}+\mathbf{v}$ where A is a matrix and $\mathbf{v}\sim\mathcal{N}(0,\Sigma_v)$ is a random variable independent of $\mathbf{x}$. Let $(\mathbf{x}_G,\mathbf{y}_G)$ be zero-mean, jointly Gaussian random variables, and suppose that $(\mathbf{x},\mathbf{y})$ and $(\mathbf{x}_G,\mathbf{y}_G)$ have the same covariance matrix. Then $\mathbf{y}_G$ can be written as $\mathbf{y}_G=A\mathbf{x}_G+\mathbf{v}$ with $\mathbf{v}\sim\mathcal{N}(0,\Sigma_v)$ independent of $\mathbf{x}_G$.

Proof: Observe

$$\operatorname{cov}\begin{bmatrix}\mathbf{x}\\ \mathbf{y}\end{bmatrix}=\begin{bmatrix}\Sigma_x & \Sigma_x A^\top\\ A\Sigma_x & A\Sigma_x A^\top+\Sigma_v\end{bmatrix}. \qquad (50)$$

Since it must be that $\mathbf{x}_G\sim\mathcal{N}(0,\Sigma_x)$, introducing a matrix R with full column rank such that $\Sigma_x=RR^\top$, $\mathbf{x}_G$ can be written as $\mathbf{x}_G=R\mathbf{z}_G$ with $\mathbf{z}_G\sim\mathcal{N}(0,I)$. Since $(\mathbf{x}_G,\mathbf{y}_G)$ are jointly Gaussian, there exists a matrix $\bar{A}$ such that

$$\mathbf{y}_G=\bar{A}\mathbf{x}_G+\mathbf{v},\;\;\mathbf{v}\sim\mathcal{N}(0,\bar{\Sigma}_v) \qquad (51)$$

where $\bar{\mathbf{v}}$ is independent of $\mathbf{x}_G$. Thus

$$\operatorname{cov}\begin{bmatrix}\mathbf{x}_G\\ \mathbf{y}_G\end{bmatrix}=\begin{bmatrix}\Sigma_x & \Sigma_x\bar{A}^\top\\ \bar{A}\Sigma_x & \bar{A}\Sigma_x\bar{A}^\top+\bar{\Sigma}_v\end{bmatrix}. \qquad (52)$$

By comparing (50) and (52), it can be seen that $\bar{A}=A+S$ with S satisfying $SR=0$, and $\bar{\Sigma}_v=\Sigma_v$. Then from (51),

$$\begin{aligned}
\mathbf{y}_G&=(A+S)\mathbf{x}_G+\mathbf{v}\\
&=(A+S)R\mathbf{z}_G+\mathbf{v}\\
&=AR\mathbf{z}_G+\mathbf{v}\\
&=A\mathbf{x}_G+\mathbf{v},\;\;\mathbf{v}\sim\mathcal{N}(0,\Sigma_v).
\end{aligned}$$

■

Lemma 7: Let $\mathbb{P}_{\mathbf{x},\mathbf{y}}$ be a zero-mean joint probability measure on $(\mathcal{X}\times\mathcal{Y},\mathcal{B}_{\mathcal{X}}\otimes\mathcal{B}_{\mathcal{Y}})$, $\mathcal{X}=\mathbb{R}^n$, $\mathcal{Y}=\mathbb{R}^m$, with a covariance matrix $\Sigma_{x,y}$. Let $\mathbb{G}_{\mathbf{x},\mathbf{y}}$ be a zero-mean Gaussian joint probability measure with the same covariance matrix $\Sigma_{x,y}$. If there exists a subset $\mathcal{C}_y \subseteq \mathbb{R}^m$ with $\mathbb{P}_{\mathbf{y}}(\mathcal{C}_y)>0$ such that $\mathbb{P}_{\mathbf{x}|y}$ admits density for every $y\in\mathcal{C}_y$, then $\mathbb{G}_{\mathbf{x}|y}$ admits density for every $y\in\text{supp}(\mathbb{G}_{\mathbf{y}})$.

Proof: Suppose $\mathbb{P}_{\mathbf{x},\mathbf{y}}(x,y)$ and $\mathbb{G}_{\mathbf{x},\mathbf{y}}(x,y)$ share a covariance matrix

$$\Sigma_{x,y}=\begin{bmatrix}\Sigma_{xx} & \Sigma_{xy}\\ \Sigma_{yx} & \Sigma_{yy}\end{bmatrix}\succeq 0.$$

Since $\mathbb{G}_{\mathbf{x},\mathbf{y}}$ is a zero-mean Gaussian distribution, we know

$$\mathbb{G}_{\mathbf{x}|y}\sim\mathcal{N}(\hat{x}(y),\Sigma_e)\ \ \forall y\in\mathcal{Y} \tag{53}$$

with $\hat{x}(y)=\Sigma_{xy}\Sigma_{yy}^{\dagger}y$, $\Sigma_e=\Sigma_{xx}-\Sigma_{xy}\Sigma_{yy}^{\dagger}\Sigma_{yx}$, where $\Sigma_{yy}^{\dagger}$ is the Moore–Penrose pseudoinverse of Σ_{yy}. To show contrapositive, assume that there exists $y\in\text{supp}(\mathbb{G}_{\mathbf{y}})$ such that $\mathbb{G}_{\mathbf{x}|y}(x|y)$ does not admit a density. From (53), this means that Σ_e is a singular matrix. For every $y\in\mathcal{Y}$, define a covariance matrix $M(y)\triangleq\int_{\mathcal{X}}(x-\hat{x}(y))(x-\hat{x}(y))^\top \mathrm{d}\mathbb{P}_{\mathbf{x}|y}(x|y)$. Observe that

$$\begin{aligned}\int_{\mathcal{Y}}M(y)\mathrm{d}\mathbb{P}(y)&=\int_{\mathcal{X},\mathcal{Y}}(x-\hat{x}(y))(x-\hat{x}(y))^\top\mathrm{d}\mathbb{P}(x,y)\\&=\int_{\mathcal{X},\mathcal{Y}}(x-\hat{x}(y))(x-\hat{x}(y))^\top\mathrm{d}\mathbb{G}(x,y)\\&=\Sigma_e.\end{aligned}\tag{54}$$

Since Σ_e is singular, there exists a full row rank matrix $U\in\mathbb{R}^{r\times n}$, $1\le r\le n$ such that $U\Sigma_eU^\top=0$. From (54), it follows that $UM(y)U^\top=0,\mathbb{P}_{\mathbf{y}}-a.e.$. For every $y\in\mathcal{Y}$, define a subset $\mathcal{C}_{x|y}\subset\mathbb{R}^n$ by

$$\mathcal{C}_{x|y}=\{x\in\mathbb{R}^n: U(x-\hat{x}(y))=0\}.$$

By construction, $\mathbb{P}_{\mathbf{x}|y}(\mathcal{C}_{x|y}|y)=1,\quad \mathbb{P}_{\mathbf{y}}-a.e.$. However, clearly $\mathbb{L}_{\mathbf{x}}(\mathcal{C}_{x|y})=0$, where $\mathbb{L}_{\mathbf{x}}(x)$ is the Lebesgue measure on $\mathbb{R}^n$. Thus $\mathbb{P}_{\mathbf{x}|y}\ll\mathbb{L}_{\mathbf{x}}$ fails to hold $\mathbb{P}_{\mathbf{y}}-a.e.$. Hence, $\mathbb{P}_{\mathbf{x}|y}$ fails to admit a density $\mathbb{P}_{\mathbf{y}}-a.e.$. This is a contradiction to the assumption that there exists a subset $\mathcal{C}_y\subseteq\mathbb{R}^m$ with $\mathbb{P}_{\mathbf{y}}(\mathcal{C}_y)>0$ such that $\mathbb{P}_{\mathbf{x}|y}$ admits density for every $y\in\mathcal{C}_y$. ■

C. *Proof of Lemma 2*

Lemma 8: Let $\mathbb{P}$ be a zero-mean Borel probability measure on $\mathbb{R}^n$ with covariance matrix Σ. Suppose $\mathbb{G}$ is a zero-mean Gaussian probability measure on $\mathbb{R}^n$ with the same covariance matrix Σ. Then $\text{supp}(\mathbb{P})\subseteq\text{supp}(\mathbb{G})$.

Proof: When Σ is positive definite, the claim is trivial since $\text{supp}(\mathbb{G})=\mathbb{R}^n$. So assume that Σ is singular. Then, there exists an orthonormal matrix

$$U=\begin{bmatrix}U_1\\U_2\end{bmatrix}\ \text{with}\ U_1\in\mathbb{R}^{p\times n},\ U_2\in\mathbb{R}^{(n-p)\times n},\ p<n$$

such that $U\Sigma U^\top=\text{diag}(\Sigma_{zz},0)$, where $\Sigma_{zz}\in\mathbb{R}^{p\times p}$ is a positive definite matrix. Notice that $U_2x=0$ for every $x\in\text{supp}(\mathbb{G})$, and $U_2x\neq 0$ for every $x\in\text{supp}(\mathbb{G})^c$. Suppose $\text{supp}(\mathbb{P})\subseteq\text{supp}(\mathbb{G})$ does not hold, i.e., there exists a closed set $\mathcal{C}\subseteq\text{supp}(\mathbb{G})^c$ such that $\mathbb{P}(\mathcal{C})>0$. Then

$$\begin{aligned}U_2\Sigma U_2^\top&=\int U_2xx^\top U_2^\top d\mathbb{P}(x)\\&=\int_{\text{supp}(\mathbb{G})}U_2xx^\top U_2^\top d\mathbb{P}(x)+\int_{\text{supp}(\mathbb{G})^c}U_2xx^\top U_2^\top d\mathbb{P}(x)\\&=\int_{\text{supp}(\mathbb{G})^c}U_2xx^\top U_2^\top d\mathbb{P}(x)\succeq\int_{\mathcal{C}}U_2xx^\top U_2^\top d\mathbb{P}(x)\end{aligned}$$

Since $U_2x\neq 0$ for every $x\in\mathcal{C}$, the last expression is a non-zero positive semidefinite matrix. However, by construction, we have $U_2\Sigma U_2^\top=0$. Thus, the above inequality leads to a contradiction. ■

Lemma 9: Let $\mathbb{P}(x^{T+1},u^T)$ be a joint probability measure generated by a policy $\gamma_{\mathbb{P}}=\{\mathbb{P}(u_t|x^t,u^{t-1})\}_{t=1}^T$ and (1).

(a) For each $t=1,\cdots,T$, $\mathbb{P}(x_{t+1}|u^t)$ and $\mathbb{P}(x_{t+1}|x_t,u^t)$ are non-degenerate Gaussian probability measures for every x_t and u^t.

Moreover, if $I_{\mathbb{P}}(\mathbf{x}_t;\mathbf{u}_t|\mathbf{u}^{t-1})<+\infty$ for all $t=1,\cdots,T$, then the following statements hold.

(b) For every $t=1,\cdots,T$,

$$\mathbb{P}(x_t|u^t)\ll\mathbb{P}(x_t|u^{t-1}),\ \ \mathbb{P}(u^t)-a.e.,\ \text{and}$$
$$I_{\mathbb{P}}(\mathbf{x}_t;\mathbf{u}_t|\mathbf{u}^{t-1})=\int\log\left(\frac{\mathrm{d}\mathbb{P}(x_t|u^t)}{\mathrm{d}\mathbb{P}(x_t|u^{t-1})}\right)\mathrm{d}\mathbb{P}(x_t,u^t).$$

(c) For every $t=1,\cdots,T$,

$$\mathbb{P}(x_t|x_{t+1},u^t)\ll\mathbb{P}(x_t|u^{t-1}),\ \ \mathbb{P}(x_{t+1},u^t)-a.e..$$

Moreover, the following identity holds $\mathbb{P}(x_{t+1},u^t)-a.e.$:

$$\frac{\mathrm{d}\mathbb{P}(x_t|u^t)}{\mathrm{d}\mathbb{P}(x_t|u^{t-1})}=\frac{\mathrm{d}\mathbb{P}(x_{t+1}|u^t)}{\mathrm{d}\mathbb{P}(x_{t+1}|x_t,u^t)}\frac{\mathrm{d}\mathbb{P}(x_t|x_{t+1},u^t)}{\mathrm{d}\mathbb{P}(x_t|u^{t-1})}.\tag{55}$$

Proof:

(a) This is clear since $\mathbb{P}(x^{T+1},u^T)$ is constructed using (1).

(b) By definition of conditional mutual information, $I_{\mathbb{P}}(\mathbf{x}_t;\mathbf{u}_t|\mathbf{u}^{t-1})<+\infty$ requires $I_{\mathbb{P}}(\mathbf{x}_t;\mathbf{u}_t|u^{t-1})<+\infty,\mathbb{P}_{\mathbf{u}^{t-1}}-a.e.$. For a fixed u^{t-1}, $I_{\mathbb{P}}(\mathbf{x}_t;\mathbf{u}_t|u^{t-1})<+\infty$ requires $\mathbb{P}_{\mathbf{x}_t,\mathbf{u}_t|u^{t-1}}\ll\mathbb{P}_{\mathbf{x}_t|u^{t-1}}\otimes\mathbb{P}_{\mathbf{u}_t|u^{t-1}}$ by definition of mutual information. By Lemma 4, this implies

$$\mathbb{P}_{\mathbf{x}_t|u^t}\ll\mathbb{P}_{\mathbf{x}_t|u^{t-1}}\ \text{and}\tag{56}$$
$$\frac{\mathrm{d}\mathbb{P}_{\mathbf{x}_t,\mathbf{u}_t|u^{t-1}}}{\mathrm{d}(\mathbb{P}_{\mathbf{x}_t|u^{t-1}}\otimes\mathbb{P}_{\mathbf{u}_t|u^{t-1}})}=\frac{\mathrm{d}\mathbb{P}_{\mathbf{x}_t|u^t}}{\mathrm{d}\mathbb{P}_{\mathbf{x}_t|u^{t-1}}}\tag{57}$$

must hold $\mathbb{P}(u_t|u^{t-1})-a.e.$. Since this is the case $\mathbb{P}(u^{t-1})-a.e.$, we have both (56) and (57) $\mathbb{P}(u^t)-a.e.$. Hence

$$\begin{aligned}&I_{\mathbb{P}}(\mathbf{x}_t;\mathbf{u}_t|\mathbf{u}^{t-1})\\&=\int\log\left(\frac{\mathrm{d}\mathbb{P}_{\mathbf{x}_t,\mathbf{u}_t|u^{t-1}}}{\mathrm{d}(\mathbb{P}_{\mathbf{x}_t|u^{t-1}}\otimes\mathbb{P}_{\mathbf{u}_t|u^{t-1}})}\right)\mathrm{d}\mathbb{P}(x_t,u^t)\\&=\int\log\left(\frac{\mathrm{d}\mathbb{P}_{\mathbf{x}_t|u^t}}{\mathrm{d}\mathbb{P}_{\mathbf{x}_t|u^{t-1}}}\right)\mathrm{d}\mathbb{P}(x_t,u^t).\end{aligned}$$

(c) Let $B_{X_t}\in\mathcal{B}_{\mathcal{X}_t}$, $B_{X_{t+1}}\in\mathcal{B}_{\mathcal{X}_{t+1}}$, $B_{U^t}\in\mathcal{B}_{\mathcal{U}^t}$ be arbitrary Borel sets. Since both $\mathbb{P}(x_{t+1}|u^t)$ and $\mathbb{P}(x_{t+1}|x_t,u^t)$ are non-degenerate Gaussian probability measures, there exists a continuous map $f:\mathcal{X}_t\times\mathcal{X}_{t+1}\times\mathcal{U}^t\to(0,+\infty)$ such that

$$k(x_t, x_{t+1}, u^t) = \frac{\mathrm{d}\mathbb{P}(x_{t+1}|u^t)}{\mathrm{d}\mathbb{P}(x_{t+1}|x_t, u^t)}, \text{ or} \tag{58a}$$

$$\int_{B_{X_{t+1}}} k(x_t, x_{t+1}, u^t)\mathrm{d}\mathbb{P}(x_{t+1}|x_t, u^t) = \int_{B_{X_{t+1}}} \mathrm{d}\mathbb{P}(x_{t+1}|u^t). \tag{58b}$$

In what follows, we suppress the arguments of $k(x_t, x_{t+1}, u^t)$ and simply write it as k. Next, we express

$$\int_{B_{X_t}\times B_{X_{t+1}}\times B_{U^t}} k\mathrm{d}\mathbb{P}(x_t, x_{t+1}, u^t) \tag{59}$$

in two different ways:

$$(59) = \int_{B_{X_{t+1}}\times B_{U^t}} \int_{B_{X_t}} k\mathrm{d}\mathbb{P}(x_t|x_{t+1}, u^t)\mathrm{d}\mathbb{P}(x_{t+1}, u^t); \tag{60a}$$

$$(59) = \int_{B_{X_t}\times B_{U^t}} \int_{B_{X_{t+1}}} k\mathrm{d}\mathbb{P}(x_{t+1}|x_t, u^t)\mathrm{d}\mathbb{P}(x_t, u^t) \tag{60b}$$

$$= \int_{B_{X_t}\times B_{U^t}} \int_{B_{X_{t+1}}} \mathrm{d}\mathbb{P}(x_{t+1}|u^t)\mathrm{d}\mathbb{P}(x_t, u^t) \tag{60c}$$

$$= \int_{B_{U^t}} \int_{B_{X_t}} \int_{B_{X_{t+1}}} \mathrm{d}\mathbb{P}(x_{t+1}|u^t)\mathrm{d}\mathbb{P}(x_t|u^t)\mathrm{d}\mathbb{P}(u^t) \tag{60d}$$

$$= \int_{B_{X_{t+1}}\times B_{U^t}} \int_{B_{X_t}} \mathrm{d}\mathbb{P}(x_t|u^t)\mathrm{d}\mathbb{P}(x_{t+1}, u^t) \tag{60e}$$

$$= \int_{B_{X_{t+1}}\times B_{U^t}} \mathbb{P}_{\mathbf{x}_t|u^t}(B_{X_t}|u^t)\mathrm{d}\mathbb{P}(x_{t+1}, u^t). \tag{60f}$$

Notice that (58b) is used in step (60c). Comparing (60a) and (60f), we have the following identity $\mathbb{P}(x_{t+1}, u^t) - a.e.$:

$$\int_{B_{X_t}} k\mathrm{d}\mathbb{P}(x_t|x_{t+1}, u^t) = \mathbb{P}_{\mathbf{x}_t|u^t}(B_{X_t}|u^t). \tag{61}$$

Since $k(x_t, x_{t+1}, u^t)$ assumes values in $(0, +\infty)$, the first claim follows from (61).

Now, the next equalities hold $\mathbb{P}(x_{t+1}, u^t) - a.e.$, which establishes the second claim.

$$\int_{B_{X_t}} \frac{\mathrm{d}\mathbb{P}(x_{t+1}|u^t)}{\mathrm{d}\mathbb{P}(x_{t+1}|x_t, u^t)} \frac{\mathrm{d}\mathbb{P}(x_t|x_{t+1}, u^t)}{\mathrm{d}\mathbb{P}(x_t|u^{t-1})} \mathrm{d}\mathbb{P}(x_t|u^{t-1})$$
$$= \int_{B_{X_t}} k(x_t, x_{t+1}, u^t)\mathrm{d}\mathbb{P}(x_t|x_{t+1}, u^t)$$
$$= \mathbb{P}_{\mathbf{x}_t|u^t}(B_{X_t}|u^t).$$

The identity (58a) is used in the first step, and (61) is used in the second step. ∎

Lemma 10: Let $\mathbb{P}(x^{T+1}, u^T)$ be a joint probability measure generated by a policy $\gamma_{\mathbb{P}} = \{\mathbb{P}(u_t|x^t, u^{t-1})\}_{t=1}^T$ and (1), and $\mathbb{G}(x^{T+1}, u^T)$ be a zero-mean jointly Gaussian probability measure having the same covariance as $\mathbb{P}(x^{T+1}, u^T)$. For every $t = 1, \cdots, T$, we have

(a) $\mathbf{u}^{t-1}$–$(\mathbf{x}_t, \mathbf{u}_t)$–$\mathbf{x}_{t+1}$ form a Markov chain in $\mathbb{G}$. Moreover, for every $t = 1, \cdots, T$, we have

$$\mathbb{G}(x_{t+1}|x_t, u^t) = \mathbb{G}(x_{t+1}|x_t, u_t)$$
$$= \mathbb{P}(x_{t+1}|x_t, u_t)$$
$$= \mathbb{P}(x_{t+1}|x_t, u^t)$$

all of which have a nondegenerate Gaussian distribution $\mathcal{N}(A_t x_t + B_t u_t, W_t)$.

(b) For each $t = 1, \cdots, T$, $\mathbb{G}(x_t|x_{t+1}, u^t)$ is a non-degenerate Gaussian measure for every $(x_{t+1}, u^t) \in \text{supp}(\mathbb{G}(x_{t+1}, u^t))$.

Proof: (a) Since $\mathbf{u}^{t-1}$–$(\mathbf{x}_t, \mathbf{u}_t)$–$\mathbf{x}_{t+1}$ form a Markov chain in $\mathbb{P}$, and $\mathbf{x}_{t+1}$ and $(\mathbf{x}_t, \mathbf{u}_t)$ are related by a linear map (1), by Lemma 5 (borrowed from [58]), $\mathbf{u}^{t-1}$–$(\mathbf{x}_t, \mathbf{u}_t)$–$\mathbf{x}_{t+1}$ form a Markov chain also in $\mathbb{G}$. Notice that $\mathbb{P}(x_{t+1}|x_t, u^t) = \mathbb{P}(x_{t+1}|x_t, u_t)$ and $\mathbb{G}(x_{t+1}|x_t, u^t) = \mathbb{G}(x_{t+1}|x_t, u_t)$ hold since $\mathbf{u}^{t-1}$–$(\mathbf{x}_t, \mathbf{u}_t)$–$\mathbf{x}_{t+1}$ form a Markov chain both in $\mathbb{P}$ and $\mathbb{G}$. Since $\mathbb{P}(x_{t+1}, x_t, u_t)$ and $\mathbb{G}(x_{t+1}, x_t, u_t)$ have the same covariance, by Lemma 6, a linear relationship (1) holds both in $\mathbb{P}$ and $\mathbb{G}$. Thus $\mathbb{G}(x_{t+1}|x_t, u_t) = \mathbb{P}(x_{t+1}|x_t, u_t)$.

(b) From Lemma 9 (c), $\mathbb{P}(x_t|x_{t+1}, u^t)$ admits a density $\mathbb{P}(x_{t+1}u^t) - a.e.$. Thus, by Lemma 7, $\mathbb{G}(x_t|x_{t+1}, u^t)$ admits density for every $(x_{t+1}, u^t) \in \text{supp}(\mathbb{G}(x_{t+1}, u^t))$. ∎

If the left hand side of (23) is finite, by Lemma 9, it can be written as follows.

$$\sum\nolimits_{t=1}^{T} I_{\mathbb{P}}(\mathbf{x}_t; \mathbf{u}_t|\mathbf{u}^{t-1})$$
$$= \sum\nolimits_{t=1}^{T} \int \log\left(\frac{\mathrm{d}\mathbb{P}(x_t|u^t)}{\mathrm{d}\mathbb{P}(x_t|u^{t-1})}\right) \mathrm{d}\mathbb{P}(x^{T+1}, u^T)$$
$$= \int \log\left(\prod_{t=1}^{T} \frac{\mathrm{d}\mathbb{P}(x_t|u^t)}{\mathrm{d}\mathbb{P}(x_t|u^{t-1})}\right) \mathrm{d}\mathbb{P}(x^{T+1}, u^T)$$
$$= \int \log\left(\prod_{t=1}^{T} \frac{\mathrm{d}\mathbb{P}(x_t|x_{t+1}, u^t)}{\mathrm{d}\mathbb{P}(x_t|u^{t-1})} \frac{\mathrm{d}\mathbb{P}(x_{t+1}|u^t)}{\mathrm{d}\mathbb{P}(x_{t+1}|x_t, u^t)}\right) \mathrm{d}\mathbb{P}(x^{T+1}, u^T)$$

$$= \int \log\left(\frac{\mathrm{d}\mathbb{P}(x_1|x_2, u_1)}{\mathrm{d}\mathbb{P}(x_1)}\right) \mathrm{d}\mathbb{P}(x^2, u_1) \tag{62a}$$

$$+ \sum_{t=2}^{T} \int \log\left(\frac{\mathrm{d}\mathbb{P}(x_t|x_{t+1}, u^t)}{\mathbb{P}(x_t|x_{t-1}, u^{t-1})}\right) \mathrm{d}\mathbb{P}(x^{t+1}, u^t) \tag{62b}$$

$$+ \int \log\left(\frac{\mathrm{d}\mathbb{P}(x_{T+1}|u^T)}{\mathrm{d}\mathbb{P}(x_{T+1}|x_T, u^T)}\right) \mathrm{d}\mathbb{P}(x^{T+1}, u^T) \tag{62c}$$

The result of Lemma 9 (c) is used in the third equality. In the final step, the the chain rule for the Radon-Nikodym derivatives [57, Proposition 3.9] is used multiple times for telescoping cancellations. We show that each term in (62a), (62b) and (62c) does not increase by replacing the probability measure $\mathbb{P}$ with $\mathbb{G}$. Here we only show the case for (62b), but a similar technique is also applicable to (62a) and (62c).

$$\begin{aligned}
&\int \log\left(\frac{\mathrm{d}\mathbb{P}(x_t|x_{t+1},u^t)}{\mathrm{d}\mathbb{P}(x_t|x_{t-1},u^{t-1})}\right)\mathrm{d}\mathbb{P}(x^{t+1},u^t)\\
&\quad-\int \log\left(\frac{\mathrm{d}\mathbb{G}(x_t|x_{t+1},u^t)}{\mathrm{d}\mathbb{G}(x_t|x_{t-1},u^{t-1})}\right)\mathrm{d}\mathbb{G}(x^{t+1},u^t) \qquad (63a)\\
&=\int \log\left(\frac{\mathrm{d}\mathbb{P}(x_t|x_{t+1},u^t)}{\mathrm{d}\mathbb{P}(x_t|x_{t-1},u^{t-1})}\right)\mathrm{d}\mathbb{P}(x^{t+1},u^t)\\
&\quad-\int \log\left(\frac{\mathrm{d}\mathbb{G}(x_t|x_{t+1},u^t)}{\mathrm{d}\mathbb{G}(x_t|x_{t-1},u^{t-1})}\right)\mathrm{d}\mathbb{P}(x^{t+1},u^t) \qquad (63b)\\
&=\int\log\left(\frac{\mathrm{d}\mathbb{P}(x_t|x_{t+1},u^t)}{\mathrm{d}\mathbb{P}(x_t|x_{t-1},u^{t-1})}\frac{\mathrm{d}\mathbb{G}(x_t|x_{t-1},u^{t-1})}{\mathrm{d}\mathbb{G}(x_t|x_{t+1},u^t)}\right)\mathrm{d}\mathbb{P}(x^{t+1},u^t)\\
&=\int \log\left(\frac{\mathrm{d}\mathbb{P}(x_t|x_{t+1},u^t)}{\mathrm{d}\mathbb{G}(x_t|x_{t+1},u^t)}\right)\mathrm{d}\mathbb{P}(x^{t+1},u^t) \qquad (63c)\\
&=\int\left[\int\log\left(\frac{\mathrm{d}\mathbb{P}(x_t|x_{t+1},u^t)}{\mathrm{d}\mathbb{G}(x_t|x_{t+1},u^t)}\right)\mathrm{d}\mathbb{P}(x_t|x_{t+1},u^t)\right]\mathrm{d}\mathbb{P}(x_{t+1},u^t)\\
&=\int D\left(\mathbb{P}(x_t|x_{t+1},u^t)\|\mathbb{G}(x_t|x_{t+1},u^t)\right)\mathrm{d}\mathbb{P}(x_{t+1},u^t)\\
&\geq 0.
\end{aligned}$$

Due to Lemma 10, $\log\frac{\mathrm{d}\mathbb{G}(x_t|x_{t+1},u^t)}{\mathrm{d}\mathbb{G}(x_t|x_{t-1},u^{t-1})}$ in (63a) is a quadratic function of x^{t+1} and u^t everywhere on $\mathrm{supp}(\mathbb{G}(x^{t+1},u^t))$. This is also the case everywhere on $\mathrm{supp}(\mathbb{P}(x^{t+1},u^t))$ since it follows from Lemma 8 that $\mathrm{supp}(\mathbb{P}(x^{t+1},u^t))\subseteq\mathrm{supp}(\mathbb{G}(x^{t+1},u^t))$. Since $\mathbb{P}$ and $\mathbb{G}$ have the same covariance, $\mathrm{d}\mathbb{G}(x^{t+1},u^t)$ can be replaced by $\mathrm{d}\mathbb{P}(x^{t+1},u^t)$ in (63b). In (63c), the chain rule of the Radon-Nikodym derivatives is used invoking that $\mathbb{P}(x_t|x_{t-1},u^{t-1})=\mathbb{G}(x_t|x_{t-1},u^{t-1})$ from Lemma 10 (a).

D. Proof of Lemma 3

Clearly $\mathbb{G}(x_1)=\mathbb{Q}(x_1)$ holds. Following an induction argument, assume that the claim holds for $t=k-1$. Then

$$\begin{aligned}
&\mathrm{d}\mathbb{Q}(x_{k+1},u^k)\\
&=\int_{\mathcal{X}_k}\mathrm{d}\mathbb{Q}(x_k,x_{k+1},u^k)\\
&=\int_{\mathcal{X}_k}\mathrm{d}\mathbb{P}(x_{k+1}|x_k,u_k)\mathrm{d}\mathbb{Q}(x_k,u^k) \qquad (64a)\\
&=\int_{\mathcal{X}_k}\mathrm{d}\mathbb{P}(x_{k+1}|x_k,u_k)\mathrm{d}\mathbb{Q}(u_k|x_k,u^{k-1})\mathrm{d}\mathbb{Q}(x_k,u^{k-1}) \qquad (64b)\\
&=\int_{\mathcal{X}_k}\mathrm{d}\mathbb{P}(x_{k+1}|x_k,u_k)\mathrm{d}\mathbb{Q}(u_k|x_k,u^{k-1})\mathrm{d}\mathbb{G}(x_k,u^{k-1}) \qquad (64c)\\
&=\int_{\mathcal{X}_k}\mathrm{d}\mathbb{P}(x_{k+1}|x_k,u_k)\mathrm{d}\mathbb{G}(x_k,u^k) \qquad (64d)\\
&=\int_{\mathcal{X}_k}\mathrm{d}\mathbb{G}(x_k,x_{k+1},u^k) \qquad (64e)\\
&=\mathrm{d}\mathbb{G}(x_{k+1},u^k).
\end{aligned}$$

The integral signs "$\int_{B_{X_{k+1}}\times B_{U^k}}$" in front of each of the above expressions are omitted for simplicity. Equations (64a) and (64b) are due to (25) and (26) respectively. In (64c), the induction assumption $\mathbb{G}(x_k,u^{k-1})=\mathbb{Q}(x_k,u^{k-1})$ is used. Identity (64d) follows from the definition (24). The result of Lemma 10(b) was used in (64e).

E. Proof of Theorem 2 (Outline only)

First, it can be shown that the three-stage separation principle continues to hold for the infinite horizon problem (6). The same idea of proof as in Section VI is applicable; for every policy $\gamma_{\mathbb{P}}=\{\mathbb{P}(u_t|x^t,u^{t-1})\}_{t\in\mathbb{N}}$, there exists a linear-Gaussian policy $\gamma_{\mathbb{Q}}=\{\mathbb{Q}(u_t|x^t,u^{t-1})\}_{t\in\mathbb{N}}$ which is at least as good as $\gamma_{\mathbb{P}}$. Second, the optimal certainty equivalence controller gain is time-invariant. This is because, since (A,B) is stabilizable, for every finite t, the solution S_t of the Riccati recursion (11) converges to the solution S of (17) as $T\to\infty$ [59, Theorem 14.5.3]. Third, the optimal AWGN channel design problem becomes an SDP over an infinite sequence $\{P_{t|t},\Pi_t\}_{t\in\mathbb{N}}$ similar to (12) in which "$\sum_{t=1}^T$" is replaced by "$\limsup_{T\to\infty}\frac{1}{T}\sum_{t=1}^T$" and parameters A_t,W_t,S_t,Θ_t are time-invariant. It is shown in [60] that the optimality of this SDP over $\{P_{t|t},\Pi_t\}_{t\in\mathbb{N}}$ is attained by a time-invariant sequence $P_{t|t}=P,\Pi_t=\Pi\ \forall t\in\mathbb{N}$, where P and Π are the optimal solution to (18).

F. Proof of Corollary 1

We write $v^*(A,W)\triangleq\lim_{D\to+\infty}R(D)$ to indicate its dependency on A and W. From (18), we have

$$v^*(A,W)= \qquad (65)$$
$$\begin{cases}\inf\limits_{P,\Pi} & \frac{1}{2}\log\det\Pi^{-1}+\frac{1}{2}\log\det W\\ \text{s.t.} & \Pi\succ 0, P\preceq APA^\top+W, \begin{bmatrix}P-\Pi & PA^\top\\ AP & APA^\top+W\end{bmatrix}\succeq 0.\end{cases}$$

Due to the strict feasibility, Slater's constraint qualification [61] guarantees that the duality gap is zero. Thus, we have an alternative representation of $v^*(A,W)$ using the dual problem of (65).

$$v^*(A,W)= \qquad (66)$$
$$\begin{cases}\sup\limits_{X,Y} & \frac{1}{2}\log\det X_{11}-\frac{1}{2}\mathrm{Tr}(X_{22}+Y)W+\frac{1}{2}\log\det W+\frac{n}{2}\\ \text{s.t.} & A^\top YA-Y+X_{11}+X_{12}A+A^\top X_{21}+A^\top X_{22}A\preceq 0,\\ & Y\succeq 0, X=\begin{bmatrix}X_{11} & X_{12}\\ X_{21} & X_{22}\end{bmatrix}\succeq 0.\end{cases}$$

The primal problem (65) can be also rewritten as

$$\begin{aligned}
&v^*(A,W)\\
&=\begin{cases}\inf\limits_{P} & \frac{1}{2}\log\det(APA^\top+W)-\frac{1}{2}\log\det P\\ \text{s.t.} & P\preceq APA^\top+W, P\in\mathbb{S}^n_{++}\end{cases} \qquad (67)\\
&=\begin{cases}\inf\limits_{P,C,V} & -\frac{1}{2}\log\det(I-V^{-\frac{1}{2}}CPC^\top V^{-\frac{1}{2}})\\ \text{s.t.} & P^{-1}-(APA^\top+W)^{-1}=C^\top V^{-1}C\\ & P\in\mathbb{S}^n_{++}, V\in\mathbb{S}^n_{++}, C\in\mathbb{R}^{n\times n}.\end{cases} \qquad (68)
\end{aligned}$$

To see that (67) and (68) are equivalent, note that the feasible set of P in (67) and (68) are the same. Also

$$\begin{aligned}
&\tfrac{1}{2}\log\det(APA^\top+W)-\tfrac{1}{2}\log\det P\\
&=-\tfrac{1}{2}\log\det(APA^\top+W)^{-1}-\tfrac{1}{2}\log\det P\\
&=-\tfrac{1}{2}\log\det(P^{-1}-C^\top V^{-1}C)-\tfrac{1}{2}\log\det P\\
&=-\tfrac{1}{2}\log\det(I-P^{\frac{1}{2}}C^\top V^{-1}CP^{\frac{1}{2}})\\
&=-\tfrac{1}{2}\log\det(I-V^{-\frac{1}{2}}CPC^\top V^{-\frac{1}{2}})
\end{aligned}$$

The last step follows from Sylvester's determinant theorem.

1) Case 1: When all eigenvalues of A satisfy $|\lambda_i| \geq 1$: We first show that if all eigenvalues of A are outside the open unit disc, then $v^*(A, W) = \sum_{\lambda_i \in \sigma(A)} \log|\lambda_i|$, where $\sigma(A)$ is the set of all eigenvalues of A counted with multiplicity. To see that $v^*(A, W) \leq \sum_{\lambda_i \in \sigma(A)} \log|\lambda_i|$, note that the value $\sum_{\lambda_i \in \sigma(A)} \log|\lambda_i| + \epsilon$ with arbitrarily small $\epsilon > 0$ can be attained by $P = kI$ in (67) with sufficiently large $k > 0$. To see that $v^*(A, W) \geq \sum_{\lambda_i \in \sigma(A)} \log|\lambda_i|$, note that the value $\sum_{\lambda_i \in \sigma(A)} \log|\lambda_i|$ is attained by the dual problem (66) with $X = [A \quad -I]^\top W^{-1}[A \quad -I]$ and $Y = 0$.

2) Case 2: When all eigenvalues of A satisfy $|\lambda_i| < 1$: In this case, we have $v^*(A, W) = 0$. The fact that $v^*(A, W) \geq 0$ is immediate from the expression (67). To see that $v^*(A, W) = 0$, consider $P = P^*$ in (67) where $P^* \succ 0$ is the unique solution to the Lyapunov equation $P^* = AP^*A^\top + W$.

3) Case 3: General case: In what follows, we assume without loss of generality that A has a structure (e.g., a Jordan form)

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

where all eigenvalues of $A_1 \in \mathbb{R}^{n_1 \times n_1}$ satisfy $|\lambda_i| \geq 1$ and all eigenvalues of $A_2 \in \mathbb{R}^{n_2 \times n_2}$ satisfy $|\lambda_i| < 1$. We first recall the following basic property of the algebraic Riccati equation.

Lemma 11: Suppose $V \succ 0$ and (A, C) is a detectable pair and $0 \prec W_1 \preceq W_2$. Then, we have $\tilde{P} \preceq \tilde{Q}$ where $\tilde{P}$ and $\tilde{Q}$ are the unique positive definite solutions to

$$A\tilde{P}A^\top - \tilde{P} - A\tilde{P}C^\top(C\tilde{P}C^\top + V)^{-1}C\tilde{P}A^\top + W_1 = 0 \quad (69)$$

$$A\tilde{Q}A^\top - \tilde{Q} - A\tilde{Q}C^\top(C\tilde{Q}C^\top + V)^{-1}C\tilde{Q}A^\top + W_2 = 0. \quad (70)$$

Proof: Consider Riccati recursions

$$\tilde{P}_{t+1} = A\tilde{P}_tA^\top - A\tilde{P}_tC^\top(C\tilde{P}_tC^\top + V)^{-1}C\tilde{P}_tA^\top + W_1 \quad (71)$$

$$\tilde{Q}_{t+1} = A\tilde{Q}_tA^\top - A\tilde{Q}_tC^\top(C\tilde{Q}_tC^\top + V)^{-1}C\tilde{Q}_tA^\top + W_2 \quad (72)$$

with $\tilde{P}_0 = \tilde{Q}_0 \succ 0$. Since (RHS of (71)) $\preceq$ (RHS of (72)) for every t, we have $\tilde{P}_t \preceq \tilde{Q}_t$ for every t (see also [62, Lemma 2.33] for the monotonicity of the Riccati recursion). Under the detectability assumption, we have $\tilde{P}_t \to \tilde{P}$ and $\tilde{Q}_t \to \tilde{Q}$ as $t \to +\infty$ [59, Theorem 14.5.3]. Thus $\tilde{P} \preceq \tilde{Q}$. ∎

Using the above lemma, we obtain the following result.

Lemma 12: $0 \prec W_1 \preceq W_2$, then $v^*(A, W_1) \leq v^*(A, W_2)$.

Proof: Due to the characterization (68) of $v^*(A, W_2)$, there exist $Q \succ 0, V \succ 0, C \in \mathbb{R}^{n \times n}$ such that $v^*(A, W_2) = -\frac{1}{2}\log\det(I - V^{-\frac{1}{2}}CQC^\top V^{-\frac{1}{2}})$ and

$$Q^{-1} - (AQA^\top + W_2)^{-1} = C^\top V^{-1}C. \quad (73)$$

Setting $\tilde{Q} \triangleq AQA^\top + W_2 \succ 0$, it is elementary to show that (73) implies $\tilde{Q}$ satisfies the algebraic Riccati equation (70). Setting $\tilde{L} \triangleq A\tilde{Q}C^\top(C\tilde{Q}C^\top + V)^{-1}$, (70) implies a Lyapunov inequality $(A - \tilde{L}C)\tilde{Q}(A - \tilde{L}C)^\top - \tilde{Q} \prec 0$, showing that $A - \tilde{L}C$ is Schur stable. Hence (A, C) is a detectable pair. By Lemma 11, a Riccati equation (69) admits a positive definite solution $\tilde{P} \preceq \tilde{Q}$. Setting $P \triangleq (\tilde{P}^{-1} + C^\top V^{-1}C)^{-1}$, P satisfies

$$P^{-1} - (APA^\top + W_1)^{-1} = C^\top V^{-1}C \quad (74)$$

Moreover, we have $P \preceq Q$ since

$$0 \prec Q^{-1} = \tilde{Q}^{-1} + C^\top V^{-1}C \preceq \tilde{P}^{-1} + C^\top V^{-1}C = P^{-1}.$$

Since P satisfies (74), we have thus constructed a feasible solution (P, C, V) that upper bounds $v^*(A, W_1)$. That is,

$$\begin{aligned} v^*(A, W_2) &= -\tfrac{1}{2}\log\det(I - V^{-\frac{1}{2}}CQC^\top V^{-\frac{1}{2}}) \\ &\geq -\tfrac{1}{2}\log\det(I - V^{-\frac{1}{2}}CPC^\top V^{-\frac{1}{2}}) \\ &\geq v^*(A, W_1). \end{aligned}$$

∎

Next, we prove that $v^*(A, W)$ is both upper and lower bounded by $\sum_{\lambda_i \in \sigma(A_1)} \log|\lambda_i|$. To establish an upper bound, note that the following inequalities hold with a sufficiently large $\delta > 0$ with $W \preceq \delta I_n$.

$$\begin{aligned} v^*(A, W) &\leq v^*(A, \delta I_n) \\ &\leq v^*(A_1, \delta I_{n_1}) + v^*(A_2, \delta I_{n_2}) = \sum_{\lambda_i \in \sigma(A_1)} \log|\lambda_i|. \end{aligned}$$

Lemma 12 is used in the first step. To see the second inequality, consider the primal representation (65) of $v^*(A, \delta I_n)$. If we restrict decision variables to have block-diagonal structures

$$P = \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}, \quad \Pi = \begin{bmatrix} \Pi_1 & 0 \\ 0 & \Pi_2 \end{bmatrix}$$

according to the partitioning $n = n_1 + n_2$, then the original primal problem (65) with $(A, \delta I_n)$ is decomposed into a problem in terms of decision variables (P_1, Π_1) with data $(A_1, \delta I_{n_1})$ and a problem in terms of decision variables (P_2, Π_2) with data $(A_2, \delta I_{n_2})$. Due to the additional structural restriction, the sum of $v^*(A_1, \delta I_{n_1})$ and $v^*(A_2, \delta I_{n_2})$ cannot be smaller than $v^*(A, \delta I_n)$. Finally, by the arguments in Cases 1 and 2, we have $v^*(A_1, \delta I_{n_1}) = \sum_{\lambda_i \in \sigma(A_1)} \log|\lambda_i|$ and $v^*(A_2, \delta I_{n_2}) = 0$.

To establish a lower bound, we show the following inequalities using a sufficiently small $\epsilon > 0$ such that $\epsilon I \preceq W$.

$$\begin{aligned} v^*(A, W) &\geq v^*(A, \epsilon I_n) \\ &\geq v^*(A_1, \epsilon I_{n_1}) + v^*(A_2, \epsilon I_{n_2}) = \sum_{\lambda_i \in \sigma(A_1)} \log|\lambda_i|. \end{aligned}$$

The first inequality is due to Lemma 12. To prove the second inequality, consider the dual representation (66) of $v^*(A, \epsilon I_n)$. By restricting decision variables $X_{11}, X_{12}, X_{21}, X_{22}$ and Y to have block-diagonal structures according to the partitioning $n = n_1 + n_2$, the original dual problem is decomposed into two problems of the form (66) with $(A_1, \epsilon I_{n_1})$ and $(A_2, \epsilon I_{n_2})$. Since the additional constraints in the dual problem never increase the optimal value, we have the second inequality. Discussions in Cases 1 and 2 are again used in the last step.

G. Proof of Theorem 3

To prove Theorem 3, we reduce the original problem (38) for partially observable plants to a problem for fully observable plants so that the results obtained in Section IV are applicable. To this end, a key technique is the *innovations approach* [63], which has been used in the context of zero-delay rate-distortion theory for a similar purpose [64].

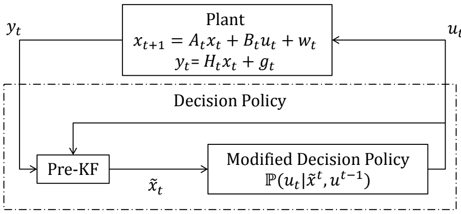

Fig. 9. The space of original decision policy is parametrized by the modified decision policies.

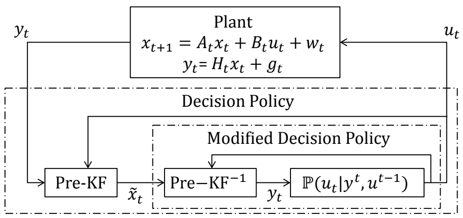

Fig. 10. Proof of Lemma 15.

First, observe that the least-mean square error estimate $\tilde{\mathbf{x}}_t = \mathbb{E}(\mathbf{x}_t|\mathbf{y}^t, \mathbf{u}^{t-1})$ can be computed recursively by the pre-Kalman filter (39). In particular, it is apparent from (39) that the expectation $\mathbb{E}(\mathbf{x}_t|\mathbf{y}^t, \mathbf{u}^{t-1})$ can be written as a linear function $\mathcal{L}_t^{KF}$ of $\mathbf{y}^t$ and $\mathbf{u}^{t-1}$, i.e.,

$$\int x_t \mathbb{P}(dx_{t+1}|y^t, u^{t-1}) = \mathcal{L}_t^{KF}(y^t, u^{t-1}). \tag{75}$$

The Kalman filter is also known as the whitening filter since it has a property that the *innovation* process $\boldsymbol{\psi}_t \triangleq \tilde{L}_{t+1}(\mathbf{y}_{t+1} - H_{t+1}\tilde{\mathbf{x}}_{t+1|t})$ is white.

Lemma 13: The innovation $\boldsymbol{\psi}_t$ is a zero-mean, white (i.e., independent of $\tilde{\mathbf{x}}^t, \mathbf{u}^t$ and $\boldsymbol{\psi}^{t-1}$) Gaussian random variable such that $\boldsymbol{\psi}_t \sim \mathcal{N}(0, \Psi_t)$ with $\Psi_t = \tilde{L}_{t+1}(H_{t+1}\tilde{P}_{t+1|t}H_{t+1}^\top + G_{t+1})\tilde{L}_{t+1}^\top$.

Proof: See, e.g., [65, Section 10.1]. ■

The next observation is important to show that the pre-Kalman filter can be always introduced without loss of performance.

Lemma 14: The pre-Kalman filter (39) is causally invertible; that is, for every $t = 1, \cdots, T$, $(\mathbf{y}^t, \mathbf{u}^{t-1})$ can be reconstructed from $(\tilde{\mathbf{x}}^t, \mathbf{u}^{t-1})$.

Proof: Since we are assuming $W_t \succ 0$ and H_t has full row rank, $\tilde{L}_t$ has full column rank. Thus $\tilde{L}_t^\dagger \triangleq (\tilde{L}_t^\top \tilde{L}_t)^{-1}\tilde{L}_t^\top$ exists for every $t = 1, \cdots, T$. From (39), it is clear that $\mathbf{y}_t$ can be constructed by

$$\mathbf{y}_t = \tilde{L}_t^\dagger \tilde{\mathbf{x}}_t + \tilde{L}_t^\dagger(\tilde{L}_t H_t - I)(A_{t-1}\tilde{\mathbf{x}}_{t-1} + B_{t-1}\mathbf{u}_{t-1}). \tag{76}$$

■

Lemma 15: Without loss of performance, one can assume a decision architecture in Fig. 9, where the "Pre-KF" block represents the pre-Kalman filter (39).

Proof: Suppose $\gamma = \mathbb{P}(u^T \| y^T)$ is an optimal solution to the original problem (38). Construct a modified decision policy block in Figure 9 as shown in Figure 10 using the optimal policy $\gamma = \mathbb{P}(u^T \| y^T)$ for the original problem, where the "Pre-KF^{-1}" block represents the causal inverse (76) of the pre-Kalman filter. Then, the "decision policy" block in Fig. 10 is equivalent to the "decision policy" block in the original problem in Figure 7. Thus, we have shown by construction that there exists a modified decision policy with which the "decision policy" block in Figure 9 attains optimality in the the original problem (38). ■

Lemma 16: For every $t = 1, \cdots, T$,

$$\mathbb{E}\|\mathbf{x}_{t+1}\|_{Q_t}^2 = \text{Tr}(Q_t\tilde{P}_{t+1|t+1}) + \mathbb{E}\|\tilde{\mathbf{x}}_{t+1}\|_{Q_t}^2.$$

Proof: Observe

$$\begin{aligned}
&\mathbb{E}\|\mathbf{x}_{t+1}\|_{Q_t}^2 = \mathbb{E}\|\mathbf{x}_{t+1} - \tilde{\mathbf{x}}_{t+1} + \tilde{\mathbf{x}}_{t+1}\|_{Q_t}^2 \\
&= \mathbb{E}\|\mathbf{x}_{t+1} - \tilde{\mathbf{x}}_{t+1}\|_{Q_t}^2 + \mathbb{E}\|\tilde{\mathbf{x}}_{t+1}\|_{Q_t}^2 + 2\mathbb{E}\tilde{\mathbf{x}}_{t+1}^\top Q_t(\mathbf{x}_{t+1} - \tilde{\mathbf{x}}_{t+1}).
\end{aligned}$$

Since $\mathbb{E}\|\mathbf{x}_{t+1} - \tilde{\mathbf{x}}_{t+1}\|_{Q_t}^2 = \text{Tr}(Q_t\tilde{P}_{t+1|t+1})$, it suffices to prove $\mathbb{E}\tilde{\mathbf{x}}_{t+1}^\top Q_t(\mathbf{x}_{t+1} - \tilde{\mathbf{x}}_{t+1}) = 0$. This can be directly verified as in (77).

$$\begin{aligned}
&\mathbb{E}\left\{\tilde{\mathbf{x}}_{t+1}^\top Q_t(\mathbf{x}_{t+1} - \tilde{\mathbf{x}}_{t+1})\right\} \\
&= \int \tilde{x}_{t+1}^\top Q_t(x_{t+1} - \tilde{x}_{t+1})\mathbb{P}(dx_{t+1}, dy^{t+1}, du^t) \\
&= \int \mathcal{L}_{t+1}^{KF}(y^{t+1}, u^t)^\top Q_t\big(x_{t+1} - \mathcal{L}_{t+1}^{KF}(y^{t+1}, u^t)\big) \\
&\qquad\qquad\qquad\qquad \times \mathbb{P}(dx_{t+1}, dy^{t+1}, du^t) \\
&= \int \mathcal{L}_{t+1}^{KF}(y^{t+1}, u^t)^\top Q_t \\
&\times \left(\int \big(x_{t+1} - \mathcal{L}_{t+1}^{KF}(y^{t+1}, u^t)\big)\mathbb{P}(dx_{t+1}|y^{t+1}, u^t)\right) \mathbb{P}(dy^{t+1}, du^t) \\
&= \int \mathcal{L}_{t+1}^{KF}(y^{t+1}, u^t)^\top Q_t \\
&\times \left(\int x_{t+1}\mathbb{P}(dx_{t+1}|y^{t+1}, u^t) - \mathcal{L}_{t+1}^{KF}(y^{t+1}, u^t)\right) \mathbb{P}(dy^{t+1}, du^t) \\
&= 0
\end{aligned} \tag{77}$$

The last step is due to (75). ■

Finally, we are ready to reduce the original problem (38) for partially observable plants to a problem for fully observable plants. Let $\tilde{\gamma} = \mathbb{P}(u^T \| \tilde{x}^T)$ be a modified decision policy in Fig. 9, and let $\tilde{\Gamma}$ be the space of such policies. Notice that if a policy $\tilde{\gamma} \in \tilde{\Gamma}$ is fixed, then the system equation (1) and the pre-Kalman filter equation (39) uniquely define a joint probability measure $\mathbb{P}(x^{T+1}, y^T, \tilde{x}^T, u^T)$. Expectation and the mutual information with respect to this probability measure will be denoted by $\mathbb{E}_{\tilde{\gamma}}$ and $I_{\tilde{\gamma}}$. By Lemma 15, our main optimization problem (4) can be equivalently written as

$$\min_{\tilde{\gamma} \in \tilde{\Gamma}} \quad \sum\nolimits_{t=1}^{T} I_{\tilde{\gamma}}(\mathbf{y}^t; \mathbf{u}_t | \mathbf{u}^{t-1}) \tag{78a}$$

$$\text{s.t.} \quad \sum\nolimits_{t=1}^{T} \mathbb{E}_{\tilde{\gamma}}\left(\|\mathbf{x}_{t+1}\|_{Q_t}^2 + \|\mathbf{u}_t\|_{R_t}^2\right) \leq D. \tag{78b}$$

By Lemma 14, given $\mathbf{u}^{t-1}$, $\mathbf{y}^t$ can be reconstructed from $\tilde{\mathbf{x}}^t$ and *vice versa*. Thus, we have $I_{\tilde{\gamma}}(\mathbf{y}^t; \mathbf{u}_t|\mathbf{u}^{t-1}) = I_{\tilde{\gamma}}(\tilde{\mathbf{x}}^t; \mathbf{u}_t|\mathbf{u}^{t-1})$. Therefore, using Lemma 16, problem (78) can be further rewritten as

$$\min_{\tilde{\gamma} \in \tilde{\Gamma}} \sum\nolimits_{t=1}^{T} I_{\tilde{\gamma}}(\tilde{\mathbf{x}}^t; \mathbf{u}_t | \mathbf{u}^{t-1}) \tag{79}$$

$$\text{s.t.} \sum\nolimits_{t=1}^{T} \mathbb{E}_{\tilde{\gamma}}\left(\|\tilde{\mathbf{x}}_{t+1}\|_{Q_t}^2 + \|\mathbf{u}_t\|_{R_t}^2\right) + \text{Tr}(Q_t\tilde{P}_{t+1|t+1}) \leq D.$$

Now, notice that the terms $\mathrm{Tr}(Q_t\tilde{P}_{t+1|t+1})$ do not depend on $\tilde{\gamma}$. Thus, by rewriting the pre-Kalman filter (39) equation as

$$\tilde{\mathbf{x}}_{t+1} = A_t\tilde{\mathbf{x}}_t + B_t\mathbf{u}_t + \boldsymbol{\psi}_t \tag{80}$$

and considering (80) as a new "system" with white Gaussian process noise $\boldsymbol{\psi}_t \sim \mathcal{N}(0, \Psi_t)$, problem (79) can be written as

$$\min_{\tilde{\gamma}\in\tilde{\Gamma}} \quad I_{\tilde{\gamma}}(\tilde{\mathbf{x}}^T \to \mathbf{u}^T) \tag{81a}$$

$$\text{s.t.} \quad J_{\tilde{\gamma}}(\tilde{\mathbf{x}}^{T+1}, \mathbf{u}^T) \leq \tilde{D} \tag{81b}$$

where $\tilde{D} = D - \sum_{t=1}^{T}\mathrm{Tr}(Q_t\tilde{P}_{t+1|t+1})$. Since the state $\tilde{X}_t$ of (80) is fully observable by the modified control policy $\tilde{\gamma}$, (81) is now the problem for fully observable systems considered in Section IV.

References

[1] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, "Feedback control under data rate constraints: An overview," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 108–137, 2007.

[2] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 138–162, 2007.

[3] J. Baillieul and P. J. Antsaklis, "Control and communication challenges in networked real-time systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 9–28, 2007.

[4] S. Yüksel and T. Başar, *Stochastic networked control systems*, ser. Systems & Control Foundations & Applications. New York, NY: Springer, 2013, vol. 10.

[5] K. You, N. Xiao, and L. Xie, *Analysis and Design of Networked Control Systems*, ser. Communications and Control Engineering. Springer London, 2015.

[6] S. Tatikonda, "Control under communication constraints," *PhD thesis, Massachusetts Institute of Technology*, 2000.

[7] J. Baillieul, "Feedback coding for information-based control: operating near the data-rate limit," *The 41st IEEE Conference on Decision and Control (CDC)*, 2002.

[8] J. Hespanha, A. Ortega, and L. Vasudevan, "Towards the control of linear systems with minimum bit-rate," *The 15th International Symposium on Mathematical Theory of Networks and Systems (MTNS)*, pp. 1–15, 2002.

[9] G. N. Nair and R. J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM Journal on Control and Optimization*, vol. 43, no. 2, pp. 413–436, 2004.

[10] G. N. Nair, R. J. Evans, I. M. Mareels, and W. Moran, "Topological feedback entropy and nonlinear stabilization," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1585–1597, 2004.

[11] V. S. Borkar and S. K. Mitter, "LQG control with communication constraints," *Communications, Computation, Control, and Signal Processing*, pp. 365–373, 1997.

[12] 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, 2004.

[13] A. S. Matveev and A. V. Savkin, "The problem of LQG optimal control via a limited capacity communication channel," *Systems & Control Letters*, vol. 53, no. 1, pp. 51–64, 2004.

[14] C. D. Charalambous and A. Farhadi, "LQG optimality and separation principle for general discrete time partially observed stochastic systems over finite capacity communication channels," *Automatica*, vol. 44, no. 12, pp. 3181–3188, 2008.

[15] M. Fu, "Linear Quadratic Gaussian control with quantized feedback," *The 2009 American Control Conference (ACC)*, 2009.

[16] K. You and L. Xie, "Linear Quadratic Gaussian control with quantised innovations Kalman filter over a symmetric channel," *IET Control Theory & Applications*, vol. 5, no. 3, pp. 437–446, 2011.

[17] J. S. Freudenberg, R. H. Middleton, and J. H. Braslavsky, "Minimum variance control over a Gaussian communication channel," *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1751–1765, 2011.

[18] L. Bao, M. Skoglund, and K. H. Johansson, "Iterative encoder-controller design for feedback control over noisy channels," *IEEE Transactions on Automatic Control*, vol. 56, no. 2, pp. 265–278, 2011.

[19] S. Yuksel, "Jointly optimal LQG quantization and control policies for multi-dimensional systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, 2014.

[20] M. Huang, G. N. Nair, and R. J. Evans, "Finite horizon LQ optimal control and computation with data rate constraints," *The 44th IEEE Conference on Decision and Control (CDC)*, 2005.

[21] M. D. Lemmon and R. Sun, "Performance-rate functions for dynamically quantized feedback systems," *The 45th IEEE Conference onDecision and Control (CDC)*, 2006.

[22] E. Silva, M. S. Derpich, and J. Ostergaard, "A framework for control system design subject to average data-rate constraints," *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1886–1899, 2011.

[23] E. Silva, M. S. Derpich, and J. Østergaard, "An achievable data-rate region subject to a stationary performance constraint for LTI plants," *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1968–1973, 2011.

[24] E. Silva, M. Derpich, J. Ostergaard, and M. Encina, "A characterization of the minimal average data rate that guarantees a given closed-loop performance level," *IEEE Transactions on Automatic Control*, 2015.

[25] P. Iglesias, "An analogue of Bode's integral for stable nonlinear systems: Relations to entropy," *The 40th IEEE Conference on Decision and Control (CDC)*, 2001.

[26] G. Zang and P. A. Iglesias, "Nonlinear extension of Bode's integral based on an information-theoretic interpretation," *Systems & Control Letters*, vol. 50, no. 1, pp. 11–19, 2003.

[27] N. Elia, "When Bode meets Shannon: Control-oriented feedback communication schemes," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1477–1488, 2004.

[28] 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*, vol. 53, no. 7, pp. 1604–1615, 2008.

[29] H. Marko, "The bidirectional communication theory – A generalization of information theory," *IEEE Transactions on Communications*, vol. 21, no. 12, pp. 1345–1351, 1973.

[30] J. Massey, "Causality, feedback and directed information," *International Symposium on Information Theory and Its Applications (ISITA)*, pp. 27–30, 1990.

[31] G. Kramer, "Capacity results for the discrete memoryless network," *IEEE Transactions on Information Theory*, vol. 49, no. 1, pp. 4–21, 2003.

[32] C. Gourieroux, A. Monfort, and E. Renault, "Kullback causality measures," *Annales d'Economie et de Statistique*, pp. 369–410, 1987.

[33] P.-O. Amblard and O. J. Michel, "On directed information theory and Granger causality graphs," *Journal of computational neuroscience*, vol. 30, no. 1, pp. 7–16, 2011.

[34] J. Jiao, T. A. Courtade, K. Venkat, and T. Weissman, "Justification of logarithmic loss via the benefit of side information," *IEEE Transactions on Information Theory*, vol. 61, no. 10, pp. 5357–5365, 2015.

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

[36] J. L. Massey and P. C. Massey, "Conservation of mutual and directed information," *IEEE International Symposium on Information Theory (ISIT)*, 2005.

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

[38] T. Tanaka, K. H. Johansson, T. Oechtering, H. Sandberg, and M. Skoglund, "Rate of prefix-free codes in LQG control systems," *IEEE International Symposium on Information Theory (ISIT)*, 2016.

[39] M. S. Derpich, E. I. Silva, and J. Østergaard, "Fundamental inequalities and identities involving mutual and directed informations in closed-loop systems," *arXiv preprint arXiv:1301.6427*, 2013.

[40] R. Zamir and M. Feder, "On universal quantization by randomized uniform/lattice quantizers," *IEEE Transactions on Information Theory*, vol. 38, no. 2, 1992.

[41] T. Tanaka, K. H. Johansson, and M. Skoglund, "Optimal block length for data-rate minimization in networked LQG control," *The 6th IFAC Workshop on Distributed Estimation and Control in Networked Systems (NecSys)*, pp. 133–138, 2016.

[42] K. You and L. Xie, "Minimum data rate for mean square stabilizability of linear systems with Markovian packet losses," *IEEE Transactions on Automatic Control*, vol. 56, no. 4, pp. 772–785, 2011.

[43] H. S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proceedings of the IEEE*, vol. 59, no. 11, pp. 1557–1566, 1971.

[44] T. Tanaka and H. Sandberg, “SDP-based joint sensor and controller design for information-regularized optimal LQG control,” *The 54th IEEE Conference on Decision and Control (CDC)*, 2015.

[45] J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg, “Feedback stabilization over signal-to-noise ratio constrained channels,” *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1391–1403, 2007.

[46] T. Tanaka, K.-K. Kim, P. A. Parrilo, and S. K. Mitter, “Semidefinite programming approach to Gaussian sequential rate-distortion trade-offs,” *IEEE Transactions on Automatic Control (To appear)*, 2014.

[47] C. D. Charalambous, P. A. Stavrou, and N. U. Ahmed, “Nonanticipative rate distortion function and relations to filtering theory,” *IEEE Transactions on Automatic Control*, vol. 59, no. 4, pp. 937–952, 2014.

[48] F. Rezaei, N. Ahmed, and C. D. Charalambous, “Rate distortion theory for general sources with potential application to image compression,” *International Journal of Applied Mathematical Sciences*, vol. 3, no. 2, pp. 141–165, 2006.

[49] C. D. Charalambous and P. A. Stavrou, “Optimization of directed information and relations to filtering theory,” *The 2014 European Control Conference (ECC)*, 2014.

[50] I. Csiszár, “On an extremum problem of information theory,” *Studia Scientiarum Mathematicarum Hungarica*, vol. 9, no. 1, pp. 57–71, 1974.

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

[52] E. Shafieepoorfard and M. Raginsky, “Rational inattention in scalar LQG control,” *The 52nd IEEE Conference on Decision and Control (CDC)*, 2013.

[53] P. A. Stavrou, T. Charalambous, and C. D. Charalambous, “Filtering with fidelity for time-varying Gauss-Markov processes,” *The 55th IEEE Conference on Decision and Control (CDC)*, 2016.

[54] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, “SDPT3 - A MATLAB software package for semidefinite programming, version 1.3,” *Optimization methods and software*, vol. 11, no. 1-4, pp. 545–581, 1999.

[55] J. Löfberg, “YALMIP: A toolbox for modeling and optimization in MATLAB,” *The 2004 IEEE International Symposium on Computer Aided Control Systems Design*, 2004.

[56] D. A. Harville, *Matrix algebra from a statistician’s perspective*. Springer, 1997, vol. 1.

[57] G. Folland, *Real analysis: modern techniques and their applications*. Hoboken, NJ, USA: John Wiley & Sons, 1999.

[58] Y.-H. Kim, “Feedback capacity of stationary gaussian channels,” *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 57–85, 2010.

[59] T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation*. Prentice-Hall information and system sciences series, 2000.

[60] T. Tanaka, “Semidefinite representation of sequential rate-distortion function for stationary Gauss-Markov processes,” *The 2015 IEEE Multi-Conference on Systems and Control (MSC)*, 2015.

[61] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2009.

[62] P. R. Kumar and P. Varaiya, *Stochastic systems: estimation, identification and adaptive control*. Prentice-Hall, 1986.

[63] T. Kailath, “An innovations approach to least-squares estimation–Part I: Linear filtering in additive white noise,” *IEEE Transactions on Automatic Control*, vol. 13, no. 6, pp. 646–655, 1968.

[64] T. Tanaka, “Zero-delay rate-distortion optimization for partially observable Gauss-Markov processes,” *The 52nd IEEE Conference on Decision and Control (CDC)*, 2015.

[65] D. Simon, *Optimal state estimation: Kalman, H-infinity, and nonlinear approaches*. John Wiley & Sons, 2006.