LQG Control with Minimum Directed Information: Semidefinite Programming Approach

Takashi Tanaka\textsuperscript{1}  Peyman Mohajerin Esfahani\textsuperscript{2}  Sanjoy K. Mitter\textsuperscript{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.

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.

\textsuperscript{1}TT is with the Department of Aerospace Engineering and Engineering Mechanics at the University of Texas at Austin, USA. ttanaka@utexas.edu; \textsuperscript{2}PME is with the Delft Center for Systems and Control at the Delft University of Technology, Netherlands. P.MohajerinEsfahani@tudelft.nl; \textsuperscript{3} SM is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, USA. mitter@mit.edu.
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 \( x \). 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 \( \mathcal{X}^t \) 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 \( x \) is denoted by \( \mathbb{P} \). A Gaussian distribution with mean \( \mu \) and covariance \( \Sigma \) is denoted by \( \mathcal{N}(\mu, \Sigma) \). The relative entropy of \( x \) from \( y \) is defined by
\[
D(x || y) \triangleq \int \log_2 \frac{dP(x)}{dQ(x)} dP(x) \quad \text{if} \quad P \ll Q
\]
where \( P \ll Q \) means that \( P \) is absolutely continuous with respect to \( Q \), and \( \frac{dP(x)}{dQ(x)} \) denotes the Radon-Nikodym derivative. The mutual information between \( x \) and \( y \) is defined by
\[
I(x; y) \triangleq D(x || x \otimes y), \quad \text{where} \quad x \otimes y \text{ and } x \otimes x \text{ are joint and product probability measures respectively. The entropy of a discrete random variable } x \text{ with probability mass function } P(x_i) \text{ is defined by } H(x) \triangleq -\sum_i P(x_i) \log_2 P(x_i).
\]

II. PROBLEM FORMULATION

Consider a linear time-varying stochastic plant
\[
x_{t+1} = A_t x_t + B_t u_t + w_t, \quad t = 1, \cdots, T, \quad (1)
\]
where \( x_t \) is an \( \mathbb{R}^n \)-valued state of the plant, and \( u_t \) is the control input. We assume that initial state \( x_0 \sim \mathcal{N}(0, P_{01}) \), \( P_{10} > 0 \) and noise process \( w_t \sim \mathcal{N}(0, W_t) \), \( W_t > 0 \), \( t = 0, 1, \cdots, T \) 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 \( \Gamma \) be the space of decision policies, i.e., the space of sequences of Borel measurable stochastic kernels
\[
\mathbb{P}(u_t | x_t) \triangleq \{ \mathbb{P}(u_t | x^t, u^{t-1}) \}_{t=1}^{T}.
\]
A decision policy \( \gamma \in \Gamma \) is evaluated by two criteria:

(i) the LQG control cost
\[
J(x^T, u^T) \triangleq \sum_{t=1}^{T} \mathbb{E} \left[ \| x_{t+1} \|^2_{Q_t} + \| u_t \|^2_{R_t} \right]; \quad (2)
\]
(ii) and directed information
\[
I(x^T \rightarrow u^T) \triangleq \sum_{t=1}^{T} I(x^t; u_t | u^{t-1}). \quad (3)
\]

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 \( \gamma \). In what follows, we often write
The length of a codeword information and the standard mutual information:

\[ I(x^T; u^T) = \sum_{t=1}^{T} I(x_t; u_t) \]

Directly, 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

\[ x_{t+1} = Ax_t + Bu_t + w_t, \quad t \in \mathbb{N} \]

with \( w_t \sim \mathcal{N}(0, W) \), 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 \( R > 0 \). Let \( \Gamma \) be the space of Borel-measurable stochastic kernels \( \mathbb{P}(u^\infty \mid x^\infty) \). The problem of interest is

\[ \text{Di}(D) \triangleq \min_{\gamma \in \Gamma} \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} I(x_t; u_t) \]

s.t. \( \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} J_0(x_{t+1}; u_t) \leq D \).

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 \( t = 1, \ldots, T \), let \( \mathcal{A}_t \subset \{0, 1, 00, 01, 10, 11, 000, \ldots\} \) be a set of uniquely decodable variable-length codewords [37, Ch.5]. Assume that codewords are generated by a causal policy

\[ \mathbb{P}(a^\infty \mid x^\infty) \triangleq \{ \mathbb{P}(a_t \mid x^t, a^{t-1}) \}_{t=1,2,\ldots} \]

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

\[ \mathbb{P}(u^\infty \mid a^\infty) \triangleq \{ \mathbb{P}(u_t \mid a^t, a^{t-1}) \}_{t=1,2,\ldots} \]

The length of a codeword \( a_t \in \mathcal{A}_t \) is denoted by a random variable \( l_t \). Let \( \Gamma' \) be the space of triplets

\[ (x_t, x_{t+1} = Ax_t + Bu_t + w_t, u_t) \]

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.

\[ \text{Fig. 2. LQG control over noiseless binary channel.} \]

\[ \{ \mathbb{P}(a^\infty \mid x^\infty), \mathcal{A}^\infty, \mathbb{P}(u^\infty \mid a^\infty) \}. \]

Introduce a quadratic control cost

\[ J(x_{t+1}; u^T) \triangleq \sum_{t=1}^{T} \mathbb{E} \left( ||x_{t+1}||_2^2 + ||u_t||_R^2 \right) \]

with \( Q > 0 \) and \( R > 0 \). 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

\[ \min \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}(l_t) \]

s.t. \( \limsup_{T \to \infty} \frac{1}{T} J(x_{t+1}; u^T) \leq D \).

It is difficult to evaluate \( \text{Di}(D) \) directly since (7) is a highly complex optimization problem. Nevertheless, Silva et al. [22] observed that \( \text{Di}(D) \) is closely related to \( \text{Di}(D) \) defined by (6).

The following result is due to [38].

\[ \text{Di}(D) \leq \text{Di}(D) + \frac{r}{2} \log \frac{4\pi e}{12} + 1 \forall D > 0. \]

Here, \( r \) is an integer no greater than the state space dimension of the plant. 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

\[ I(x^T; u^T) \leq I(x^T; a^T ||u^{T-1}_+), \]

where the right hand side is Kramer’s notation [31] for causally conditioned directed information

\[ \sum_{t=1}^{T} I(x^t; a_t ||u^{t-1}, u^{t-1}). \]

**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 \( x_t \) is affected by feedback. See also [39] for relevant inequalities involving directed information.

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

\[ I(x^T; u^T) \leq \sum_{t=1}^{T} I(x^t; a_t ||u^{t-1}_+, u^{t-1}) \]

\[ = \sum_{t=1}^{T} \left( H(a_t ||u^{t-1}, u^{t-1}) - H(a_t ||u^t, a^{t-1}, u^{t-1}) \right) \]

\[ \leq \sum_{t=1}^{T} H(a_t ||u^{t-1}, u^{t-1}) \]

\[ \leq \sum_{t=1}^{T} H(a_t) \]

\[ \leq \sum_{t=1}^{T} \mathbb{E}(l_t). \]
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

\[ y_t = C_t \tilde{x}_t + v_t, \quad v_t \sim \mathcal{N}(0, V_t), \quad V_t \succeq 0 \]  \hspace{1cm} (10)

where \( v_1, t = 1, \ldots, T \) are mutually independent.

2. The Kalman filter computing \( \tilde{x}_t = \mathbb{E}(x_t | y^t, u^{t-1}) \).

3. The certainty equivalence controller \( u_t = K_t \tilde{x}_t \).

The role of the mechanism (10) is noteworthy. Recall that in the current problem setting in Figure 1, the state vector \( x_t \) 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 \( x_t \) 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:

\[
S_t = \begin{cases} 
Q_t & \text{if } t = T \\
Q_t + \Phi_{t+1} & \text{if } t = 1, \ldots, T - 1 
\end{cases} 
\]  \hspace{1cm} (11a)

\[
\Phi_t = A_t^T (S_t - S_t B_t (B_t^T S_t B_t + R_t)^{-1} B_t^T S_t) A_t 
\]  \hspace{1cm} (11b)

\[
K_t = -(B_t^T S_t B_t + R_t)^{-1} B_t^T S_t A_t 
\]  \hspace{1cm} (11c)

\[
\Theta_t = K_t (B_t^T S_t B_t + R_t) K_t^T. 
\]  \hspace{1cm} (11d)

Positive semidefinite matrices \( \Theta_t \) will be used in Step 2.

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

\[
\min \{ P_{1|t}, \Pi_t \}_{t=1}^T \frac{1}{2} \sum_{t=1}^T \log \det \Pi_t^{-1} + c_1 \]  \hspace{1cm} (12a)

subject to:

\[
\sum_{t=1}^T \text{Tr}(\Theta_t P_{1|t}) + c_2 \leq D 
\]  \hspace{1cm} (12b)

\[
\Pi_t \succeq 0, \quad t = 1, \ldots, T - 1 
\]  \hspace{1cm} (12c)

\[
P_{t+1|t} \leq A_t P_{t|t} A_t^T + W_t, \quad t = 1, \ldots, T - 1 
\]  \hspace{1cm} (12d)

\[
P_{t|t} - \Pi_t A_t P_{t|t} A_t^T + W_t \geq 0, \quad t = 1, \ldots, T - 1 
\]  \hspace{1cm} (12e)

The constraint (12c) is imposed for every \( t = 1, \ldots, T \), while (12e) and (12f) are for every \( t = 1, \ldots, T - 1 \).

Constants \( c_1 \) and \( c_2 \) are given by

\[
c_1 = \frac{1}{2} \log \det P_{1|0} + \frac{1}{2} \sum_{t=1}^{T-1} \log \det W_t 
\]

\[
c_2 = \text{Tr}(N_t P_{1|t}) + T \sum_{t=1}^T \text{Tr}(W_t S_t). 
\]

Define signal-to-noise ratio matrices \( \{ \text{SNR}_t \}_{t=1}^T \) by

\[
\text{SNR}_t \triangleq P_{t|t}^{-1} - P_{t-1|t-1}, \quad t = 1, \ldots, T 
\]

\[
P_{t|t} \leq \hat{A}_{t-1} P_{t-1|t-1} A^T_{t-1} + W_{t-1}, \quad t = 2, \ldots, T 
\]

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

\[
\text{SNR}_t = C_t V_t^{-1} C_t^T, \quad t = 1, \ldots, T. 
\]  \hspace{1cm} (13)

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

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

\[
L_t = P_{t|t-1} C_t^T (C_t P_{t|t-1} C_t^T + V_t)^{-1}. 
\]  \hspace{1cm} (14)

Construct a Kalman filter by

\[
\hat{x}_{t|t-1} = x_{t|t-1} + L_t (y_t - C_t \hat{x}_{t|t-1}) 
\]  \hspace{1cm} (15a)

\[
\hat{x}_{t+1|t} = A_t \hat{x}_t + B_t u_t. 
\]  \hspace{1cm} (15b)

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

An optimization problem (12) plays a key role in the proposed synthesis. Intuitively, (12) "schedules" the optimal
sequence of covariance matrices \( \{P_0, t\} \) 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). The constraint (16c) is imposed for every time-invariant and infinite-horizon problems (5) and (6). Theorem 1 shows a noteworthy fact that there exists a virtual sensor mechanism to realize it 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 Appendix E.

**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):

\[
\begin{align}
\min_{\{P_t, t\}} & \quad \frac{1}{2} \sum_{t=1}^{T} \log \det \Delta_t^{-1} + c_1 \quad \text{(16a)} \\
\text{s.t.} & \quad \sum_{t=1}^{T} \text{Tr}(\Theta_t P_{t(t)}) + c_2 \leq D \quad \text{(16b)} \\
& \quad \Delta_t \succ 0, \quad \text{(16c)} \\
& \quad P_{1|1} \preceq P_{1|0}, \quad P_{T|T} = \Delta_T, \quad \text{(16d)} \\
& \quad P_{t+1|t+1} \preceq A_t P_t A_t^\top + F_t F_t^\top, \quad \text{(16e)} \\
& \quad \left[ I - F_t A_t P_t A_t^\top - F_t F_t^\top \right] \succeq 0. \quad \text{(16f)}
\end{align}
\]

The constraint (16c) is imposed for every \( t = 1, \ldots, T \), while (16d) and (16f) are for every \( t = 1, \ldots, T - 1 \). Constants \( c_1 \) and \( c_2 \) are given by \( c_1 = \frac{1}{2} \sum_{t=1}^{T} \log \det P_{t|0} + \sum_{t=1}^{T-1} \log \det A_t \) and \( c_2 = \text{Tr}(A_t P_{t|0}) + \sum_{t=1}^{T} \text{Tr}(F_t S_t F_t) \). This formulation requires that \( A_t, t = 1, \ldots, T - 1 \) 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
  \[
  A^\top S A - S - A^\top S B (B^\top S B + R)^{-1} B^\top S A + Q = 0 \quad \text{(17)}
  \]
  and determine the optimal feedback control gain by \( K = -(B^\top S B + R)^{-1} B^\top S A \). Set \( \Theta = K^\top (B^\top S B + R) K \).

- **Step 2** (Virtual sensor design) Choose \( P \) and \( \Pi \) as the solution to a max-det problem:
  \[
  \begin{align}
  \min_{P, \Pi} & \quad \frac{1}{2} \log \det \Pi^{-1} + \frac{1}{2} \log \det W \quad \text{(18a)} \\
  \text{s.t.} & \quad \text{Tr}(\Theta P) + \text{Tr}(WS) \leq D, \quad \text{(18b)} \\
  & \quad \Pi \succ 0, \quad \text{(18c)} \\
  & \quad P \preceq APA^\top + W, \quad \text{(18d)} \\
  & \quad \left[ P - \Pi APA^\top \right] \succeq 0. \quad \text{(18e)}
  \end{align}
  \]

Define \( \hat{P} \triangleq APA^\top + W, \text{SNR} \triangleq P^{-1} - \hat{P}^{-1} \) and set \( r = \text{rank}(\text{SNR}) \). Choose a virtual sensor \( y_t = Cx_t + v_t, \quad v_t \sim N(0, V) \) with matrices \( C \in \mathbb{R}^{r \times n} \) and \( V \in \mathbb{S}^n_{++} \) such that \( C^\top V^{-1} C = \text{SNR} \).

- **Step 3** (Filter design) Design a time-invariant Kalman filter
  \[
  \dot{x}_t = x_{t-1} + L(z_t - Cx_{t-1}) \\
  \dot{x}_{t+1} = A\dot{x}_t + B u_t
  \]
  with \( L = \hat{P} C^\top (\hat{P} C^\top + V)^{-1} \).

**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 \( \text{DI}(D) \) 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 \( \text{DI}(D) \) 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 \( \lim_{D \to +\infty} \text{DI}(D) \). 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 \( \lambda_i \) of \( A \) such that \( |\lambda_i| \geq 1 \) counted with multiplicity. Then,

\[
\lim_{D \to +\infty} \text{DI}(D) = \sum_{\lambda_i \in \sigma_+ (A)} \log |\lambda_i|. 
\]

**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 \( \lim_{T \to +\infty} \frac{1}{T} \text{I}(X_t \to U_t) \) under the stabilization scheme proposed there and study how it is compared to the right hand side of (19).
"Gaussian Sequential Rate-Distortion Problem"

\[ \mathcal{X}_t \xrightarrow{\text{Sensor}} \mathcal{Y}_t \xrightarrow{\text{Filter}} \hat{\mathcal{X}}_t \xrightarrow{\text{Controller}} \mathcal{U}_t \]

"LQG Optimal Control Problem"

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 \( y_t \) to \( x_t \). The proposed sensor design algorithm has already appeared in [44]. In this paper we strengthen the result by showing that the designed linear sensor 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 control theory [43]. Unlike the standard framework in which the filter-controller separation principle in the standard LQG filter and filter-controller separation principles.

The horizontal asymptote \( \sum_{\lambda \in \sigma(A)} \log |\lambda| = 1.169 \) 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.

E. 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

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

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 unrestrained amount of information about the state is available. This corresponds to the performance of the state-feedback linear-quadratic regulator (LQR). The horizontal asymptote \( \sum_{\lambda \in \sigma(A)} \log |\lambda| = 1.169 \) 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 \( y_t = Cx_t + v_t, \), \( v_t \sim \mathcal{N}(0, V) \) to achieve this performance is given by

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

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

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

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

\[
C = \begin{bmatrix}
-0.876 & 0.271 & -0.169 & -0.362
\end{bmatrix}, \quad V = 1.775
\]
Figure 6 shows the closed-loop responses of the state trajectories and their Kalman estimates are shown. D_\Gamma subsets Fig. 6. Closed-loop performances of the controllers designed for performance D. DI Fig. 5. (Top left) Data rate DI (middle), and DI (right) Singular values are shown in block bars. An SDP solver SDPT3 [54] with SNR singular values of 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.

VI. DERIVATION OF MAIN RESULT

This section is devoted to prove Theorem 1. We first define subsets D_0, D_1, and D_2 of the policy space D as follows.

D_0 : The space of policies with three-stage separation structure explained in Section IV.

D_1 : The space of linear controllers without state memory followed by linear deterministic feedback control. Namely, a policy P(\|x^T\|T) in D_1 can be expressed as a composition of

y_t = C_t x_t + v_t, v_t \sim N(0, V_t) \tag{20}

and u_t = l_t(y^t), where C_t \in \mathbb{R}^{r \times n}, r_t is some nonnegative integer, V_t > 0, and l_t(\cdot) is a linear map.

D_2 : The space of linear policies without state memory. Namely, a policy P(u^T|X^T) in D_2 can be expressed as

u_t = M_t x_t + N_t u_t^{-1} + g_t, g_t \sim N(0, G_t) \tag{21}

with some matrices M_t, N_t, and G_t \geq 0.

A. Proof outline

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

\inf_{\gamma \in D_1; J_t \leq D} I_{\gamma}(x^T \rightarrow u^T) \tag{22a}
\geq \inf_{\gamma \in D_2; J_t \leq D} \sum_{t=1}^{T} I_{\gamma}(x_t; u_t|u_t^{-1}) \tag{22b}
\geq \inf_{\gamma \in D_2; J_t \leq D} \sum_{t=1}^{T} I_{\gamma}(x_t; y_t|y_t^{-1}) \tag{22c}
\geq \inf_{\gamma \in D_2; J_t \leq D} \sum_{t=1}^{T} I_{\gamma}(x_t; y_t|y_t^{-1}) \tag{22d}
\geq \inf_{\gamma \in D_2; J_t \leq D} \sum_{t=1}^{T} I_{\gamma}(x^T \rightarrow u^T). \tag{22f}

Since D_0 \subseteq D, clearly (22a) \leq (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 D_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 D_0, D_1 \equiv C_t^T 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_{\mathcal{P}} = \{P(u_t|x^T, u_t^{-1})\}_{t=1}^{T} \in \Gamma \ that attains a finite objective value in (22b), there exists \gamma_{Q} = \{Q(u_t|x^T, u_t^{-1})\}_{t=1}^{T} \in \Gamma_2 such that J_t = J_0 and

\sum_{t=1}^{T} I_{\mathcal{P}}(x_t; u_t|u_t^{-1}) \geq \sum_{t=1}^{T} I_{\mathcal{Q}}(x_t; u_t|u_t^{-1})

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

\hat{P}(u_t|x^T, u_t^{-1}) \triangleq \hat{P}(u_t+\mathbb{E}_{\mathcal{P}}(u_t)|x^T+\mathbb{E}_{\mathcal{P}}(x^T), u_t^{-1}+\mathbb{E}_{\mathcal{P}}(u_t^{-1}))

which generates a zero-mean joint distribution \hat{P}(x^{T+1}, u^T). We have I_{\mathcal{P}} = I_{\mathcal{Q}} in view of the translation invariance of
mutual information, and \( J_\rho \leq J_Y \) 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.

\[
\sum_{t=1}^{T} I_p(x_t; u_t|u^{t-1}) \geq \sum_{t=1}^{T} I_G(x_t; u_t|u^{t-1})
\]  

(23)

**Proof:** See Appendix C.

Next, we are going to construct a policy \( \gamma_Q = \{Q(u_t|x_t, u^{t-1})\}_{t=1}^{T} \in \Gamma_2 \) using a jointly Gaussian measure \( \mathbb{G}(x^{T+1}, u^T) \). Let \( E_t x_t + F_t u^{t-1} \) be the least mean-square error estimate of \( u_t \) given \( (x_t, u^{t-1}) \) 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 \( \mathbb{Q}(u_t|x_t, u^{t-1}) = \mathbb{N}(E_t x_t + F_t u^{t-1}, V_t) \). By construction, \( \mathbb{Q}(u_t|x_t, u^{t-1}) \) satisfies

\[
d\mathbb{G}(x_t, u^t) = d\mathbb{Q}(u_t|x_t, u^{t-1})d\mathbb{G}(x_t, u^{t-1}).
\]

(24)

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

\[
d\mathbb{Q}(x^t, u^t) = d\mathbb{P}(x_t|x_{t-1}, u_{t-1})d\mathbb{Q}(x^{t-1}, u^{t-1})
\]

(25)

\[
d\mathbb{Q}(x^t, u^t) = d\mathbb{Q}(u_t|x_t, u^{t-1})d\mathbb{Q}(x^{t-1}, u^{t-1})
\]

(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 \( \gamma \) with a new policy \( \gamma_Q \) does not change the LQG control cost.

\[
J_{\gamma_Q} = \int (||x_{t+1}||_Q^2 + ||u_t||_R^2) d\mathbb{P}(x_{t+1}, u^t)
\]

(27a)

\[
= \int (||x_{t+1}||_Q^2 + ||u_t||_R^2) d\mathbb{G}(x_{t+1}, u^t)
\]

(27b)

\[= J_{\gamma}.
\]

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

\[
\sum_{t=1}^{T} I_p(x_t; u_t|u^{t-1}) \geq \sum_{t=1}^{T} I_G(x_t; u_t|u^{t-1})
\]

(28a)

\[= \sum_{t=1}^{T} I_Q(x_t; u_t|u^{t-1}).
\]

(28b)

The inequality (28a) is due to Lemma 2. In (28b), \( I_Q(x_t; u_t|u^{t-1}) = I_Q(x_t; u_t|u^{t-1}) \) holds for every \( t = 1, \cdots, T \) because of Lemma 3.

\[\text{(30)}\]

**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

\[
I_{\gamma_2}(x_t; u_t|u^{t-1}) = I_{\gamma_1}(x_t; y_t|y^{t-1})
\]

(29)

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

\[
u_t = M_t x_t + N_t u^{t-1} + g_t, \quad g_t \sim \mathcal{N}(0, G_t).
\]

Define \( \tilde{y}_t \equiv M_t x_t + g_t \). If we write \( N_t u^{t-1} = N_t u^{t-1} + \cdots + N_1 u_1 \), it can be seen that \( u^t \) and \( \tilde{y}^t \) are related by an invertible linear map

\[
\begin{bmatrix}
\tilde{y}_t \\
\vdots \\
\tilde{y}_1
\end{bmatrix} = \begin{bmatrix}
I \\
\vdots \\
-N_{2,1} \\
\vdots \\
-N_{1,1} & \cdots & -N_{1,t-1} & I
\end{bmatrix}
\begin{bmatrix}
u_1 \\
\vdots \\
u_t
\end{bmatrix}
\]

(30)

for every \( t = 1, \cdots, T \). Hence,

\[
I(x_t; u_t|u^{t-1}) = I(x_t; \tilde{y}_t + N_t u^{t-1} | \tilde{y}^{t-1}, u^{t-1})
\]

(31)

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

\[
\text{Im}(M_t) \subseteq \text{Im}(G_t) = \text{Im}(E_t^T)
\]

(32)

Otherwise, the component of \( u_t \) in \( \text{Im}(G_t)^\perp \) depends deterministically on \( x_t \) and (31) is unbounded. Now, define \( y_t \equiv E_t \tilde{y}_t = E_t M_t x_t + E_t g_t \), \( g_t \sim \mathcal{N}(0, G_t) \). Then, we have

\[
E_t^T y_t = E_t^T E_t M_t x_t + E_t^T E_t g_t, \quad g_t \sim \mathcal{N}(0, G_t)
\]

(33)

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

\[
I(x_t; y_t|y^{t-1}) = I(x_t; y_t|y^{t-1}).
\]

(33)

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

\[
u_t = l_t(y^t).
\]

(34)

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

\[\text{D. Proof of inequality (22e)}\]

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

\[
I_{\gamma}(x_t; y_t|y^{t-1})
\]

(35)
Moreover, for every fixed sensor equation (20), covariance matrices are determined by the Kalman filtering formula
\[ P_{t|t} = ((A_{t-1}P_{t-1|t-1}A_{t-1}^T + W_{t-1})^{-1} + \text{SNR}_t)^{-1}. \]

Hence, conditional mutual information (35) depends only on the choice of \{\text{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 \(J_\gamma\) 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 \(J_\gamma\) is a composition of a Kalman filter \(\hat{x}_t = E(x_t|y^t, u^{t-1})\) and a certainty equivalence controller \(u_t = K_t \hat{x}_t\). This implies that an optimal solution \(\gamma\) can always be found in the class \(\Gamma_0\), 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 \(J_\gamma\) explicitly as
\[ J_\gamma = \text{Tr}(N_1 P_{1|0}) + \sum_{t=1}^T \left( \text{Tr}(W_t S_t) + \text{Tr}(\Theta_t P_{t|t}) \right). \]

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
\[ I_\gamma(x^T \rightarrow u^T) \leq I_\gamma(x^T \rightarrow y^T | u^{t-1}) \]
\[ = \sum_{t=1}^T I_\gamma(x^T; y_t^T | y^{t-1}, u^{t-1}) \]
\[ = \sum_{t=1}^T I_\gamma(x^T; y_t^T | y^{t-1}) + I_\gamma(x^T; y_t^T | y^{t-1}) \]
\[ = \sum_{t=1}^T I_\gamma(x^T; y_t^T | y^{t-1}). \]
The last equality holds since, by construction, \(y_t = C_t x_t + v_t\) is conditionally independent of \(x^T\) given \(x_t\).

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 \{\text{P}_{t|t}, \text{SNR}_t\}_{t=1}^T as
\[ \min_{\gamma_1} \sum_{t=2}^T \left( \frac{1}{2} \log \det(A_{t-1}P_{t-1|t-1}A_{t-1}^T + W_{t-1}) - \frac{1}{2} \log \det P_{t|t} \right) \]
\[ + \frac{1}{2} \log \det P_{1|0} - \frac{1}{2} \log \det P_{1|1} \]
\[ \text{s.t. } \text{Tr}(N_1 P_{1|0}) + \sum_{t=1}^T \left( \text{Tr}(W_t S_t) + \text{Tr}(\Theta_t P_{t|t}) \right) \leq D, \]
\[ P_{t|1}^{-1} = P_{1|1}^{-1} + \text{SNR}_1, \]
\[ P_{t|t}^{-1} = (A_{t-1}P_{t-1|t-1}A_{t-1}^T + W_{t-1})^{-1} + \text{SNR}_t, \]
\[ t = 2, ..., T. \]

This problem can be reformulated as a max-det problem as follows. First, variables \{\text{SNR}_t\}_{t=1}^T are eliminated from the problem by replacing the last three constraints with equivalent conditions
\[ 0 < P_{1|1} \preceq P_{1|0}, \]
\[ 0 < P_{t|t} \preceq A_{t-1}P_{t-1|t-1}A_{t-1}^T + W_{t-1}, \]
\[ t = 2, ..., T. \]

Second, the following equalities can be used for \(t = 1, ..., T - 1\) to rewrite the objective function:
\[ \frac{1}{2} \log \det(A_{t}P_{t|t}A_{t}^T + W_{t}) - \frac{1}{2} \log \det P_{t|t} \]
\[ = \frac{1}{2} \log \det(P_{t|t}^{-1} + A_{t}^T W_{t}^{-1} A_{t}) + \frac{1}{2} \log \det W_{t} \]
\[ = \inf_{\Pi_t} \frac{1}{2} \log \det \Pi_t^{-1} + \frac{1}{2} \log \det W_t \]
\[ \text{s.t. } 0 < \Pi_t \preceq (P_{t|t}^{-1} + A_{t}^T W_{t}^{-1} A_{t})^{-1} \]
\[ = \inf_{\Pi_t} \frac{1}{2} \log \det \Pi_t^{-1} + \frac{1}{2} \log \det W_t \]
\[ \text{s.t. } \Pi_t > 0, P_{t|t} - \Pi_t - A_{t}P_{t|t}A_{t}^T + W_{t} \succeq 0. \]

In step (37a), we have used the matrix determinant theorem [56, Theorem 18.1.1]. An additional variable \(\Pi_t\) 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).

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 \(x_t\) 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 \(y_t = H_t x_t + g_t\) are given. We assume that initial state \(x_1 \sim \mathcal{N}(0, P_{1|0})\), \(P_{1|0} > 0\) and noise processes \(w_t \sim \mathcal{N}(0, W_t)\), \(W_t > 0\), \(g_t \sim \mathcal{N}(0, G_t)\), \(G_t > 0\), \(t = 1, ..., T\) are mutually independent. We also assume that \(H_t\) has full row rank for \(t = 1, ..., T\). Consider the following problem:
\[ \min_{\gamma \in \mathcal{G}} \quad I_\gamma(y^T \rightarrow u^T) \]
\[ \text{s.t. } J_\gamma(x^{T+1}, u^T) \leq D. \]
where \( \Gamma \) 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
\[
\hat{x}_t = \hat{x}_t|_{t-1} + L_t(y_t - H_t \hat{x}_t|_{t-1})
\]
where the Kalman gains \( \{L_t\}_{t=1}^{T+1} \) are computed by
\[
\begin{align*}
L_t &= \hat{P}_t|_{t-1}H_t^T(\hat{H}_t\hat{P}_t|_{t-1}H_t^T + G_t)^{-1} \\
\hat{P}_t|_{t-1} &= (I - \hat{L}_tH_t)\hat{P}_t|_{t-1} \\
\hat{L}_t &= A_t\hat{x}_t + B_t\hat{u}_t,
\end{align*}
\]
Matrices \( \Psi_t \) are used in Step 3.

**Step 2.** (Controller design) Determine feedback control gains \( K_t \) via the backward Riccati recursion:
\[
\begin{align*}
S_t &= \begin{cases} 
  Q_t & \text{if } t = T \\
  Q_t + N_{t+1} & \text{if } t = 1, \ldots, T - 1
\end{cases} \\
M_t &= B_t^T S_t B_t + R_t \\
N_t &= A_t^T (S_t - S_t B_t M_t^{-1} B_t^T S_t) A_t \\
K_t &= -M_t^{-1} B_t^T S_t A_t \\
\Theta_t &= K_t^T M_t K_t
\end{align*}
\]
Positive semidefinite matrices \( \Theta_t \) will be used in Step 3.

**Step 3.** (Virtual sensor design) Solve a max-det problem with respect to \( \{P_t|_{t}, \Pi_t\}_{t=1}^{T} \):
\[
\begin{align*}
\min & \quad \frac{1}{2} \sum_{t=1}^{T} \log \det \Pi_t^{-1} + c_1 \\
\text{s.t.} & \quad \sum_{t=1}^{T} \text{Tr}(\Theta_t P_t|_{t}) + c_2 \leq D \\
& \quad \Pi_t > 0, \\
& \quad P_{1|1} \leq P_{1|0}, P_{T|T} = \Pi_T, \\
& \quad P_{t+1|t+1} \leq A_t P_t|_{t} A_t^T + \Psi_t, \\
& \quad \begin{bmatrix} P_{t|t} - \Pi_t & P_{t|t} A_t^T \\ A_t P_{t|t} & A_t P_{t|t} A_t^T + \Psi_t \end{bmatrix} \succeq 0.
\end{align*}
\]

The constraint (41c) is imposed for every \( t = 1, \ldots, T \), while (41e) and (41f) are for every \( t = 1, \ldots, T - 1 \). Constants \( c_1 \) and \( c_2 \) are given by
\[
\begin{align*}
c_1 &= \frac{1}{2} \log \det P_{1|0} + \frac{1}{2} \sum_{t=1}^{T-1} \log \det \Psi_t \\
c_2 &= \text{Tr}(N_t P_t|_{0}) + \sum_{t=1}^{T} \text{Tr}(\Psi_t S_t).
\end{align*}
\]
If \( \Psi_t \) is singular for some \( t \), consider an alternative max-det problem suggested in Remark 1. Set \( r_t = \text{rank}(P_{t|t}^{-1} - P_{t|t-1}^{-1}) \), where
\[
P_{t|t-1} = A_{t-1} P_{t-1|t-1} A_{t-1}^T + W_{t-1}, t = 2, \ldots, T.
\]
Choose matrices \( C_t \in \mathbb{R}^{r_t \times n_t} \) and \( V_t \in \mathbb{S}^{r_t}_{++} \) so that
\[
C_t^T V_t^{-1} C_t = P_{t|t}^{-1} - P_{t|t-1}^{-1}
\]
for \( t = 1, \ldots, T \). 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
\[
\hat{x}_t = \hat{x}_t|_{t-1} + L_t(z_t - C_t\hat{x}_t|_{t-1})
\]
where Kalman gains \( \hat{L}_t \) are computed by
\[
\hat{L}_t = P_{t|t-1} C_t^T (C_t P_{t|t-1} C_t^T + V_t)^{-1}.
\]
If \( r_t = 0 \), \( L_t \) is a null matrix and (43a) is simply replaced by \( \hat{x}_t = \hat{x}_t|_{t-1} \).

**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 $t = 1, \ldots, T$.

\begin{align*}
I(x^t; a_t | u^{t-1}, u^t) &= I(x^t; a_t | u^{t-1}, u^t) - I(x^t; u_t | u^{t-1}) \\
&= I(x^t; a_t | u^{t-1}, u^t) - I(x^t; u_t | u^{t-1}) \quad (45a) \\
&= I(x^t; a_t | u^t) - I(x^t; a_t^1 | u^{t-1}) - I(x^t; a_t^1 | u^{t-1}) \quad (45b) \\
&= I(x^t; a_t^1 | u^t) - I(x^t; a_t^1 | u^{t-1}) \quad (45c) \\
&= I(x^t; a_t | u^t) - I(x^t; a_t^1 | u^{t-1}) \quad (45d)
\end{align*}

When $t = 1$, the above identity is understood to mean $I(x_1; a_1) = I(x_1; u_1)$ which clearly holds as $\mathbf{X}_1 = \mathbf{a}_1 - \mathbf{u}_1$ form a Markov chain. Equation (45a) holds because $I(x^t; a_t, u_t | u^{t-1}) = I(x^t; a_t | u^t, u^{t-1})$ and the second term is zero since $x^t - (a^t, u^t) - u_t$ form a Markov chain. Equation (45b) is obtained by applying the chain rule for mutual information in two different ways:

\begin{align*}
I(x^t; a_t, u_t | u^{t-1}) &= I(x^t; a_t | u^{t-1}, u^t) + I(x^t; u_t | u^{t-1}) \\
&= I(x^t; a_t | u^{t-1}, u^t) + I(x^t; a_t | u^t) \\
&= I(x^t; a_t | u^t) - I(x^t; a_t | u^{t-1}).
\end{align*}

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

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

\begin{align*}
\sum_{t=1}^T [I(x^t; a_t | u^{t-1}, u^t) - I(x^t; u_t | u^{t-1})] &= 0 \quad (46a) \\
\sum_{t=1}^T [I(x^t; a_t | u^t) - I(x^t; a_t^1 | u^{t-1})] &= 0 \quad (46b)
\end{align*}

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

\[ \mathbb{P}_{x,y}(B_x \times B_y) = \int_{B_y} \mathbb{P}_{x|y}(B_x | y) \mathbb{P}_y(dy) \quad (47) \]

for every $B_x \in \mathcal{B}_x$ and $B_y \in \mathcal{B}_y$. If $\mathbb{P}_{x,y} \ll \mathbb{P}_x \otimes \mathbb{P}_y$, then $\mathbb{P}_{x|y} \ll \mathbb{P}_x, \mathbb{P}_y - a.e.$, and

\[ \frac{d\mathbb{P}_{x,y}}{d(\mathbb{P}_x \otimes \mathbb{P}_y)} = \frac{d\mathbb{P}_{x|y}}{d\mathbb{P}_x} \mathbb{P}_y - a.e. \]

**Proof:** Suppose $\mathbb{P}_{x,y} \ll \mathbb{P}_x \otimes \mathbb{P}_y$, and let $f(x, y) = \frac{d\mathbb{P}_{x,y}}{d(\mathbb{P}_x \otimes \mathbb{P}_y)}$ be the Radon-Nikodym derivative. For every $B_x \in \mathcal{B}_x$ and $B_y \in \mathcal{B}_y$, we have

\[ \mathbb{P}(B_x \times B_y) = \int_{B_x \times B_y} f(x, y) d(\mathbb{P}_x \otimes \mathbb{P}_y)(x, y) \]

\[ = \int_{B_y} \int_{B_x} f(x, y) d\mathbb{P}_{x|y}(x) d\mathbb{P}_y(y) \quad (48) \]

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

\[ \mathbb{P}_{x|y}(B_x | y) = \int_{B_x} f(x, y) d\mathbb{P}_{x|y}(x) \mathbb{P}_y(y) - a.e. \quad (49) \]

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

**Lemma 5:** Let $\mathbb{P}_{x,y,z}$ 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 $\mathbb{P}_{x,y,z}$ has a covariance matrix $\Sigma_{x,y,z}$, and there exists a matrix $L$ such that $z - Ly$ is independent of $x$ and $y$ on $\mathbb{P}_{x,y,z}$. Then:

(i) $\mathbb{P}_{x,y,z}$ is a zero-mean, jointly Gaussian probability measure with the same covariance matrix $\Sigma_{x,y,z}$.

**Proof:** See [58, Lemma 3.2].

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

**Proof:** Observe

\[ \text{cov} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \Sigma_x & \Sigma_x A^T \\ A \Sigma_x & A \Sigma_x A^T + \Sigma_v \end{bmatrix} \quad (50) \]

Since it must be that $x_G \sim \mathcal{N}(0, \Sigma_x)$, introducing a matrix $R$ with full column rank such that $\Sigma_x = R R^T$, $x_G$ can be written as $x_G = R x_G$ with $x_G \sim \mathcal{N}(0, I)$. Since $(x_G, y_G)$ are jointly Gaussian, there exists a matrix $A$ such that

\[ y_G = A x_G + v, \quad v \sim \mathcal{N}(0, \Sigma_v) \quad (51) \]

where $v$ is independent of $x_G$. Thus

\[ \text{cov} \begin{bmatrix} x_G \\ y_G \end{bmatrix} = \begin{bmatrix} \Sigma_x & \Sigma_x A^T \\ A \Sigma_x & A \Sigma_x A^T + \Sigma_v \end{bmatrix} \quad (52) \]

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

\[ y_G = (A + S) x_G + v = (A + S) R x_G + v = A R x_G + v = A x_G + v, \quad v \sim \mathcal{N}(0, \Sigma_v). \]
Lemma 7: Let \( P_{x,y} \) be a zero-mean joint probability measure on \((X \times Y, B_X \otimes B_Y)\), \( X = \mathbb{R}^n \), \( Y = \mathbb{R}^m \), with a covariance matrix \( \Sigma_{x,y} \). Let \( G_{x,y} \) be a zero-mean Gaussian joint probability measure with the same covariance matrix \( \Sigma_{x,y} \). If there exists a subset \( C_y \subseteq \mathbb{R}^m \) with \( P_Y(C_y) > 0 \) such that \( P_y \) admits density for every \( y \in C_y \), then \( G_{x,y} \) admits density for every \( y \in supp(G_y) \).

Proof: Suppose \( P_{x,y}(x,y) \) and \( G_{x,y}(x,y) \) share a covariance matrix

\[
\Sigma_{x,y} = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix} > 0.
\]

Since \( G_{x,y} \) is a zero-mean Gaussian distribution, we know

\[
G_{x|y} \sim N(\hat{x}(y), \Sigma_c) \quad \forall y \in Y
\]

with \( \hat{x}(y) = \Sigma_{xy} \Sigma_{yy}^{-1} y \). The Moore–Penrose pseudo-inverse of \( \Sigma_{yy} \). To show contrapositive, assume that there exists \( y \in supp(G_y) \) such that \( G_{x|y}(x|y) \) does not admit a density. From (53), this means for every \( y \) in \( Y \), define a covariance matrix \( M(y) = f_x((x - \hat{x}(y))(x - \hat{x}(y))^\top dP_{x|y}(x|y). \)

Observe that

\[
\int_Y M(y) dP(y) = \int_{X \times Y} (x - \hat{x}(y))(x - \hat{x}(y))^\top dP(x,y)
\]

\[
= \int_{X \times Y} (x - \hat{x}(y))(x - \hat{x}(y))^\top dG(x,y)
\]

\[
= \Sigma_c.
\]

Since \( \Sigma_c \) is singular, there exists a full row rank matrix \( U \in \mathbb{R}^{r \times n}, 1 \leq r \leq n \) such that \( U \Sigma_c U^\top = 0 \). From (54), it follows that \( U M(y) U^\top = 0, P_Y \) a.e.. For every \( y \in Y \), define a subset \( C_{x|y} \subseteq \mathbb{R}^n \) by

\[
C_{x|y} = \{ x \in \mathbb{R}^n : U(x - \hat{x}(y)) = 0 \}.
\]

By construction, \( P_{x|y}(C_{x|y}) = 1, P_Y \) a.e.. However, clearly \( \mathbb{L}_x(C_{x|y}) = 0 \), where \( \mathbb{L}_x \) is the Lebesgue measure on \( \mathbb{R}^n \). Thus \( P_{x|y} \ll \mathbb{L}_x \) fails to hold for \( \mathbb{L}_y \) a.e.. Hence, \( P_{x|y} \) fails to admit a density \( \mathbb{L}_y \) a.e.. This is in contradiction to the assumption that there exists a subset \( C_y \subseteq \mathbb{R}^m \) with \( G_y(C_y) > 0 \) such that \( P_{x|y} \) admits density for every \( y \in C_y \).

C. Proof of Lemma 2

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

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

\[
U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}
\]

with \( U_1 \in \mathbb{R}^{p \times n}, U_2 \in \mathbb{R}^{(n-p) \times n}, p < n \) such that \( U \Sigma U^\top = diag(\Sigma_{zz}, 0) \), where \( \Sigma_{zz} \in \mathbb{R}^{p \times p} \) is a positive definite matrix. Notice that \( U_2 x = 0 \) for every \( x \in supp(G) \), and \( U_2 x \neq 0 \) for every \( x \in supp(G)^c \). Suppose \( supp(P) \subseteq supp(G) \) does not hold, i.e., there exists a closed set \( C \subseteq supp(G)^c \) such that \( P(C) > 0. \)

\[
U_2 \Sigma U_2^\top = \int U_2 x x^\top U_2^\top dP(x)
\]

\[
= \int_{supp(G)^c} U_2 x x^\top U_2^\top dP(x) + \int_{supp(G)^c} U_2 x x^\top U_2^\top dP(x)
\]

\[
= \int_{supp(G)^c} U_2 x x^\top U_2^\top dP(x)
\]

Since \( U_2 x \neq 0 \) for every \( x \in C \), the last expression is a non-negative 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 \( P(x^{T+1}, u^T) \) be a joint probability measure generated by a policy \( \gamma = \{ P(u|x^t, u^{t-1}) \}_{t=1}^T \) and (1).

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

Moreover, if \( I_p(x_t; u_t|u^{t-1}) < +\infty \) for all \( t = 1, \ldots, T \), then the following statements hold.

(b) For every \( t = 1, \ldots, T \),

\[
P(x_t|u^t) \ll P(x_t|u^{t-1}), \quad P(u^t) - a.e.,
\]

and

\[
I_p(x_t; u_t|u^{t-1}) = \int \log \left( \frac{dP(x_t|u^t)}{dP(x_t|x_t|u^{t-1})} \right) dP(x_t, u^t).
\]

(c) For every \( t = 1, \ldots, T \),

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

Moreover, the following identity holds \( P(x_t, u^t) - a.e.: \)

\[
\frac{dP(x_t|u^t)}{dP(x_t|x_t, u^t)} = \frac{dP(x_t|u^t)}{dP(x_t|x_t|u^{t-1})}.
\]

Proof:

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

(b) By definition of conditional mutual information, \( I_p(x_t; u_t|u^{t-1}) < +\infty \) requires \( I_p(x_t; u_t|u^{t-1}) < +\infty, P(u_t) - a.e.. \) For a fixed \( u^{t-1} \), \( I_p(x_t; u_t|u^{t-1}) < +\infty \) requires \( P(x_t|u^{t-1}) = \int \log \left( \frac{dP(x_t|u^t)}{dP(x_t|u^{t-1})} \right) dP(x_t, u^t) \) by definition of mutual information. By Lemma 4, this implies

\[
P(x_t|u^t) \ll P(x_t|u^{t-1}) \quad \text{and}
\]

\[
\frac{dP(x_t, u_t|u^{t-1})}{dP(x_t|x_t|u^{t-1})} = \frac{dP(x_t, u_t|u^{t-1})}{dP(x_t|x_t|u^{t-1})}
\]

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

\[
I_p(x_t; u_t|u^{t-1}) = \int \log \left( \frac{dP(x_t, u_t|x_t|u^{t-1})}{dP(x_t|x_t|u^{t-1})} \right) dP(x_t, u^t);
\]

\[
= \int \log \left( \frac{dP(x_t, u_t|x_t|u^{t-1})}{dP(x_t|x_t|u^{t-1})} \right) dP(x_t, u^t).
\]

(c) Let \( B_{X_t} \subseteq B_{X_{t+1}}, B_{X_{t+1}} \subseteq B_{X_{t+1}}, B_{U^t} \subseteq B_{U^t} \) be arbitrary Borel sets. Since both \( P(x_{t+1}|u^t) \) and \( P(x_{t+1}|x_t, u^t) \) are non-degenerate Gaussian probability measures, there exists a continuous map \( f : \mathcal{X}_t \times \mathcal{X}_{t+1} \times \mathcal{U}^t \rightarrow (0, +\infty) \) such that
\[ k(x_t, x_{t+1}, u^t) = \frac{dP(x_{t+1}|u^t)}{dP(x_{t+1}|x_t, u^t)}, \quad \text{or} \quad (58a) \]

\[ \int_{B_{x_{t+1}}} k(x_t, x_{t+1}, u^t) dP(x_{t+1}|x_t, u^t) = \int_{B_{x_{t+1}}} dP(x_{t+1}|u^t). \quad (58b) \]

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

\[ \int_{B_{x_t} \times B_{x_{t+1}} \times B_{u^t}} k dP(x_t, x_{t+1}, u^t) \]

in two different ways:

\[ (59) = \int_{B_{x_t} \times B_{x_{t+1}} \times B_{u^t}} k dP(x_t|x_{t+1}, u^t) dP(x_{t+1}, u^t); \quad (60a) \]

\[ (59) = \int_{B_{x_t} \times B_{x_{t+1}} \times B_{u^t}} k dP(x_t|x_{t+1}, u^t) dP(x_{t+1}|u^t). \quad (60b) \]

\[ = \int_{B_{x_t} \times B_{u^t}} \int_{B_{x_{t+1}}} dP(x_{t+1}|u^t) dP(x_t|u^t) dP(u^t) \quad (60c) \]

\[ = \int_{B_{x_t} \times B_{u^t}} \int_{B_{x_{t+1}}} dP(x_{t+1}|u^t) dP(x_t|x_{t+1}, u^t) dP(u^t) \quad (60d) \]

\[ = \int_{B_{x_t} \times B_{u^t}} \mathbb{P}_{x_t}(u^t|B_{x_t}) dP(x_{t+1}, u^t). \quad (60f) \]

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

\[ \int_{B_{x_t}} k dP(x_t|x_{t+1}, u^t) = \mathbb{P}_{x_t}(u^t|B_{x_t}). \quad (61) \]

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 \( P(x_{t+1}, u^t) - a.e., \) which establishes the second claim.

\[ \int_{B_{x_t}} \frac{dP(x_{t+1}|u^t)}{dP(x_{t+1}|x_t, u^t)} \frac{dP(x_{t+1}|x_{t+1}, u^t)}{dP(x_{t+1}|x_t, u^t)} dP(x_t|u^t-1) \]

\[ = \int_{B_{x_t}} k(x_t, x_{t+1}, u^t) dP(x_{t+1}, u^t) \]

\[ = \mathbb{P}_{x_t}(u^t|B_{x_t}). \]

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

**Lemma 10:** Let \( \mathbb{P}(x_{t+1}, u^T) \) be a joint probability measure generated by a policy \( \gamma_P = \{ P(u_t|x_t, u^t-1) \}_{t=1}^T \) and (1), and \( G(x_{t+1}, u^T) \) 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) \( u^t-1-(x_t, u_t)-x_{t+1} \) form a Markov chain in \( G \). 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 \( \mathcal{N}(Ax_t + Bu_t, W_t) \).

The result of Lemma 9 (c) is used in the third equality. In the final step, 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 \( G \). Here we only show the case for (62b), but
a similar technique is also applicable to (62a) and (62c).

\[
\int \log \left( \frac{dP(x_t|x_{t-1}, u')} {dP(x_t|x_{t-1}, u')} \right) dP(x_{t-1}, u')
\]

\[
= \int \log \left( \frac{dG(x_t|x_{t-1}, u')} {dG(x_t|x_{t-1}, u')} \right) dG(x_{t-1}, u') \tag{63a}
\]

\[
= \int \log \left( \frac{dP(x_t|x_{t-1}, u')} {dG(x_t|x_{t-1}, u')} \right) dP(x_{t-1}, u')
\]

\[
= \int \log \left( \frac{dG(x_t|x_{t-1}, u')} {dG(x_t|x_{t-1}, u')} \right) dP(x_{t-1}, u') \tag{63b}
\]

\[
= \int \log \left( \frac{dG(x_t|x_{t-1}, u')} {dG(x_t|x_{t-1}, u')} \right) dP(x_{t-1}, u')
\]

\[
= \int \int \log \left( \frac{dP(x_t|x_{t-1}, u')} {dG(x_t|x_{t-1}, u')} \right) dP(x_t|x_{t-1}, u')
\]

\[
\geq 0.
\]

Due to Lemma 10, \( \log \frac{dG(x_t|x_{t-1}, u')}{dG(x_t|x_{t-1}, u')} \) in (63a) is a quadratic function of \( x_t+1 \) and \( u' \) everywhere on \( \text{supp}(G(x_t+1, u')) \). This is also the case everywhere on \( \text{supp}(P(x_t+1, u')) \) since it follows from Lemma 8 that \( \text{supp}(P(x_t+1, u')) \subseteq \text{supp}(G(x_t+1, u')) \). Since \( P \) and \( G \) have the same covariance, \( dG(x_t+1, u') \) can be replaced by \( dP(x_t+1, u') \) in (63b). In (63c), the chain rule of the Radon-Nikodym derivatives is used invoking that \( P(x_t|x_{t-1}, u'=1) = G(x_t|x_{t-1}, u'=1) \) from Lemma 10 (a).

D. Proof of Lemma 3

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

\[
\frac{dQ(x_{k+1}, u^k)} {dQ(x_k, x_{k+1}, u^k)}
\]

\[
= \int \frac{dP(x_{k+1}|x_k, u_k) dQ(x_k, u^k)} {dQ(x_k, x_{k+1}, u^k)}
\]

\[
= \int \frac{dP(x_{k+1}|x_k, u_k) dQ(x_k, u^k)} {dQ(x_k, x_{k+1}, u^k)}
\]

\[
= \int \frac{dP(x_{k+1}|x_k, u_k) dQ(x_k, u^k)} {dQ(x_k, x_{k+1}, u^k)}
\]

\[
= \int \frac{dG(x_{k+1}|x_k, u_k) dG(x_k, u^k)} {dQ(x_k, x_{k+1}, u^k)}
\]

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 \( G(x_k, u^{k-1}) = Q(x_k, u^{k-1}) \) 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 \( \gamma_\infty = \{P(u_t|\mathcal{X}_t, u^{t-1})\}_{t \in \mathbb{N}} \), there exists a linear-Gaussian policy \( \gamma_\mathcal{G} = \{Q(u_t|\mathcal{X}_t, u^{t-1})\}_{t \in \mathbb{N}} \) which is at least as good as \( \gamma_\infty \). 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 \( \{P_{ti}, \Pi_t\}_{t \in \mathbb{N}} \) similar to (12) in which \( \sum_{t=1}^\infty \) is replaced by \( \lim_{t \to \infty} \frac{1}{t} \sum_{t=1}^T \) and parameters \( \beta_t, W_t, S_t, \Theta_t \) are time-invariant. It is shown in [60] that the optimality of this SDP over \( \{P_{ti}, \Pi_t\}_{t \in \mathbb{N}} \) is attained by a time-invariant sequence \( P_{ti} = P, \Pi_t = \Pi \forall t \in \mathbb{N} \), where \( P \) and \( \Pi \) are the optimal solution to (18).

E. 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

\[
v^*(A, W) =
\begin{cases}
\inf_{P, \Pi} \frac{1}{2} \log \det \Pi - \frac{1}{2} \log \det W & \text{s.t. } \Pi > 0, P \succeq APA^T + W, \left[ P - \Pi \right] APA^T + W \succeq 0.
\end{cases}
\]

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).

\[
v^*(A, W) =
\begin{cases}
\sup_{X, Y} \frac{1}{2} \log \det X_{11} - \frac{1}{2} \log \det X_{22} + Y_{11} + Y_{12} + Y_{21}, A \succeq 0, Y \succeq 0, X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}
\end{cases}
\]

The primal problem (65) can be also rewritten as

\[
v^*(A, W) =
\begin{cases}
\inf_{P} \frac{1}{2} \log \det (APA^T + W) - \frac{1}{2} \log \det P & \text{s.t. } P \succeq APA^T + W, P \in \mathbb{S}^n_+
\end{cases}
\]

\[
= \begin{cases}
\inf_{P, C, V} -\frac{1}{2} \log \det (I - V^{-\frac{1}{2}} CPC^T V^{-\frac{1}{2}}) & \text{s.t. } P \succeq APA^T + W, P = C^T C \succeq 0, C \in \mathbb{R}^{n \times n}
\end{cases}
\]

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

\[
\frac{1}{2} \log \det (APA^T + W) - \frac{1}{2} \log \det P
\]

\[
= -\frac{1}{2} \log \det (APA^T + W) - \frac{1}{2} \log \det P
\]

\[
= -\frac{1}{2} \log \det (P - C^T C) - \frac{1}{2} \log \det P
\]

\[
= -\frac{1}{2} \log \det (I - P^T C^T V^{-1} C P^T)
\]

\[
= -\frac{1}{2} \log \det (I - V^{-\frac{1}{2}} CPC^T V^{-\frac{1}{2}})
\]
The last step follows from Sylvester’s determinant theorem.

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

2) **Case 2:** When all eigenvalues of $A$ satisfy $|\lambda| < 1$: 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 $P^* > 0$ is the unique solution to the Lyapunov equation $P^* = APA^T + W$.

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

**Lemma 11:** Suppose $V > 0$ and $(A,C)$ is a detectable pair and $0 < W_1 \leq W_2$. Then, we have $P \preceq Q$ where $P$ and $Q$ are the unique positive definite solutions to

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

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

**Proof:** Consider Riccati recursions

$$\hat{P}_{t+1} = A\hat{P}_t A - A\hat{P}_t C^T (C\hat{P}_t C^T + V)^{-1} C\hat{P}_t A + W_1 \quad (71)$$

$$\hat{Q}_{t+1} = A\hat{Q}_t A - A\hat{Q}_t C^T (C\hat{Q}_t C^T + V)^{-1} C\hat{Q}_t A + W_2 \quad (72)$$

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

Using the above lemma, we obtain the following result.

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

**Proof:** Due to the characterization (68) of $v^*(A,W_2)$, there exist $Q > 0, V > 0, C \in \mathbb{R}^{n \times n}$ such that $v^*(A,W_2) = -\frac{1}{2} \log \det (V^{-\frac{1}{2}} CQC^T V^{-\frac{1}{2}})$ and $Q^{-1} - (AQA^T + W_2)^{-1} = C^T V^{-1} C$. (73)

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

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

Moreover, we have $P \preceq Q$ since

$$0 < Q^{-1} = \hat{Q}^{-1} + C^T V^{-1} C \preceq \hat{P}^{-1} + C^T V^{-1} C = P^{-1}$$

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

$$v^*(A,W_2) = -\frac{1}{2} \log \det (I - V^{-\frac{1}{2}} CQC^T V^{-\frac{1}{2}})$$

$$\geq -\frac{1}{2} \log \det (I - V^{-\frac{1}{2}} CPC^T V^{-\frac{1}{2}})$$

$$\geq v^*(A,W_1).$$

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

$$v^*(A,W) \leq v^*(A,\delta I_n)$$

$$\leq v^*(A_1,\delta I_{n_1}) + v^*(A_2,\delta I_{n_2}) = \sum_{\lambda \in \sigma(A_1)} \log |\lambda|.$$
Since we are assuming the Kalman filter is also known as the whitening filter since it is equivalent to the “decision policy” block in the original problem (38).

Lemma 15: For every $t = 1, \cdots, T$,
\[
\mathbb{E}\|\tilde{x}_{t+1}\|^2_{Q_t} = \text{Tr}(Q_t \tilde{P}_{t+1|t+1}) + \mathbb{E}\|\tilde{x}_{t+1}\|^2_{Q_t}.
\]

Proof: Observe
\[
\mathbb{E}\|\tilde{x}_{t+1}\|^2_{Q_t} = \mathbb{E}\|\tilde{x}_{t+1} + \tilde{x}_t\|^2_{Q_t} = \mathbb{E}\|\tilde{x}_{t+1}\|^2_{Q_t} + 2\text{E}\tilde{x}_{t+1} Q_t (\tilde{x}_{t+1} - \tilde{x}_t).
\]

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

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

\[
\min_{\gamma \in \Gamma} \sum_{t=1}^{T} I_{\gamma}(y^t; u^t|u^{t-1})
\]
\[\text{s.t. } \sum_{t=1}^{T} \mathbb{E}_{\gamma}(\|x_{t+1}\|^2_{Q_t} + \|u_t\|^2_{R_t}) \leq D.
\]

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

\[
\min_{\gamma \in \Gamma} \sum_{t=1}^{T} I_{\gamma}(\tilde{x}^t; u^t|u^{t-1})
\]
\[\text{s.t. } \sum_{t=1}^{T} \mathbb{E}_{\gamma}(\|\tilde{x}_{t+1}\|^2_{Q_t} + \|u_t\|^2_{R_t}) + \text{Tr}(Q_t \tilde{P}_{t+1|t+1}) \leq D.
\]
Now, notice that the terms $\text{Tr}(Q_tP_{t+1|t+1})$ do not depend on $\tilde{\gamma}$. Thus, by rewriting the pre-Kalman filter (39) equation as
\[
\tilde{x}_{t+1} = A_t\tilde{x}_t + B_tu_t + \psi_t
\]
and considering (80) as a new “system” with white Gaussian process noise $\psi_t \sim \mathcal{N}(0, \Psi_t)$, problem (79) can be written as
\[
\min_{\tilde{\gamma} \in \tilde{\gamma}^*} I_t(\tilde{x}^T \to u^T) \quad (81a)
\]
s.t. \( J_t(\tilde{x}^T_{t+1}, u^T_t) \leq \tilde{D} \quad (81b) \)
where \( \tilde{D} = D - \sum_{t=1}^T \text{Tr}(Q_tP_{t+1|t+1}) \). 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


