Minimum-Information LQG Control
Part II: Retentive Controllers
Roy Fox† and Naftali Tishby†

Abstract— Retentive (memory-utilizing) sensing-acting agents, when they are distributed or power-constrained, operate under limitations on the communication between their sensing, memory and acting components. This requires them to trade off the external cost that they incur with the capacity of their communication channels. This is a sequential rate-distortion problem of minimizing the rate of information required for the controller’s operation, under a constraint on its external cost. In this paper we formulate this problem with a simple information term, by viewing the memory reader as one more sensor, and the memory writer as one more actuator. This allows us to reduce the bounded retentive control problem to the memoryless one, studied in Part I of this work [1]. We further investigate the form of the resulting optimal solution, and demonstrate its interesting phenomenology.

I. INTRODUCTION

In a feedback-control system, the internal state of the agent interacts with the external state of the world, through sensors that pay attention to the agent’s environment, and actuators that apply intention to it, in a perception-action cycle [2].

In Part I† of this work [1], we discussed how communication from the sensor to the actuator is central to the agent’s ability to act upon the perceived information. As devices become smaller and more ubiquitous, power efficiency and physical restrictions dictate that communication becomes a limiting factor in the agent’s operation.

A related but often overlooked resource is memory bandwidth. We can think of memory as a communication channel from the past internal state of the controller to its future internal state. When memory resources are remote, communication constraints apply to them as well. Even local memory is limited by its capacity to store information, and by the capacity of the internal communication channels to and from the memory components. This limitation is evidenced by the hierarchical design of memory in modern digital computers, which places larger capacity on the channels to closer, but smaller, cache memory components [3].

Classic optimal control theory [4] was unconcerned with the costs and the limitations of communicating the information needed for the controller’s operation. In the past two decades, however, a large body of research has been dedicated to this issue (see references in: [5]–[8]).

The perception-action cycle between a controller and its environment (Fig. 1) consists of multiple channels, and the capacity of any of them can be limited. Accordingly, various information rates can be considered. Our guiding principle in this work is to measure the information complexity of the controller’s internal representation, by asking “How much information does the controller have on the past?”. The past is informative of the future [9], and some information in past observations is useful in controlling the future. We therefore seek a trade-off between the external cost incurred by the system, and the internal cost of the communication resources spent by the controller in reducing that external cost. This trade-off is often formulated as an optimization problem, where one cost is constrained and the other minimized.

Memoryless controllers were discussed in Part I of this work [1]. When the controller has no internal memory, it can only attend, perhaps selectively, to its most recent input observation. The degree of this attention, measured by the amount of Shannon information about the input observation that is utilized in the output control, is a lower bound on the required capacity of the communication channel between the controller’s sensor and its actuator.

When the controller is retentive (memory-utilizing), it does maintain an internal memory state, and that state can have information on more than the most recent observation. In a sense, we can consider the reader of the memory state to be one more sensor, and the writer of the memory state one more actuator. The controller receives information of the past through both memory and sensory channels (Fig. 2), and the amount of information that it keeps of the past is a lower bound on the total capacity of both these channels [10].

We refer the reader to Part I [1, Sec. I] for further discussion of related prior work.

In this paper we make two contributions. First, we present a method for the design of controllers that are optimal under a constraint on both their memory and sensory channel capacity. To our knowledge, this is the first explicit treatment of the channel capacity of the memory process in the context of continuous state-space systems.

Second, we provide a reduction from the problem of bounded retentive control, to the problem of bounded memoryless control. This reduction is conceptually convenient and constructive, allowing us to treat both problems using the same framework, and providing insight into the structure of the optimal retentive controller.

In Sec. II we define the LQG task and restate the results of Part I. In Sec. III we present the retentive control model, its reduction to memoryless control, and the structure of the resulting optimal solution. In Sec. IV we illustrate our results with an example.
II. PRELIMINARIES

A. Control task

We consider the same closed-loop control problem detailed in Part I [1, Sec. II]. In time $t$, a plant in state $x_t \in \mathbb{R}^n$ emits an observation $y_t \in \mathbb{R}^k$, then takes in a control input $u_t \in \mathbb{R}^l$ and undergoes a stochastic state transition. We focus on discrete-time systems with linear dynamics, Gaussian noise and quadratic cost rate (LQG). For simplicity, all elements are taken to be homogeneous, i.e. centered at the origin, and time-invariant. Our results hold without these assumptions, with the appropriate adjustments.

Definition 1: A linear-Gaussian time-invariant (LTI) plant $(A, B, C, \Sigma_x, \Sigma_e)$ has state dynamics

$$x_{t+1} = Ax_t + Bu_t + \xi_t; \quad \xi_t \sim \mathcal{N}(0, \Sigma_x),$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times l}$, $0 \leq \Sigma_x \in \mathbb{S}^+_n$, and $\xi_t$ is independent of $(x^t, y^t)$. The observation dynamics are

$$y_t = Cx_t + \epsilon_t; \quad \epsilon_t \sim \mathcal{N}(0, \Sigma_e),$$

where $C \in \mathbb{R}^{k \times n}$, $\Sigma_e \in \mathbb{S}^+_k$, and $\epsilon_t$ is independent of $(y^{t-1}, u^{t-1}, x^t)$.

Definition 2: A linear-quadratic-Gaussian (LQG) task $(A, B, C, \Sigma_x, \Sigma_e, Q, R)$ involves a LTI plant and the cost rate

$$J_t = \frac{1}{2}(x_t^T Q x_t + u_t^T R u_t),$$

where $Q \in \mathbb{S}^+_n$ and $R \in \mathbb{S}^+_l$. The task is to achieve a low long-term average expected cost rate, with respect to the distribution induced by the plant and the controller $\pi$

$$J_\pi = \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_\pi[J_t].$$

As motivated in Part I, we are particularly interested in linear-Gaussian time-invariant (LTI) controllers, which induce, jointly with a LTI plant, a stationary Gaussian process, independent of any initial conditions. With $\Sigma_x \in \mathbb{S}^+_n$, $\Sigma_y \in \mathbb{S}^+_k$, and $\Sigma_u \in \mathbb{S}^+_l$, the stationary covariances of the state, the observation and the control, respectively, we have

$$\Sigma_y = C \Sigma_x C^T + \Sigma_e,$$

and the reverse relation

$$x_t = K y_t + \kappa_t; \quad \kappa_t \sim \mathcal{N}(0, \Sigma_\kappa)$$

$$K = \Sigma_x C^T \Sigma_y^{-1}$$

$$\Sigma_\kappa = \Sigma_x - \Sigma_x C^T \Sigma_y^{-1} C \Sigma_x,$$

with $^{-1}$ the Moore-Penrose pseudoinverse. Assuming that the process has mean 0, the stationary expected cost rate is

$$J_\pi = \frac{1}{2}(tr(Q \Sigma_x) + tr(R \Sigma_u)).$$

B. Bounded memoryless control

In this section we restate in algorithmic form the main result of Part I [1, Sec. IV].

Definition 3: A memoryless linear-Gaussian time-invariant (LTI) controller has control law of the form

$$u_t = H y_t + \eta_t; \quad \eta_t \sim \mathcal{N}(0, \Sigma_\eta),$$

where $H \in \mathbb{R}^{l \times k}$, $\Sigma_\eta \in \mathbb{S}^+_l$, and $\eta_t$ is independent of $y_t$.

The controller is bounded, and operates under limits on its capacity to process the observation and produce the control. Namely, with the Shannon information rate

$$I_t = I[y_t; u_t] = \mathbb{E} \left[ \log \frac{f(y_t, u_t)}{f(y_t) f(u_t)} \right],$$

where $f$ denotes the various probability density functions, as indicated by their arguments, we are interested in a LTI controller $\pi$ that minimizes the long-term average rate

$$I_\pi = \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} I_t,$$

under the constraint that it achieves some guarantee level $c$ expected cost rate.

Problem 1: Given a LQG task, the bounded memoryless LTI controller optimization problem is

$$\min_{\pi} I_\pi$$

s.t. $J_\pi \leq c,$

with $I_\pi$ as in (4), where

$$I_t = I[y_t; u_t],$$

and with $u_t$ as in (3).

To solve the optimization problem, we consider the minimum mean square error (MMSE) estimators

$$\hat{x}_y = \mathbb{E}[x_t | y_t] = Ky_t$$

$$\hat{x}_u = \mathbb{E}[x_t | u_t] = \Sigma_{x;u} \Sigma_u^{-1} u_t,$$

for the state given the observation and the control, respectively. Since $\hat{x}_u$ is a sufficient statistic of $u_t$ for $x_t$, we can reverse their causality, basing $u_t$ on $\hat{x}_u$, instead of vice versa. This puts the control law in the form

$$\hat{x}_y = Ky_t$$

$$\hat{x}_u = W \hat{x}_y + \omega_t; \quad \omega_t \sim \mathcal{N}(0, \Sigma_\omega)$$

$$u_t = L \hat{x}_u.$$

The optimal memoryless controller satisfies the conditions of Theorem 1 in Part I, Sec. IV-A, restated below in algorithmic form. To numerically find the optimal solution, we can interpret these conditions as update equations. We can iteratively apply these updates, until a fixed point is reached.

We split the equations into two parts, a forward iteration (Algorithm 1) updating the marginal distributions and the inference policy, and a backward iteration (Algorithm 2) updating the cost-to-go and the control policy. We can alternate between Algorithms 1 and 2, and iterate until the solution converges to a fixed point of the equations.

III. BOUNDED RETENTIVE CONTROLLERS

A. Control model

In this section we discuss retentive (memory-utilizing) controllers with bounded communication resources. A retentive controller has an internal memory state $z_t$ in some space $Z$. The memory allows the controller to output a control that
Algorithm 1 Forward iteration

\begin{algorithm}[H]
\begin{algorithmic}
\Function{FORWARD}{$\Sigma_{x}, \Sigma_{\hat{x}_u}, L, N; \beta$}
\State Update
\State $\Sigma_x \leftarrow (A + BL) \Sigma_{\hat{x}_u} (A + BL)^\top + A (\Sigma_x - \Sigma_{\hat{x}_u}) A^\top + \Sigma_\xi$
\State $\Sigma_y \leftarrow C \Sigma_x C^\top + \Sigma_\xi$
\State $K \leftarrow \Sigma_x C^\top \Sigma_y^{-1}$
\State $\Sigma_{\hat{x}_u} \leftarrow K \Sigma_y K^\top$
\State $V, \Lambda \leftarrow \text{EVD}(\Sigma_{\hat{x}_u}^{1/2} N \Sigma_{\hat{x}_u}^{1/2})$
\State $D \leftarrow \text{diag} \left\{ 1 - \beta^{-1} \lambda_i^{-1} \quad \lambda_i > \beta^{-1} \right\}$
\State $\Sigma_{\hat{x}_u} \leftarrow \Sigma_{\hat{x}_u}^{1/2} VDV^\top \Sigma_{\hat{x}_u}^{1/2}$
\EndFunction
\end{algorithmic}
\end{algorithm}

Indirectly depends on past input observations, and not only on the most recent observation. The controller takes as input an observation $y_t$ and outputs a control $u_t$, while making a memory state transition from $z_{t-1}$ to $z_t$. Thus, in each time step, there are two inputs, $z_{t-1}$ and $y_t$, and two outputs, $z_t$ and $u_t$.

**Definition 4:** A controller is retentive if it satisfies the following independence properties:

1) The memory state depends only on the previous memory state and the current observation, that is, $z_t$ is independent of $(z_{t-2}, y_{t-1}, u_{t-1}, x')$ given $z_{t-1}$ and $y_t$.

2) The control depends only on the memory state, that is, $u_t$ is independent of $(z_{t-1}, u_{t-1}, x', y')$ given $z_t$.

A system including a retentive controller satisfies the Bayesian network in Fig. 1.

As motivated in Part I for the memoryless case, we are particularly interested in controllers where both the memory state update and the control are linear-Gaussian time-invariant (LTI). The conditions under which such controllers are optimal for our bounded control problem are beyond our current scope.

**Definition 5:** A retentive linear-Gaussian time-invariant (LTI) controller has memory state space that is a vector space $Z = \mathbb{R}^d$, and control law of the form

\begin{align}
\begin{aligned}
z_t &= F z_{t-1} + Gy_t + \zeta_t, \\
u_t &= L z_t + \nu_t;
\end{aligned}
\end{align}

where $F \in \mathbb{R}^{d \times d}$, $G \in \mathbb{R}^{d \times k}$, $\Sigma_{\xi} \in \mathbb{S}^d_+$, $L \in \mathbb{R}^{m \times d}$, $\Sigma_{\nu} \in \mathbb{S}^m_+$, $\zeta_t$ is independent of $(z_{t-1}, y_t)$, and $\nu_t$ is independent of $z_t$.

We are interested in reducing the information complexity of implementing this controller. To measure this complexity, we consider the capacity of a memoryless communication channel from the sensor-reader to the actuator-writer (Fig. 2). The encoder and the decoder themselves are memoryless, but the memory component has perfect fidelity, making everything written by the actuator available for the sensor to read in the next step.

If $Z = \{0, 1\}^r$ is the set of binary strings of length $r$, then the controller can process at most $r$ bits of information per time step

\[ I[z_{t-1}, y_t; z_t, u_t] = \mathbb{H}[z_{t-1}, y_t] \leq \mathbb{H}[z_t] \leq r \log 2. \]

Similarly to the memoryless case (Part I [1, Sec. III-B]), the information rate is generally not a tight lower bound on the capacity of a discrete memory. But here again, if the controller is LTI, then there exists a perfectly matched memoryless additive Gaussian noise channel. As shown in the Supplementary Material (SM), Appx. I, the capacity of this channel optimally equals the information rate $\mathbb{H}[z_{t-1}, y_t; z_t, u_t]$, and a constraint on the information rate is equivalent to a constraint on the power available for transmission on the channel.

The retentive controller optimization problem is therefore similar to Problem 1, but with the information rate including both the memory and the sensory channels.
**Problem 2:** Given a LQG task, the bounded retentive LTI controller optimization problem is

\[
\min_{\pi} \mathcal{I}_\pi \\
\text{s.t.} \quad J_\pi \leq c,
\]

with \(\mathcal{I}_\pi\) as in (4), where

\[
\mathcal{I}_t = \mathbb{I}[z_{t-1}, y_t; z_t, u_t],
\]

and with \(z_t\) and \(u_t\) as in (6).

Note that here there is no additional constraint or cost on the precision of \(u_t\) given \(z_t\), implying that optimally \(\Sigma_u = 0\).

There is an interesting connection between the retentive information rate \(\mathcal{I}_\pi\), with \(\mathcal{I}_t\) as in (7), and the long-term average of the directed information rate [11], [12], defined by

\[
\mathbb{I}[\{y_t\} \to \{z_t\}] = \limsup_{T \to \infty} \frac{1}{T} \mathbb{I}[y^T \to z^T] = \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}[y^t; z_t|z^{t-1}].
\]

By the independence properties of the retentive controller, and by the chain rule for information [13], we have

\[
\mathbb{I}[z_{t-1}, y_t; z_t, u_t] = \mathbb{I}[z_{t-1}, y_t; z_t] = \mathbb{I}[z^{t-1}, y_t; z_t] = \mathbb{I}[z^{t-1}; z_t] + \mathbb{I}[y^t; z_t|z^{t-1}].
\]

We can thus define the following extension of the concept of directed information.

**Definition 6:** The retentive directed information from the sequence of observations \(y^T\) to the sequence of memory states \(z^T\) is

\[
\mathbb{I}[y^T \to z^T] = \sum_{t=1}^{T} \mathbb{I}[z^{t-1}, y^t; z_t].
\]

Since \(\mathbb{I}[y^T \to z^T] \geq \mathbb{I}[y^T \to z^T]\), the retentive directed information rate is always a tighter lower bound on the capacity of the channel in Fig. 2. Despite the apparent similarity to Fig. 2 in [12], notice that their encoder and decoder have unlimited memory of \(z^T\) and \(u^T\). This justifies their use of directed information, regardless of the residual term \(\mathbb{I}[z^{t-1}; z_t]\) being infinite in their optimal controller.

Some further properties of the retentive directed information can be found in the SM, Appx. VI.

**B. Reduction to memoryless controllers**

We can analyze the bounded retentive control problem (Problem 2) directly using the same tools developed in Part I [1, Sec. IV-A]. Fortunately, there is no need to repeat that entire treatment, as a simple and insightful reduction will allow us to reuse the results already obtained there for the bounded memoryless control problem (Problem 1).

We start by reformulating the problem. The following relaxation, and Lemma 1 that shows its equivalence to the original problem, allow us to reverse the causality between \(u_t\) and \(z_t\). We need a new notation for the resulting time-shifted memory state sequence, and define for each \(t\)

\[
m_t = z_{t-1}.
\]

**Definition 7:** A retentive controller is relaxed if \(u_t\) is not required to be independent of \((m_t, y_t)\) given \(m_{t+1}\). Thus the relaxed controller satisfies the Bayesian network in Fig. 3, and its control law is given by \(\pi(u_t, m_{t+1}|m_t, y_t)\).

**Lemma 1:** The relaxed controller optimization problem is equivalent to the original Problem 2.

**Proof:** The following proof does not assume that the controller is linear-Gaussian, and holds for the LTI controller as a special case.

Let \(\pi\) be a controller satisfying the Bayesian network in Fig. 3. We construct a controller \(\tilde{\pi}\) with \(\tilde{z}_t = (u_t, m_{t+1})\) for each \(t\), such that

\[
\tilde{\pi}(\tilde{z}_t|\tilde{z}_{t-1}, y_t) = \pi(u_t, m_{t+1}|m_t, y_t) \quad \tilde{\pi}(u_t|\tilde{z}_t) = \delta_{\tilde{z}_t = (u_t, \cdot)}.
\]

This controller satisfies the Bayesian network in Fig. 1, and

\[
\mathbb{I}_\pi(\tilde{z}_{t-1}, y_t; \tilde{z}_t, u_t) = \mathbb{I}_\pi((u_{t-1}, m_t), y_t; (u_t, m_{t+1})) = \mathbb{I}_\pi(m_t, y_t; u_t, m_{t+1}).
\]

We have constructed a controller \(\tilde{\pi}\) that is feasible for the unrelaxed Problem 2, while having the same performance as the relaxed controller \(\pi\), since it induces a stochastic process with the same distribution and information rate.

The structure in Fig. 3 can now be redrawn as in Fig. 4. Comparing this Bayesian network to the one in Part I, Fig. 2, we have clearly reduced the bounded retentive control problem to a special case of the bounded memoryless control problem, as stated formally in the following lemma.

**Lemma 2:** The bounded retentive LTI controller optimization problem (Problem 2) for the LQG task \(\langle A, B_x, B_u, C_y, x, y, \Sigma_x, \Sigma_y, Q_x, R_u \rangle\) is equivalent to the bounded memoryless LTI controller optimization problem.
(Problem 1) for the LQG task \( \langle A, B, C, \Sigma_x, \Sigma_x, Q, R \rangle \), where
\[
A = \begin{bmatrix} A_x & 0 \\ 0 & 0 \end{bmatrix}; \quad B = \begin{bmatrix} B_{x,u} & 0 \\ 0 & I \end{bmatrix}; \quad C = \begin{bmatrix} C_{y,x} & 0 \\ 0 & 0 \end{bmatrix}
\]
\[
\Sigma_x = \begin{bmatrix} \Sigma_x & 0 \\ 0 & \Sigma_x \end{bmatrix}; \quad \Sigma_x = \begin{bmatrix} \Sigma_x & 0 \\ 0 & 0 \end{bmatrix}
\]
\[
Q = \begin{bmatrix} Q_x & 0 \\ 0 & 0 \end{bmatrix}; \quad R = \begin{bmatrix} R_u & 0 \\ 0 & 0 \end{bmatrix}.
\]
Here all matrices are extended by \( d \) rows and \( d \) columns.

**Proof:** Given the retentive control stochastic process \( \{x_t, m_t, y_t, u_t\} \), we consider the memoryless control stochastic process \( \{\tilde{x}_t, \tilde{y}_t, \tilde{u}_t\} \) with
\[
\tilde{x}_t = \begin{bmatrix} x_t \\ m_t \end{bmatrix}; \quad \tilde{y}_t = \begin{bmatrix} y_t \\ m_t \end{bmatrix}; \quad \tilde{u}_t = \begin{bmatrix} u_t \\ m_{t+1} \end{bmatrix}.
\]
The dynamics for this process can easily be seen to be given by (1), (2), with \( A, B, C, \Sigma_x \) and \( \Sigma_x \) as in the lemma. The cost rate applies only to the \( x_t \) and \( u_t \) parts
\[
J_t = \frac{1}{2} \begin{bmatrix} x_t^T \\ m_t \\ u_t^T \\ m_{t+1} \end{bmatrix}^T \begin{bmatrix} Q_x & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & R_u & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ m_t \\ u_t \\ m_{t+1} \end{bmatrix}.
\]
The information rate is
\[
\mathcal{I}_t = \mathbb{H}[\tilde{y}_t; \tilde{u}_t] = \mathbb{H}[m_t, y_t; u_t, m_{t+1}],
\]
where the left-hand side is taken as in (5) and the right-hand side as in (7), as required.

**C. Structure of the optimal solution**

We can substitute the form of the reduction in Lemma 2 into the optimal solution in Sec. II-B, to study more explicitly the structure of the optimal solution in the retentive case. The detailed derivations can be found in the SM, Appx. VII.

For the backward process, it is useful to borrow notation from the forward process, and denote
\[
S = \begin{bmatrix} S_x & S_{x,m} \\ S_{m,x} & S_m \end{bmatrix};
\]
\[
S_{x,m} = S_x - S_{x,m} S_{m,m}^{-1} S_{m,x} \\ S_{u,m} = R + B^T S_{x,m} B.
\]

Then we can find the feedback gain
\[
L = - (R + B^T S B)^{-1} B^T S A 
\]
\[
= \begin{bmatrix} L_{u,x|m} \\ -S_{u,m}^T S_{m,m} (A_x + B_{x,u} L_{u,x|m}) \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix},
\]
with a memory-conditioned form of the classic feedback gain
\[
L_{u,x|m} = -S_{u,m}^T B_{x,u} S_{x,m} A_x.
\]
The memory-conditioned cost reduction matrix is
\[
N = L^T (R + B^T S B) L = \begin{bmatrix} N_{x|m} & 0 \\ 0 & 0 \end{bmatrix},
\]
with
\[
N_{x|m} = A_x^T (S_x - S_{x,m} + S_{x,m} B_{x,u} S_{u,m} B_{x,u}^T S_{x|m}) A_x,
\]
\[
\text{implying that rank}(D) \leq \text{rank}(N) \leq n.
\]
The \( d \) rightmost columns in (8) are 0, implying that \( \tilde{u}_t \) depends only on the estimator \( \tilde{x}_{\hat{m}_t} = E[x_t|\tilde{u}_t] \) of \( x_t \), and not on an estimator of \( m_t \). Since \( \tilde{x}_{\hat{y}_t} \) is a sufficient statistic of \( \tilde{y}_t \) for \( x_t \), we also have
\[
x_t = \tilde{x}_{\hat{y}_t} - \tilde{y}_t - \tilde{x}_{\hat{u}_t} - \tilde{x}_{\hat{u}_t} - \tilde{u}_t,
\]
with
\[
\tilde{x}_{\hat{y}_t} = E[x_t|\tilde{y}_t] = \begin{bmatrix} x_t \\ m_t \end{bmatrix},
\]
\[
\tilde{y}_t = E[y_t|\tilde{m}_t] = \begin{bmatrix} y_t \\ m_t \end{bmatrix}.
\]
We conclude that we need only consider the estimator \( \tilde{x}_{\hat{y}_t} \), which is obtained from the observation \( \tilde{y}_t \) using
\[
K = \Sigma_x C^T \Sigma_{m|m}^T = [K_{x|y|m}(I - K_{x|y|m} C_{y|x}) \Sigma_{x,m} \Sigma_{y|m}],
\]
is the Kalman gain, which performs optimal inference in the classic LQG task.

Crucially, we see that \( \tilde{x}_{\hat{y}_t} \) depends on \( m_t \) only through
\[
\tilde{x}_{m_t} = E[x_t|m_t] = \Sigma_{x|m} \Sigma_{m|m}^T m_t.
\]
This implies that, for a controller \( \pi \), we can design an equivalent controller \( \pi' \) whose memory state is the MMSE estimator \( m_t' = \hat{x}_{m_t} \). The feedback gain for \( \pi' \) is
\[
L' = \begin{bmatrix} I & 0 \\ 0 & \Sigma_{x|m}^T \Sigma_{m|m}^T \end{bmatrix} L.
\]
Note that, since \( m_t' \) is a sufficient statistic of \( m_t \) for \( x_t \), we have \( \Sigma_{x|m'} = \Sigma_{x|m} \), and \( K_{x|y|m'} = K_{x|y|m} \). \( K' \) can then be written as
\[
K' = K_{x|y|m} (I - K_{x|y|m} C_{y|x}) \Sigma_{x|m} \Sigma_{m|m}' \Sigma_{m|m}' \Sigma_{m|m} \Sigma_{m|m},
\]
in the rightmost columns omitted due to its redundancy.

The controllers \( \pi \) and \( \pi' \) generate the same control \( u_t \), and therefore incur the same external cost. At the same time, since \( m_t' \) is a function of \( m_t \), by the data-processing inequality the information rate of \( \pi' \) is at most that of \( \pi \). Thus any controller can be converted into a MMSE controller without loss of performance, allowing us to consider the MMSE controller canonical. In particular, this proves again that \( d = n \) is always sufficient for representing the memory state.

At this point, we diverge from the solution given in Sec. II-B, which is not guaranteed to be a MMSE controller. Instead, we explicitly constrain the controller to be MMSE, which in return enables us to relax some of the conditions given in Sec. II-B, which are now not necessary (and indeed do not hold at the optimum), as discussed below.

Constraining the controller to be MMSE imposes the structure
\[
\Sigma_{\tilde{x}} = \begin{bmatrix} \Sigma_{x|m} + \Sigma_{m|m} \Sigma_{m|m} \Sigma_{m|m} \Sigma_{m} \end{bmatrix},
\]
parameterized by $\Sigma_{x|m}$ and $\Sigma_m$. The reduced number of independent parameters leaves $M$ overparameterized (see SM, Appx. VII), and we can choose, without loss of performance, the structure

$$M = \begin{bmatrix} M_{x|m} + M_m & -M_m \\ -M_m & M_m \end{bmatrix}$$

with

$$M_{x|m} = \beta^{-1} Z$$

$$M_m = \beta^{-1}(C_{y|x}^T X_{y|m} Z K_{y|x|m} C_{y|x} - Z),$$

where

$$Z = \Sigma_{\tilde{x}|\tilde{x}} - \Sigma_{\tilde{x}y}^T \Sigma_{\tilde{x}y}^{-1}$$

is the signal-to-noise-ratio (SNR) matrix for the channel $\tilde{x}_{y} \rightarrow \tilde{x}_{u}$. Due to the shrinkage effect of $K_{y|x|m} C_{y|x}$

$$M_m \preceq 0 \preceq M_{x|m} + M_m.$$

The Hessian of the cost-to-go now has the form

$$S = Q + A^T S A - M = \begin{bmatrix} Q_x + A_x^T S_x A_x & -M_{x|m} \\ -M_{m} & M_m \end{bmatrix},$$

and the second-order expansion of the cost-to-go, at the optimum, has the form

$$\tilde{x}_t^T S \tilde{x}_t = x_t^T (Q_x + A_x^T S_x A_x - \beta^{-1} Z) x_t - (M_t - x_t)^T M_m (M_t - x_t).$$

The first term measures the divergence of the state $x_t$ from 0, and the second the divergence of the controller’s estimator $\hat{m}_t$ from the true state $x_t$, which is the expected form for a MMSE controller. Both terms link the SNR matrix $Z$ to the cost reduction. In this form, $S$ is again positive semidefinite, while now $M$ is generally not.

Finally, when $\beta = \infty$, we can recover the classic LQG results. Similarly to Part I [1, Sec. IV-B], we can substitute $N_{x|m}$ for $\beta^{-1} Z$, to recover the algebraic Riccati equation

$$S_{x|m} = Q_x + A_x^T S_x A_x - N_{x|m}$$

$$= Q_x + A_x^T (S_{x|m} - S_{z|m} B_{z|m} S_{z|m} B_{x|m}^T) A_x.$$

IV. EXAMPLE

As a simple example, consider the double mass-spring-damper system in Fig. 5, adapted from [14]. The continuous-time dynamics of this system are given by

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{k_1+k_2}{m_1} & -\frac{c_1+c_2}{m_1} & \frac{k_2}{m_1} & \frac{c_2}{m_1} \\ 0 & 0 & 0 & 0 \\ \frac{k_3}{m_2} & \frac{c_3}{m_2} & -\frac{k_2+k_3}{m_2} & -\frac{c_1+c_2}{m_2} \end{bmatrix}$$

$$B = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{m_1} \\ 0 \\ 0 \\ \frac{1}{m_2} \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

with $m_1 = 5\text{ kg}$, $m_2 = \sqrt{15}\text{ kg}$, $k_1 = 1\text{ N/m}$, $k_2 = 0.5\text{ N/m}$, $c_1 = c_2 = 1\text{ N sec/m}$. We discretize the time using the Tustin transformation with sampling frequency $20\text{Hz}$, and consider the isotropic noises and cost rates

$$\Sigma_\xi = I \quad \Sigma_\epsilon = I \quad Q = I \quad R = I.$$

For the memoryless control problem, we initialize a solution with $\Sigma_x = S = 0$. For the retentive control problem, we apply the reduction in Lemma 2 to obtain a reduced plant, and then initialize a solution using the classic LQG controller, as described in Sec. III-C. To the initial solution, we apply the forward-backward iterations of Sec. II-B, with fixed $\beta$, until convergence. To improve running time, we employ a reverse-annealing scheme, decreasing $\beta$ gradually over its range, and using the fixed point for one value of $\beta$ to initialize the iterations for the next value of $\beta$.

Fig. 6 and Fig. 7 show, respectively, the resulting cost-log-beta and cost-information curves, demonstrating that even this simple example exhibits interesting phenomenology.

We see that both the memoryless (blue) and the retentive (green) controllers undergo phase transitions as $\beta$ increases. The system is controllable and observable, while $\text{rank}(B) = \text{rank}(C) = 2$, allowing the memoryless controller to undergo 2 phase transitions and reach order $d = 2$, while the retentive controller undergoes 4 phase transitions, until it fully remembers and controls all modes of the system.

In the first phase transition, the controllers begin controlling a single mode, in order to reduce the external cost, at the expense of communication resources. This is not depicted in the cost-information plot (Fig. 7), since below this critical point the information is 0, and the cost is fixed.

The second phase transition first involves memory, and only occurs in the retentive controller. Below this critical point, a hypothetical order-2 retentive controller is worse than the order-1 controller, in terms of the total external and internal cost-to-go $F$ it incurs. At the critical point, the order-2 controller overtakes the order-1 controller, already with a significantly reduced cost rate and a significant information rate (see red dots in Fig. 6 and Fig. 7). The critical point is where the ratio between these costs is $\beta^{-1}$ (see (12) in Part I [1, Sec. IV-B]).

The third phase transition is again common to the memoryless and the retentive controllers, although by now the retentive controller has committed to memory much valuable information, reducing the cost much beyond the capabilities of the memoryless controller.

The fourth phase transition is phenomenologically similar to the second one.
V. DISCUSSION

In this paper we introduced the problem of optimal LQG control with bounded channel capacity in both the memory and the sensory channels. We showed how to reduce this problem to that of bounded memoryless LQG control, and studied the structure of the resulting solution. We then illustrated the interesting phenomenology of the solution with a simple example.

An aspect of this phenomenology that merits further study, is the existence of suboptimal fixed points of the iterative algorithm (Sec. II-B). For example, around the second critical point in the double mass-spring-damper system (Sec. IV), both an order-1 controller and a retentive order-2 controller are fixed points. Before the phase transition, one of these solutions is stable, while the other is metastable and suboptimal, and at the phase transition they switch. This resembles well-studied phenomena in statistical physics, and is the subject of ongoing research.

LQG control with constraints on the sensory channel capacity has now been studied in the regime of unlimited memory [12], no memory (Part I of this work [1]), and in this paper, a shared channel capacity for sensing and memory. More generally, the memory and the sensory channels can be separate, with their relative costs ranging from 0 (no memory) to 1 (shared capacity) to $\infty$ (unlimited memory), and any intermediate value. This memory-sensory trade-off has been studied in the context of finite-state systems [10], and further insight can be gained from studying this more general problem in the LQG context.

REFERENCES

[1] R. Fox and N. Tishby, “Minimum-information LQG control — Part I: Memoryless controllers,” submitted for publication. [Online]. Available: https://arxiv.org/abs/1606.01946
[2] V. Cutsuridis, A. Hussain, and J. G. Taylor, Perception-Action Cycle: Models, Architectures, and Hardware, ser. Cognitive and Neural Systems. Springer, 2011.
[3] D. A. Patterson, “Latency lags bandwidth,” Communications of the ACM, vol. 47, no. 10, pp. 71–75, 2004.
[4] D. P. Bertsekas, Dynamic Programming and Optimal Control. Athena Scientific, 1995, vol. 1.
[5] S. K. Mitter, “Control with limited information,” European Journal of Control, vol. 7, no. 2, pp. 122–131, 2001.
[6] 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.
[7] J. 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.
[8] A. S. Matveev and A. V. Savkin, Estimation and Control over Communication Networks. Birkhäuser, 2009.
[9] W. Bialek, I. Nemenman, and N. Tishby, “Predictability, complexity, and learning,” Neural computation, vol. 13, pp. 2409–2463, 2001.
[10] R. Fox and N. Tishby, “Bounded planning in passive POMDPs,” in Proceedings of the 29th International Conference on Machine Learning, 2012, pp. 1775–1782.
[11] J. L. Massey, “Causality, feedback and directed information,” in Proceedings of the International Symposium on Information Theory and Its Applications, 1990, pp. 303–305.
[12] T. Tanaka, P. M. Esfahani, and S. K. Mitter, “LQG control with minimal information: Three-stage separation principle and SDP-based solution synthesis,” arXiv preprint arXiv:1510.04214, submitted for publication.
[13] T. M. Cover and J. A. Thomas, Elements of Information Theory. Wiley, 2012.
[14] F. Creutzig, A. Globerson, and N. Tishby, “Past-future information bottleneck in dynamical systems,” Physical Review E, vol. 79, no. 4, p. 041925, 2009.
Minimum-Information LQG Control
Supplementary Material
Roy Fox† and Naftali Tishby†

APPENDIX I
PERFECTLY MATCHED CHANNEL

In this appendix we prove the form of the perfectly matched channel described in Part I† [1, Sec. III-B]. We rely on the results of Theorem 1 there (Sec. IV-A).

The main results of [2], applied to our setting, can be summarized as follows. We wish to find a memoryless channel into which we can input an encoding \( w_t \) of \( \hat{x}_{yt} \), such that \( \hat{x}_{ut} \) can be decoded from the channel output \( \hat{w}_t \). Let \( f(\hat{w}_t|w_t) \) be the channel probability density function, and \( g \) and \( h \), respectively, the encoder and the decoder, so that \( w_t = g(\hat{x}_{yt}) \) and \( \hat{x}_{ut} = h(\hat{w}_t) \). Suppose that we are concerned with the power needed to transmit \( w_t \), and so the input cost is \( w_t^Tw_t \). Then the source \( y_t \) and the channel \( w_t \rightarrow \hat{w}_t \) are perfectly matched if

1) The Kullback-Leibler divergence \( \mathbb{D}[f(\hat{w}_t|w_t)||f(\hat{w}_t)] \) equals \( c_1 w_t^2 \hat{w}_t + c_2 \), for some constants \( c_1 \geq 0 \) and \( c_2 \), and
2) \( f(\hat{x}_{ut}|\hat{x}_{yt}) \) satisfies the conditions in Theorem 1.

To meet these conditions, we can take

\[
\begin{align*}
  w_t &= D^{1/2}V^T \Sigma_{x,y}^{1/2} \tilde{x}_{yt}, \\
  \hat{w}_t &= w_t + v_t; \quad v_t \sim \mathcal{N}(0, I - D) \\
  \hat{x}_{ut} &= \Sigma_{x,y}^{1/2} V D^{1/2} \hat{w}_t.
\end{align*}
\]

Then

\[
\begin{align*}
  \Sigma_w &= D \\
  \Sigma_{\hat{w}} &= I \\
  \Sigma_{\hat{x}} &= \Sigma_{x,y}^{1/2} V D V^T \Sigma_{x,y}^{1/2} = \Sigma_{\hat{x},\hat{x}},
\end{align*}
\]

and it can be verified that

\[
\mathbb{D}[f(\hat{w}_t|w_t)||f(\hat{w}_t)] = \frac{1}{2} w_t^T \Sigma_{w}^{-1} w_t + \text{const},
\]

as required.

The capacity of the additive Gaussian noise channel with noise covariance \( I - D \), under the appropriate power constraint, is indeed achieved by a Gaussian input with covariance \( D \), and is equal to the information rate in Theorem 1. As shown in [2], this means that constraining the power \( \Sigma_w \) is equivalent to constraining the information rate \( \mathbb{D}[\hat{x}_{yt}||\hat{x}_{ut}] \). Note, however, that the matched channel noise covariance depends on the constraint, through the solution in Theorem 1.

APPENDIX II
PROOF OF LEMMA 2 OF PART I

In this appendix we restate and prove Lemma 2 of Part I [1, Sec. IV-A].

Lemma 2: Let \( x \) and \( \hat{x} \) be jointly Gaussian random variables with mean 0. The following properties are equivalent:

1) There exists a random variable \( u \), jointly Gaussian with \( x \), such that \( \hat{x}(u) = \arg\min_x \mathbb{E}[||\hat{x} - x||^2|u] = \mathbb{E}[x|u] \).
2) \( \Sigma_{\hat{x}|x} = \Sigma_{\hat{x}} \).
3) \( \Sigma_{x|\hat{x}} = \Sigma_x - \Sigma_{\hat{x}} \), where \( \Sigma_{x|\hat{x}} \) is the conditional covariance matrix of \( x \) given \( \hat{x} \), implying \( \Sigma_{x} \geq \Sigma_{\hat{x}} \).
4) \( \hat{x} = \mathbb{E}[x|\hat{x}] \).

Such \( \hat{x} \) is called a minimum mean square error (MMSE) estimator (of \( u \)) for \( x \).

Proof: (1 \( \implies \) 2) Assume without loss of generality that \( u \) has mean 0. Then

\[
\hat{x} = \Sigma_{x|u} \Sigma_{u}^{-1} u,
\]

implying

\[
\Sigma_{\hat{x}|x} = \Sigma_{x|u} \Sigma_{u}^{-1} \Sigma_{u|x} = \Sigma_{\hat{x}}.
\]

(2 \( \implies \) 3)

\[
\Sigma_{x|\hat{x}} = \Sigma_x - \Sigma_{\hat{x}} \Sigma_{\hat{x}|x} \Sigma_{\hat{x}} = \Sigma_x - \Sigma_{\hat{x}}.
\]

(3 \( \implies \) 4) Since \( x \) and \( \hat{x} \) are jointly Gaussian with mean 0, we can write for some \( T \)

\[
x = T \hat{x} + \xi; \quad \xi \sim \mathcal{N}(0, \Sigma_{x|\hat{x}}),
\]

implying

\[
\Sigma_x = T \Sigma_{\hat{x}} T^T + \Sigma_x - \Sigma_{\hat{x}},
\]

so without loss of generality \( T = I \).

(4 \( \implies \) 1) Taking \( u = \hat{x} \), we have

\[
\arg\min_{\hat{x}} \mathbb{E}[||\hat{x}' - x||^2|u]
= \arg\min_{\hat{x}'} (\hat{x}'^T \hat{x}' - 2 \hat{x}'^T \mathbb{E}[x|u] + \mathbb{E}[x^T x|u]),
\]

which is optimized by \( \hat{x}' = \mathbb{E}[x|u] \).

APPENDIX III
PROOF OF LEMMA 2 OF PART I

In this appendix we restate and prove Lemma 2 of Part I [1, Sec. IV-A].

†School of Computer Science and Engineering, The Hebrew University, {royf,tishby}@cs.huji.ac.il
†Available at https://arxiv.org/abs/1606.01946
Lemma 2: The bounded memoryless LTI controller optimization problem (Problem 1) is solved by a control law of the form
\[
\begin{align*}
\dot{x}_{yt} &= K y_t \\
\dot{x}_{ut} &= W \hat{x}_{yt} + \omega_t; \quad \omega_t \sim \mathcal{N}(0, \Sigma_{\omega}) \\
u_t &= L \hat{x}_{ut},
\end{align*}
\] (5a)

where \( W \in \mathbb{R}^{n \times k}, \Sigma_{\omega} \in \mathbb{R}^{n \times n}, L \in \mathbb{R}^{f \times n} \), \( \omega_t \) is independent of \( y_t, \hat{x}_{yt} \), is a MMSE estimator for \( \hat{x}_{yt} \), and (6)

\[
\mathbb{I}[y_t; u_t] = \mathbb{I}[\hat{x}_{yt}; \hat{x}_{ut}].
\]

Proof: Consider a LTI controller \( \pi \) of the form
\[
u_t = H y_t + \eta_t; \quad \eta_t \sim \mathcal{N}(0, \Sigma_{\eta}),
\] (III.1)
satisfying the Markov network
\[
\begin{align*}
\begin{array}{c|c}
x_t & y_t & u_t \\
\hline
\hat{x}_{yt} & \hat{x}_{ut} & \end{array}
\end{align*}
\] (III.2)

By construction, \( u'_t \) has the same joint distribution with \( \hat{x}_{ut} \) as \( u_t \) does. Since \( \hat{x}_{ut} \) is a sufficient statistic of \( u_t \) for \( x_t \), this implies that \( u'_t \) also has the same joint distribution with \( x_t \) as \( u_t \) does. Thus \( \pi' \) induces the same stochastic process \( \{x_t, u'_t\} \), and the same external cost. Note that \( u'_t \) may not have the same joint distribution with \( y_t \) as \( u_t \) does, and due to the data-processing inequality [3]
\[
\mathbb{I}[y_t; u_t] \geq \mathbb{I}[y_t; u'_t].
\]

Therefore \( \pi' \) performs at least as well as \( \pi \), and equally well when \( \pi \) is optimal.

Now (6) follows from the data-processing inequality, since our solution satisfies both structures (III.2) and (III.3). \( \hat{x}_{ut} \) is a MMSE estimator for \( \hat{x}_{yt} \), since
\[
\mathbb{E}[\hat{x}_{yt}|\hat{x}_{ut}] = \mathbb{E}[\mathbb{E}[x_t|y_t]|\hat{x}_{ut}]
= \mathbb{E}[x_t|\hat{x}_{ut}] = \hat{x}_{ut},
\]
where the second equality follows from \( x_t - y_t = \hat{x}_{yt} \).

The Lagrangian of the optimization problem ((9) in Part I) in this parameterization would depend on \( \Sigma_{\pi} \) only through the terms
\[
\frac{1}{2}(\text{tr}(R_{\Sigma}) + \text{tr}(SB_{\Sigma} B^T)).
\]
Since \( R + B^T S B \succeq 0 \) is positive semidefinite, we can take \( \Sigma_{\pi} = 0 \) without loss of performance, recovering the structure (5). Intuitively, the argument is that any noise added to \( u'_t \), beyond \( \hat{x}_{ut} \), is not helpful in compressing \( x_t \), and can only increase the external cost, without saving any communication cost.

In the other direction, let \( u_t \) satisfy the form of Lemma 2. We can rewrite \( u_t \) in the form (III.1), with
\[
H = LWK, \quad \Sigma_{\pi} = L \Sigma_{\omega} L^T.
\]

Proof of Theorem 1 of Part I

In this appendix we restate and prove Theorem 1 of Part I [1, Sec. IV-A], which relies on the following Lagrangian developed there.
\[
\begin{align*}
\mathcal{F}_{\Sigma_{\pi}, \Sigma_{\pi}, L, S; \beta} &= \frac{1}{2}(\beta^{-1} \text{log} |\Sigma_{\pi}| - \text{log} |\Sigma_{\pi}|_0) + \text{tr}(Q \Sigma_{\pi}) + \text{tr}(R L \Sigma_{\pi} L^T) \\
&\quad + \text{tr}(S ((A + BL) \Sigma_{\pi} (A + BL)^T + A \Sigma_{y} A^T + \Sigma_{\xi} - \Sigma_{\pi})).
\end{align*}
\] (9)

Theorem 1: Given \( \beta \), the Lagrangian (9) is minimized by a controller satisfying
\[
\begin{align*}
\Sigma_{\pi} &= (A + BL) \Sigma_{y} (A + BL)^T + A \Sigma_{y} A^T + \Sigma_{\xi} \\
\Sigma_{\pi} &= \frac{1}{2} \text{tr}(B V D V^T) \\
L &= - (R + B^T S B)^T B^T S A \\
S &= Q + A^T S A - M,
\end{align*}
\] (10a)

with the eigenvalue decomposition (EVD)
\[
V A V^T = \Sigma_{y}^{1/2} \Sigma_{y}^{1/2}.
\] (10j)
such that \( V \) is orthogonal with \( n - \text{rank}(\Sigma_{\pi}) \) columns spanning the kernel of \( \Sigma_{\pi} \) and \( A = \text{diag}(\lambda_i) \), and with
\[
D = \text{diag} \left\{ \begin{array}{l}
1 - \beta^{-1} \lambda_i^{-1} \quad \lambda_i > \beta^{-1} \\
0 \quad \lambda_i \leq \beta^{-1}
\end{array} \right\}.
\] (10k)

Proof: The minimum of the Lagrangian (9) must satisfy the first-order optimality conditions, i.e. that the gradient with respect to each parameter is 0 at the optimum. We start by differentiating \( \mathcal{F} \) by the feedback gain \( L \)
\[
\partial_L \mathcal{F}_{\Sigma_{\pi}, \Sigma_{\pi}, L, S; \beta} = RL \Sigma_{\pi} + B^T S (A + BL) \Sigma_{\pi} = 0,
\]
which we rewrite as
\[
(R + B^T S B)L \Sigma_{\pi} = -B^T S A \Sigma_{\pi}.
\]
As this equation shows, \( L \) is underdetermined in the kernel of \( \Sigma_{\pi} \), since these modes are always 0 in \( \hat{x}_{ut} \), and have no effect on \( u_t \). \( L \) is also underdetermined in the kernel of \( R + B^T S B \), since these modes have no cost (immediate or future), and can be controlled in any way without affecting
the solution’s performance. Thus without loss of performance we can take
\[ L = -(R + B^T SB)^T B^T SA. \]
We substitute this solution back into the Lagrangian, to get
\[ F_{\Sigma_{x}, \Sigma_{\hat{x}_u}, \beta} = \frac{1}{2} \left( \beta^{-1} \left( \log |\Sigma_{\hat{x}_y}| - \log |\Sigma_{x_y}| + 1 \right) \right) \tag{IV.4} \]
\[ + \text{tr}(M \Sigma_x) - \text{tr}(N \Sigma_{\hat{x}_u}) + \text{tr}(S \Sigma_x), \]
with
\[ M = Q + A^T SA - S \]
\[ N = L^T (R + B^T SB) L \]
\[ = A^T SB (R + B^T SB)^T B^T SA. \]
The problem of optimizing over \( \Sigma_{x_u} \), given the other parameters, can now be written, up to constants, as the semidefinite program (SDP)
\[ \max_{\Sigma_{x_u}} \log |\Sigma_{x_y} - \Sigma_{\hat{x}_u}| + \beta \text{tr}(N \Sigma_{\hat{x}_u}) \]
\[ \text{s.t.} \quad 0 \preceq \Sigma_{x_u} \preceq \Sigma_{\hat{x}_y}. \]

By Lemma V.1 in Appx. V, the optimum is achieved when \( \Sigma_{x_u} \) satisfies (10b) and (10j)-(10k).

Finally, with \( P = \Sigma_{\hat{x}_y} \Sigma_{x_y}^T \), the projection onto the support of \( \hat{x}_y \), and since the range of \( \Sigma_{\hat{x}_u} \) is contained in that subspace, we have
\[ \partial_{(\Sigma_{x},)_{i,j}} \left( \log |\Sigma_{x_y} - \Sigma_{\hat{x}_u}| - \log |\Sigma_{x_y}| \right) \]
\[ = -\partial_{(\Sigma_{x},)_{i,j}} \log |P - \Sigma_{\hat{x}_u} \Sigma_{x_y}^T| \]
\[ = -\partial_{(\Sigma_{x},)_{i,j}} \log |I - \Sigma_{\hat{x}_u} (P \Sigma_{x_y} P)^T| \]
\[ = \text{tr}((I - \Sigma_{\hat{x}_u} \Sigma_{x_y}^T)^{-1} \Sigma_{\hat{x}_u} \partial_{(\Sigma_{x},)_{i,j}} (P \Sigma_{x_y} P)^T). \]
The purpose of introducing \( P \) is to notice that even if the range of \( \Sigma_{\hat{x}_y} \) is increased, this has no effect on the Lagrangian, because these modes are orthogonal to the range of \( \Sigma_{\hat{x}_u} \). This allows us to treat \( P \) as constant, so that the range of \( P \Sigma_{\hat{x}_y} \) is constant in a neighborhood of the solution, and the derivative of the pseudoinverse is simplified in this case to
\[ \partial_{(\Sigma_{x},)_{i,j}} (P \Sigma_{x_y} P)^T \]
\[ = -\Sigma_{x_y}^T \partial_{(\Sigma_{x},)_{i,j}} (\Sigma_{x_y}) \Sigma_{x_y}^T \]
\[ = -\Sigma_{x_y}^T K C J_{i,j} C^T K^T \Sigma_{x_y}^T, \]
with \( J_{i,j} \) the matrix with 1 in position \((i,j)\) and 0 elsewhere. This yields
\[ \partial_{\Sigma_x} F_{\Sigma_{x}, \Sigma_{\hat{x}_u}, \beta} \]
\[ = \frac{1}{2} \left( (M - \beta^{-1} C^T K^T \Sigma_{x_y} (I - \Sigma_{\hat{x}_u} \Sigma_{x_y}^T)^{-1} \Sigma_{\hat{x}_u} \Sigma_{x_y}^T K C) \right) \]
\[ = \frac{1}{2} \left( (M - \beta^{-1} C^T K^T \Sigma_{x_y} (I - \Sigma_{\hat{x}_u} \Sigma_{x_y}^T)^{-1} - I) K C) \right) \]
\[ = \frac{1}{2} \left( (M - \beta^{-1} C^T K^T (\Sigma_{\hat{x}_u} - \Sigma_{x_y}) K C) = 0, \right. \]
implying (10h).

**APPENDIX V**

**SEMIDEFINITE PROGRAM SOLUTION**

In this appendix we state and prove the following solution to our SDP problem.

**Lemma VI.1:** The semidefinite program
\[ \max_{X \in \mathbb{S}^+} \log |M_1 - X| + \text{tr}(M_2 X) \]
\[ \text{s.t.} \quad X \preceq M_1, \]
with \( M_1, M_2 \geq 0 \), is optimized by
\[ X = M_1^{1/2} V D V^T M_1^{1/2}, \]
with the eigenvalue decomposition (EVD)
\[ V \Lambda V^T = M_1^{1/2} M_2 M_1^{1/2}, \]
such that \( V \) is orthogonal with \( n - \text{rank}(M_1) \) columns spanning the kernel of \( M_1 \) and \( \Lambda = \text{diag}\{\lambda_i\} \), and with
\[ D = \text{diag}\left\{ \begin{array}{c} 1 - \lambda_i^{-1} \\ \lambda_i > 1 \\ \lambda_i \leq 1 \end{array} \right\}. \]

**Proof:** Let the EVD of \( M_1 \) be
\[ U \Psi \Psi^T = M_1, \]
with \( U \) orthogonal and \( \Psi \) diagonal, having
\[ \Psi = \left[ \begin{array}{c} \Psi_+ \\ 0 \\ 0_{(n-m) \times (n-m)} \end{array} \right], \]
with \( m = \text{rank}(M_1) \). Let
\[ \Psi^T = \Psi^T - I - \Psi^T \Psi = \left[ \begin{array}{c} \Psi_+^{-1} \\ 0 \\ I \end{array} \right]. \]
By changing the variable to
\[ Y = \Psi^{1/2} U^T X U \Psi^{1/2}, \]
the constraint of the SDP becomes
\[ Y \preceq I_{m,n} = \left[ \begin{array}{c} I_{m \times m} \\ 0 \\ 0_{(n-m) \times (n-m)} \end{array} \right]. \]
\( Y \) must therefore be 0 outside the upper-left \( m \times m \) block, and the SDP is equivalent, up to constants, to
\[ \max_{Y \in \mathbb{S}^+} \log |M_{1,n} - Y| + \text{tr}(\Psi^{1/2} U^T M_2 U \Psi^{1/2} Y) \]
\[ \text{s.t.} \quad Y \preceq I_{m,n}. \]
Let the EVD of the linear coefficient be
\[ \tilde{V} \Lambda \tilde{V}^T = \Psi^{1/2} U^T M_2 U \Psi^{1/2}, \]
with
\[ \tilde{V} = \left[ \begin{array}{c} \tilde{V}_+ \\ 0 \\ I_{(n-m) \times (n-m)} \end{array} \right] \]
orthogonal and preserving the kernel of \( \Psi \), and \( \Lambda = \text{diag}\{\lambda_i\} \). We can again change the kernel of \( \Psi \), and
\[ D = \tilde{V}^T Y \tilde{V}, \]
to get

\[
\max_{D \in \mathbb{S}^+_{n}} \log |I_{m,n} - D| + \text{tr}(\Lambda D)
\]

s.t. \( D \preceq I_{m,n} \),

which can easily be solved using Hadamard’s inequality [3], to find

\[
D = \text{diag} \left\{ \frac{1 - \lambda_i^{-1}}{0, \lambda_i > 1} \right\}.
\]

Finally, the lemma follows by unmaking the variable changes and taking

\[
V = UV.
\]

**APPENDIX VI**

**PROPERTIES OF THE RETENTIVE DIRECTED INFORMATION**

In this appendix we show how the retentive directed information (Definition 6 of Part II \(^2\) [4, Sec. III-A]) relates to the multi-information of Bayesian networks [5].

Consider the Bayesian network in Fig. VI.1, which describes the process of online inference from a sequence of independent observations. The multi-information of this network, for horizon \( T \), is equal to the retentive directed information

\[
\mathbb{I}[y^T, z^T] = \mathbb{E} \left[ \log \frac{f(y^T, z^T)}{\prod_{t=1}^T f(y_t) f(z_t)} \right]
\]

\[
= \sum_{t=1}^T \mathbb{E} \left[ \log \frac{f(z_t|z_t^{t-1}, y^t)}{f(z_t)} \right] = \mathbb{I}[y^T \rightarrow z^T].
\]

An important property of the directed information is that the mutual information between two sequences can be decomposed into the sum of directed information in both directions [6]

\[
\mathbb{I}[x^T, z^T] = \mathbb{I}[x^T \rightarrow z^T] + \mathbb{I}[z^T \rightarrow x^T].
\]

Interestingly, retentive directed information extends this property to the retentive control process (Fig. 1 in Part II). This process can be thought of as consisting of four phases: observation, inference, control and state transition. Its multi-information can accordingly be decomposed [7] into the sum

\[
\mathbb{I}[x^T, y^T, z^T, u^T] = \mathbb{I}[x^T \rightarrow y^T] + \mathbb{I}[y^T \rightarrow z^T] + \mathbb{I}[z^T \rightarrow u^T] + \mathbb{I}[u^T \rightarrow x^T].
\]

---

\(^2\)Available at https://arxiv.org/abs/1606.01947

**APPENDIX VII**

**STRUCTURE OF THE OPTIMAL RETENTIVE CONTROLLER**

In this appendix we derive the structure of the optimal retentive controller summarized in Part II [4, Sec. III-C].

For the structured feedback gain \( L \) we find using the Schur complement that

\[
(R + B^T S B)^\dagger = \begin{bmatrix} R_u + B_{x;u}^T S_{x;u} B_{x;u} & B_{x;u}^T S_{x;m}^\dagger \\ S_{m;x} B_{x;u} & S_m \end{bmatrix}
\]

\[
= \begin{bmatrix} S_u^\dagger & -S_u^\dagger | B_{x;u}^T S_{x;m}^\dagger | S_{m}^\dagger \\ -S_{m;x}^\dagger S_{m;x} B_{x;u} S_u^\dagger | S_{m}^\dagger \end{bmatrix},
\]

with

\[
S_u^\dagger = S_u^\dagger + S_m^\dagger S_{m;x} B_{x;u} S_u^\dagger | B_{x;u}^T S_{x;m}^\dagger | S_{m}^\dagger,
\]

and so

\[
L = -(R + B^T S B)^\dagger B^T S A
\]

\[
= -(R + B^T S B)^\dagger B_{x;u}^T S_{x;u} A_x 0 \\
\]

\[
\begin{bmatrix} S_u^\dagger & 0 \\
S_{m;x}^\dagger (I - B_{x;u} S_u^\dagger B_{x;u}^T S_{x;m} A_x) 0 
\end{bmatrix},
\]

with

\[
L_{u;x|m} = -S_u^\dagger | B_{x;u}^T S_{x;m} A_x.
\]

We also have

\[
N = L^T (R + B^T S B) L
\]

\[
= A^T S B (R + B^T S B)^\dagger B^T S A
\]

\[
= \begin{bmatrix} S_u^\dagger & 0 \\
S_{m;x}^\dagger (I - B_{x;u} S_u^\dagger B_{x;u}^T S_{x;m} A_x) 0 
\end{bmatrix},
\]

Dually, for the structured Kalman gain \( K \) we find that

\[
\Sigma_y^\dagger = \begin{bmatrix} \Sigma_y & \Sigma_{y;m}^\dagger \\
\Sigma_{m;y} & \Sigma_m \end{bmatrix}^\dagger
\]

\[
= \begin{bmatrix} \Sigma_y^\dagger | \Sigma_{y;m}^\dagger \\
g_{m} & g_m \end{bmatrix}
\]

\[
= \begin{bmatrix} \Sigma_y^\dagger | \Sigma_{y;m}^\dagger \\
\Sigma_{m;y}^\dagger | \Sigma_{m;y}^\dagger \\
\Sigma_m + \Sigma_m^\dagger | \Sigma_{m;y}^\dagger | \Sigma_{m;y}^\dagger \\
\Sigma_m + \Sigma_m^\dagger \end{bmatrix}^\dagger,
\]

and so

\[
K = \Sigma_x C^T \Sigma_y^\dagger
\]

\[
= \begin{bmatrix} \Sigma_x C_{y;x} & \Sigma_x C_{y;x} \end{bmatrix} \begin{bmatrix} \Sigma_y & \Sigma_{y;m} \end{bmatrix}^\dagger
\]

\[
= [K_{x;y|m} (I - K_{x;y|m} C_{y;x}) \Sigma_{x;m}^\dagger],
\]

with

\[
K_{x;y|m} = \Sigma_x C_{y;x}^T \Sigma_{y|m}^\dagger.
\]
Now constraining the controller to be MMSE, we have the structure

\[
\Sigma_{\tilde{x}} = \begin{bmatrix}
\Sigma_{x|m} + \Sigma_m \\
\Sigma_m
\end{bmatrix}
\]

\[
K = \begin{bmatrix}
[K_{x:y|m} & I - K_{x:y|m}C_{y:x}]
\end{bmatrix},
\]

which we employ in differentiating \( F \) (IV.4), to get

\[
\partial_{\Sigma_{x|m}} F_{\Sigma_{x|m}, \Sigma_m, \Sigma_{\hat{x}_a}, S; \beta} = \begin{bmatrix}
I \\
0
\end{bmatrix}^T \partial_{\Sigma_{\tilde{x}}} F_{\Sigma_{\tilde{x}}, \Sigma_{\hat{x}_a}, S; \beta} \begin{bmatrix}
I \\
0
\end{bmatrix}
\]

\[
= \frac{1}{2} \left( \begin{bmatrix}
I \\
0
\end{bmatrix}^T (M - \beta^{-1}C^T K^T Z K C) \begin{bmatrix}
I \\
0
\end{bmatrix} \right) = 0
\]

\[
= \frac{1}{2} \left( \begin{bmatrix}
I \\
I
\end{bmatrix}^T (M - \beta^{-1}C^T K^T Z K C) \begin{bmatrix}
I \\
I
\end{bmatrix} \right) = 0,
\]

with

\[
Z = \Sigma_{\hat{x}_a}^{\dagger} - \Sigma_{\tilde{x}}^{\dagger}.
\]

This leaves \( M \) overparameterized, and we can choose to give it the structure

\[
M = \begin{bmatrix}
M_{x|m} + M_m & -M_m \\
-M_m & M_m
\end{bmatrix}
\]

with

\[
M_{x|m} = \beta^{-1} Z
\]

\[
M_m = \beta^{-1}(C^T_{x:y|m} Z K_{x:y|m} C_{y:x} - Z).
\]

REFERENCES

[1] R. Fox and N. Tishby, “Minimum-information LQG control — Part I: Memoryless controllers,” submitted for publication. [Online]. Available: https://arxiv.org/abs/1606.01946

[2] M. Gastpar, B. Rimoldi, and M. Vetterli, “To code, or not to code: Lossy source-channel communication revisited,” IEEE Transactions on Information Theory, vol. 49, no. 5, pp. 1147–1158, 2003.

[3] T. M. Cover and J. A. Thomas, Elements of Information Theory. Wiley, 2012.

[4] R. Fox and N. Tishby, “Minimum-information LQG control — Part II: Retentive controllers,” submitted for publication. [Online]. Available: https://arxiv.org/abs/1606.01947

[5] N. Friedman, O. Mosenzon, N. Slonim, and N. Tishby, “Multivariate information bottleneck,” in Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence, 2001, pp. 152–161.

[6] J. L. Massey, “Causality, feedback and directed information,” in Proceedings of the International Symposium on Information Theory and Its Applications, 1990, pp. 303–305.

[7] N. Tishby and D. Polani, “Information theory of decisions and actions,” in Perception-Action Cycle, ser. Cognitive and Neural Systems. Springer, 2011, pp. 601–636.