Continuous-Time Channel Gain Control for Minimum-Information Kalman-Bucy Filtering

Takashi Tanaka, Senior Member, IEEE, Vrushabh Zinage, Valery Ugrinovskii, Senior Member, IEEE, and Mikael Skoglund

Abstract—We consider the problem of estimating a continuous-time Gauss–Markov source process observed through a vector Gaussian channel with an adjustable channel gain matrix. Specifically, for various (generally time-varying) choices channel gain matrices, we study the tradeoff relationship between 1) the mean-square estimation error attainable by the classical Kalman–Bucy filter, and 2) the mutual information between the source process and its Kalman–Bucy estimate. We then formulate a novel “optimal channel gain control problem” where the objective is to control the channel gain matrix strategically to minimize the weighted sum of these two performance metrics. To develop insights into the optimal solution, we first consider the problem of controlling a time-varying channel gain over a finite time interval. A necessary optimality condition is derived based on Pontryagin’s minimum principle. For a scalar system, we show that the optimal channel gain is a piecewise constant signal with at most two switches. We also consider the problem of designing the optimal time-invariant gain to minimize the average cost over an infinite time horizon. A novel semidefinite programming heuristic is proposed and the exactness of the solution is discussed.

Index Terms—Continuous time systems, information theory, Kalman filters, networked control systems, optimal control.

I. INTRODUCTION

In this article, we consider the problem of estimating a continuous-time Gauss–Markov process from a noisy observation through a vector Gaussian channel with an adjustable channel gain matrix. For a fixed channel gain, the optimal causal estimate in the minimum mean-square-error (mse) sense is readily computable by the celebrated Kalman–Bucy filter. However, our focus in this article is on a generalized problem in which the channel gain matrix must also be designed strategically. Specifically, our objective is to design a (generally time-varying) channel gain matrix to minimize the weighted sum of 1) the mse attainable by the resulting Kalman–Bucy filter, and 2) the mutual information between the source process and its Kalman–Bucy estimate. Since choosing a “larger” channel gain results in a smaller mse and a larger mutual information, these two performance criteria are in a tradeoff relationship in general. Therefore, controlling the channel gain to attain the sweet spot is a nontrivial problem.

Related work: The problem we study is motivated by the causal source coding scenario in which an observed time series must be encoded, compressed, and transmitted over a digital communication medium to a remote decoder who tries to reproduce the original signal without delay [1], [2], [3], [4], [5], [6], [7]. Minimum-information Kalman filtering problems similar to ours appeared in [8], [9], [10], and [11] where the authors used mutual information between the source and the reproduced process to estimate the minimum bit-rate required for remote estimation. For stationary Gaussian sources, Derpich and Østergaard, [9] showed that the minimum bit-rate required to reproduce the signal within a given mse distortion criterion is lower bounded by the aforementioned mutual information and that the tightness of this lower bound is given by a constant known as the space-filling gap. Subsequent works [11] and [12] developed algorithms to compute the minimum mutual information required to estimate discrete-time Gauss–Markov processes with a desired mse level. The work of Tanaka et al. [11] proved that the optimal test channel (a stochastic kernel on the reproduced signal given the source signal that minimizes mutual information subject to a given mse distortion constraint) can always be realized by a memoryless Gaussian channel with an appropriately chosen channel gain followed by a Kalman filter. This result implies that, for discrete-time Gauss–Markov processes, computing the optimal channel gain that minimizes the mutual information under a given mse distortion constraint reveals a fundamental limitation of real-time data compression for remote estimation. Although the existing works on causal rate-distortion theory are largely limited to discrete-time settings (exceptions include [7]), this article generalizes the results to continuous-time settings.

Contribution of this article: Contributions of this article are summarized as follows.
1) We derive a new mutual information formula (Theorem 1) that generalizes both the main result of Duncan [13] and our earlier result [14, Th. 1].
2) We show the existence of a measurable solution to the optimal channel gain control problem over a finite time interval and derive a necessary optimality condition. These results are based on novel applications of Filippov’s theorem [15] and Pontryagin’s minimum principle.
3) We prove that the optimal channel gain control over a finite time interval for a scalar source process is piecewise constant with at most two discontinuities.
4) We also consider the problem of finding the optimal time-invariant channel gain for minimum-information Kalman–Bucy filtering. We propose a semidefinite programming (SDP) relaxation to compute an optimal solution candidate. We present the result of extensive numerical experiments suggesting that the relaxation is...
Fig. 1. System architecture and performance criteria.

in fact exact, although we are currently not aware of a theoretical guarantee on the exactness.

Preliminary versions of the results appeared in the authors’ conference publications [14] and [16]. The items 2) and 3) accommodate a channel gain constraint of the form $C_t; C_t \leq \gamma I$. This generalizes a special case with $\gamma = 1$ considered in [16] and avoids an oversimplification of the results. The item 4) did not appear in any of prior publications.

**Notation:** We will use notation $x_t; t \in [0, t_1)$ to denote a continuous-time signal. Bold symbols like $x$ will be used to denote random variables. We assume all the random variables considered in this article are defined on the probability space $(\Omega, F, P)$. The probability distribution $p_x$ of an $(X, A)$-valued random variable $x$ is defined by $p_x(A) = P[\omega \in \Omega : x(\omega) \in A] \forall A \in A$. If $x_1$ and $x_2$ are both $(X, A)$-valued random variables, the relative entropy from $x_2$ to $x_1$ is defined by $D(p_{x_2} \| p_{x_1}) = \int \log p_{x_1}/p_{x_2} \, dp_{x_2}$ provided that the Radon–Nikodym derivative $dp_{x_1}/dp_{x_2}$ exists, and $+\infty$ otherwise. 

The mutual information between two random variables $x$ and $y$ is defined by $I(x; y) = D(p_{xy} \| p_x p_y)$ where $p_{xy}$ and $p_x$ and $p_y$ denote the joint and product distributions, respectively. Random variables $x$, $y$, and $z$ form a Markov chain (denoted by $x \rightarrow y \rightarrow z$) if $z$ is conditionally independent of $x$ given $y$, i.e., $p_{z|xy} = p_{z|y}$.

## II. PROBLEM FORMULATION

### A. System Description

Let $(\Omega, F, P)$ be a complete probability space and suppose $F_t \subset F$ form a nondecreasing family of $\sigma$-algebras. Suppose $(w_t, F_t)$ and $(v_t, F_t)$ are $\mathbb{R}^n$-valued, mutually independent standard Wiener processes with respect to $P$. Define the random process to be estimated as an $n$-dimensional Itô process $dx_t = Ax_t \, dt + Bd w_t$, for all $t \in [0, t_1]$ with $x_{t_0} \sim N(0, X_0)$, where $X_0 \geq 0$ is a given covariance matrix and $[t_0, t_1]$ is a time interval. We assume $A$ is Hurwitz. Let $C_t : [t_0, t_1] \rightarrow \mathbb{R}^{n \times n}$ be a measurable function representing the time-variation of the channel gain setting. $x_t \triangleq C_t x_t$, the channel output is an $n$-dimensional signal as follows:

$$dy_t = z_t \, dt + dv_t, \quad t \in [t_0, t_1] \tag{1}$$

with $y_{t_0} = 0$. Based on the channel output $y_t$, the causal mmse estimate $\hat{x}_t \triangleq \mathbb{E}(x_t|F^t)$ is computed, where $F^t \subset F$ denotes the $\sigma$-algebra generated by $y_{[0, t]}$. The causal mmse estimate can be computed by the Kalman–Bucy filter as follows:

$$d\hat{x}_t = A\hat{x}_t \, dt + X_tC_t^T (y_t - C_t \hat{x}_t), \quad t \in [t_0, t_1] \tag{2}$$

with $\hat{x}_{t_0} = 0$. Here, $X_t := \mathbb{E}||x_t - \hat{x}_t||^2$, which is obtained as the unique solution to the matrix Riccati differential equation

$$\dot{X}_t = AX_t + X_tA^T - X_tC_t^T C_tX_t + BB^T, \quad t \in [t_0, t_1] \tag{3}$$

with the initial condition $X_{t_0} = X_0 \geq 0$. The system architecture considered in this article is shown in Fig. 1.

### B. Performance Criteria

In this article, we consider the problem of optimally controlling the time-varying channel gain $C_t$. The optimality is characterized in terms of the mse and the mutual information, as shown in Fig. 1.

1. **Mean-Square Error (MSE):** The first criterion is the MMSE achieved by the Kalman–Bucy filter.

$$\int_{t_0}^{t_1} \mathbb{E}||x_t - \hat{x}_t||^2 \, dt = \int_{t_0}^{t_1} \text{Tr}(X_t) \, dt. \tag{4}$$

2. **Mutual Information:** The second performance criterion is the mutual information $I(x_{t_0, t_1}; z_{t_0, t_1})$. The next theorem extends the result of Duncan [13] and provides a key formula to compute $I(x_{t_0, t_1}; z_{t_0, t_1})$ explicitly.

**Theorem 1:** Let the random processes $x_{t_0, t_1}$ and $z_{t_0, t_1}$ be defined as above. Then

$$I(x_{t_0, t_1}; z_{t_0, t_1}) = \frac{1}{2} \int_{t_0}^{t_1} \mathbb{E}||C_t (x_t - \hat{x}_t)||^2 \, dt. \tag{5}$$

Using the solution $X_t$ to the matrix Riccati differential (3) and (5) can also be written as $1/2 \int_{t_0}^{t_1} \text{Tr}(C_t X_tC_t) \, dt$.

**Proof:** The following identity is shown in [13]:

$$I(y_{t_0, t_1}; x_{t_0, t_1}) = \frac{1}{2} \int_{t_0}^{t_1} \mathbb{E}||C_t (x_t - \hat{x}_t)||^2 \, dt. \tag{6}$$

Due to the property of the Kalman–Bucy filter, $\hat{x}_{t_0, t_1}$ is a sufficient statistic of $y_{t_0, t_1}$ for $x_{t_0, t_1}$. Thus,

$$I(x_{t_0, t_1}; y_{t_0, t_1}) = I(x_{t_0, t_1}; \hat{x}_{t_0, t_1}). \tag{7}$$

To prove (5), it is left to show as follows:

$$I(x_{t_0, t_1}; z_{t_0, t_1}) = I(y_{t_0, t_1}; z_{t_0, t_1}). \tag{8}$$

This fact follows from the dependency structure of the processes $x_{t_0, t_1}$ and $z_{t_0, t_1}$. Indeed, under the law $P$, the conditional independence $p_{x_{t_0, t_1}|z_{t_0, t_1}}(x_{t_0, t_1}|z_{t_0, t_1}) = p_{x_{t_0, t_1}}(x_{t_0, t_1}|z_{t_0, t_1})$ holds because $x_{t_0, t_1}$ is a function of $x_{t_0, t_1}$. Therefore, $y_{t_0, t_1} \rightarrow x_{t_0, t_1} \rightarrow z_{t_0, t_1}$ and the data-processing inequality (see [17, Th. 2.8.1]) gives the following:

$$I(y_{t_0, t_1}; x_{t_0, t_1}) \geq I(y_{t_0, t_1}; z_{t_0, t_1}). \tag{9}$$

Furthermore, since $z_{t_0, t_1}$ is independent of $v_{t_0, t_1}$ under $P$, (I) implies the conditional independence $p_{y_{t_0, t_1}|z_{t_0, t_1}}(x_{t_0, t_1}|z_{t_0, t_1}) = p_{y_{t_0, t_1}|z_{t_0, t_1}}(x_{t_0, t_1}|z_{t_0, t_1})$. Therefore, $x_{t_0, t_1} \rightarrow z_{t_0, t_1} \rightarrow y_{t_0, t_1}$ and consequently the data-processing inequality gives the following:

$$I(x_{t_0, t_1}; z_{t_0, t_1}) \geq I(x_{t_0, t_1}; y_{t_0, t_1}). \tag{10}$$

Now since $x_{t_0, t_1} \rightarrow z_{t_0, t_1} \rightarrow y_{t_0, t_1}$ implies $y_{t_0, t_1} \rightarrow z_{t_0, t_1} \rightarrow x_{t_0, t_1}$ [18], we have the following:

$$I(y_{t_0, t_1}; z_{t_0, t_1}) \geq I(y_{t_0, t_1}; x_{t_0, t_1}). \tag{11}$$

Therefore, from (9) and (11), we can conclude that equality in (8) holds. Consequently, (5) holds. The last statement follows from the fact that $X_t = \mathbb{E}||x_t - \hat{x}_t||^2$.

Theorem 1 is more general than our previous result [14, Th. 1], since [14, Th. 1] was only applicable to time-invariant channel gains. The proof has also been simplified significantly by an application of the data-processing inequality. While Theorem 1 is related to the general result known as the I-mmsse relationship (e.g., [19] and references therein), the result of the form (5) has not appeared in the prior work to the best of the authors’ knowledge.

### C. Problem Setup

In this article, we are interested in the optimal choice of a measurable function $C_t : [t_0, t_1] \rightarrow \mathbb{R}^{n \times n}$ that minimizes the weighted sum of the
mse (4) and the mutual information (5). Notice that, in general, it is not possible to minimize these two quantities simultaneously. To see this, suppose we choose \( C_t = k C \forall t \) where \( k \geq 0 \) is a scalar and \((A, C)\) is an observable pair. As \( k \to +\infty \), the mse tends to zero, whereas the mutual information tends to \( +\infty \). Introducing a tradeoff parameter \( \alpha > 0 \), the main problem we study in this article is formulated as follows:

\[
\min_{c_t} \int_{t_0}^{t_1} \mathbb{E} \| x_t - \hat{x}_t \|^2 dt + 2 \alpha I \left( x_{[t_0,t_1]}, \hat{x}_{[t_0,t_1]} \right) \tag{12a}
\]

s.t. \( C_t^T C_t \leq \gamma I \quad \forall t \in [t_0, t_1]. \tag{12b} \)

The constraint (12b) imposes an upper bound on the allowable channel gain. Using (4) and (5), the problem (12) can be written more explicitly as follows:

\[
\min_{c_t} \int_{t_0}^{t_1} \text{Tr}(X_t) dt + \alpha \int_{t_0}^{t_1} \text{Tr} \left( C_t X_t C_t^T \right) dt \tag{13a}
\]

s.t. \( \dot{X}_t = AX_t + X_t A^T - X_t C_t^T C_t X_t + BB^T \tag{13b} \)

\( X_{t_0} = X_0, \quad C_t^T C_t \leq \gamma I \quad \forall t \in [t_0, t_1]. \tag{13c} \)

Introducing \( U_t \triangleq C_t^T C_t \geq 0 \), this can be written as an optimal control problem with state \( X_t \) and control input \( U_t \) as follows:

\[
\min_{u_t} \int_{t_0}^{t_1} \text{Tr}(X_t + \alpha U_t X_t) dt \tag{14a}
\]

s.t. \( \dot{X}_t = AX_t + X_t A^T - X_t U_t X_t + BB^T \tag{14b} \)

\( X_{t_0} = X_0, \quad U_t \geq 0, \quad U_t \leq \gamma I \quad \forall t \in [t_0, t_1]. \tag{14c} \)

The minimization is over the space of measurable functions \( U_t : [t_0, t_1] \to S^n_+ (\{ M \in \mathbb{R}^{n \times n} : M \succeq 0 \}). \)

Remark 1: The equivalence between (13) and (14) implies that optimal solutions to the main problem (12), if they exist, are not unique. Namely, if \( U_t^* \) is an optimal solution to (13), then both \( C_t^* \) and \( \hat{C}_t^* \) are optimal solutions to (13) if \( U_t^* = C_t^* C_t = \hat{C}_t^* \).

Remark 2: For simplicity, we assume that \( C_t \) is a square matrix throughout this article. A generalization to the case with \( C_t \in \mathbb{R}^{m \times n} \) where \( m \geq n \) is straightforward. However, a technical difficulty arises when \( m < n \). To see this, notice that \( C_t \in \mathbb{R}^{m \times n} \) implies \( \text{rank}(C_t^* C_t) \leq m \). Therefore, an additional nonconvex constraint \( \text{rank}(U_t) \leq m \forall t \in [t_0, t_1] \) must be included in (14) to maintain the equivalence between (13) and (14). This type of difficulty has been observed in the literature in the context of sensor design [20].

In this article, we also interested in the optimal time-invariant channel gain \( C_t = C \in \mathbb{R}^{n \times n} \) that minimizes the average cost over a long time horizon as follows:

\[
\min_{C \in \mathbb{R}^{n \times n}} \lim_{t_2 \to +\infty} \frac{1}{t_2 - t_0} \left\{ \int_{t_0}^{t_1} \mathbb{E} \| x_t - \hat{x}_t \|^2 dt + 2 \alpha I \left( x_{[t_0,t_1]}, \hat{x}_{[t_0,t_1]} \right) \right\} \tag{15a}
\]

s.t. \( X_{t_0} = X_0, \quad C^T C \leq \gamma I. \tag{15b} \)

This can be written as an equivalent optimization problem as follows:

\[
\min_{C \in \mathbb{R}^{n \times n}} \lim_{t_2 \to +\infty} \frac{1}{t_2 - t_0} \int_{t_0}^{t_1} \text{Tr}(X_t + \alpha C X_t C^T) dt \tag{16a}
\]

s.t. \( \dot{X}_t = AX_t + X_t A^T - X_t C^T C X_t + BB^T \tag{16b} \)

\( X_{t_0} = X_0, \quad C^T C \leq \gamma I. \tag{16c} \)

III. Optimality Condition

In this section, we briefly revisit some general results in optimal control theory to discuss the existence of an optimal control for (14) and to derive a necessary optimality condition. Consider the following Lagrange-type optimal control problem with a fixed end time and a free end point:

\[
\min_{u_t} \int_{t_0}^{t_1} L(x_t, u_t) dt \tag{17a}
\]

s.t. \( \dot{x}_t = f(x_t, u_t), \quad x_t \in \mathbb{R}^n, \quad u_t \in U \tag{17b} \)

\( x_{t_0} = x_0. \tag{17c} \)

Suppose that an admissible control input is a measurable function \( u_t : [t_0, t_1] \to U \) where \( U \subseteq \mathbb{R}^n \) is a compact set. We assume that \( f(x_t, u_t), \partial f(x_t, u_t) \partial u_t, L(x_t, u_t), \) and \( \partial L(x_t, u_t) / \partial x_t \) are continuous on \( \mathbb{R}^n \times U \times [t_0, t_1] \). Notice that problem (14) is in the form (17) if \( U = \{(u_0, v) : U \subseteq \mathbb{R}^n, u \leq \gamma I \} \).

A. Existence of Optimal Control

We first invoke the following useful result by Filippov [21]:

Lemma 1: (Filippov’s theorem): [15] Given a control system (17b) with \( u_t \in U \), assume that its solutions exist on a time interval \([t_0, t_1]\) for all controls and that for every \( x_t \) the set \( \{ f(x_t, u_t) : u_t \in U \} \) is compact and convex. Then, the reachable set \( R_t(x_0) \) is compact for each \( t \in [t_0, t_1] \). Theorem 1 is applicable to guarantee the existence of an optimal control for (14). Specifically, one can convert the original Lagrange-type problem (14) into an equivalent Mayer-type problem by introducing an auxiliary state \( x^{m(t)}_t \) satisfying \( x^{m(t)}_t \succeq 0 \) and \( x^{m(t)}_t = \text{Tr}(X_t + \alpha C X_t C^T) \). In the Mayer form, the original problem of minimizing (14a) becomes the problem of minimizing \( x^{m(t)}_t \) over the reachable set at \( t = t_1 \). Since the premises of Theorem 1 are satisfied by the obtained Mayer-type problem, we can conclude that the reachable set at \( t = t_1 \) is compact. Therefore, Weierstrass’ extreme value theorem guarantees the existence of an optimal solution.

B. Pontryagin Minimum Principle

We next invoke a version of Pontryagin’s minimum principle for the fixed-endtime free-endpoint optimal control problem (17).

Lemma 2 ([22, Th. 5.10]): Suppose there exists an optimal solution to (17). Let \( u^*_t : [t_0, t_1] \to U \) be an optimal control input and \( x^*_t : [t_0, t_1] \to \mathbb{R}^n \) be the corresponding state trajectory. Then, there exists a function \( p^*_t : [t_0, t_1] \to \mathbb{R}^n \) such that the following conditions hold for the Hamiltonian \( H \) defined as \( H(x_t, p_t, u_t) = L(x_t, u_t) + p_t f(x_t, u_t) \):

1) \( \dot{x}_t^* = \frac{\partial H(x^*_t, p^*_t, u^*_t)}{\partial p_t}, \quad \dot{p}_t^* = -\frac{\partial H(x^*_t, p^*_t, u^*_t)}{\partial x_t} \)

with boundary conditions \( x_{t_0} = x_0 \) and \( p_{t_1} = 0 \).

2) \( \min_{u_t \in U} H(x^*_t, p^*_t, u_t) = H(x^*_t, p^*_t, u^*_t) \) for all \( t \in [t_0, t_1] \).

For our problem (14), the Hamiltonian is defined as follows:

\[
H(X_t, P_t, U_t) = \text{Tr}(X_t + \alpha U_t X_t) + \text{Tr}(P_t (AX_t + X_t A^T - X_t U_t X_t + BB^T) + \text{Tr}(X_t) + \text{Tr}(\alpha X_t - X_t P_t U_t)).
\]

Thus, the necessary optimality condition provided by Theorem 2 is given by the following canonical equations:

\[
\dot{X}_t = AX_t + X_t A^T - X_t U_t X_t + BB^T \tag{18a}
\]
in Regions 1, 2, and 3 (see Fig. 3) as $f_1$, $f_2$, and $f_3$, respectively. From (20) and (21), we have a hybrid system as follows:

\begin{align}
\text{(Region 1) } f_1 : & \begin{cases} 
\dot{x}_t = 2ax_t + 1 \\
\dot{p}_t = -2ap_t - 1
\end{cases} \tag{22a} \\
\text{(Region 2) } f_2 : & \begin{cases} 
\dot{x}_t = 2ax_t - x_t^2 u_t + 1 \\
\dot{p}_t = 2x_t p_t u_t - 2ap_t - 1 - \alpha u_t
\end{cases} \tag{22b} \\
\text{(Region 3) } f_3 : & \begin{cases} 
\dot{x}_t = 2ax_t - \gamma x_t^2 + 1 \\
\dot{p}_t = 2\gamma x_t p_t - 2ap_t - 1 - \alpha \gamma.
\end{cases} \tag{22c}
\end{align}

1) **Local Solutions in Regions 1 and 3:** The vector field in Region 1 is characterized by the linear differential equation (22a) whose general solution is given as follows:

\[ x_t = \left(k_1 e^{2at} - 1\right)/2a, \quad p_t = \left(k_2 e^{-2at} - 1\right)/2a \tag{23} \]

where $k_1$ and $k_2$ are constants. On the other hand, the vector field in Region 3 characterized by (22c) is nonlinear. However, (22c) belongs to the class of scalar Riccati differential equations and admits an analytical solution given as follows:

\[ x_t = \frac{1}{\gamma} \left( a - c - \frac{2c}{k_3 e^{2ct} + 1} \right) := \hat{h}(t, k_3) \tag{24a} \]

\[ p_t = \frac{k_4 \left(k_3 e^{2ct} + 1\right)^2}{2ck_4 e^{2ct}} + \frac{(1 + \alpha \gamma) \left(k_3 e^{2ct} + 1\right)}{2ck_4 e^{2ct}} := \hat{h}(t, k_3, k_4) \tag{24b} \]

where, $k_3$ and $k_4$ are constants and $c = \sqrt{\alpha^2 + \gamma}$.

2) **Stationary Points:** The location of a stationary point in the phase portrait changes depending on the value of $\alpha$. Noticing that $0 < (a + \sqrt{\alpha^2 + \gamma})^2/\gamma^2 < 1/4a^2$ for all $a < 0$ and $\gamma > 0$, the following three cases can occur:

1) Case A: $1/4a^2 < \alpha$. In this case, the phase portrait has a unique stationary point in Region 1 located at $E = (x_*, p_*) = (-1/2a, -1/2a)$. It is not possible for $f_2$ to have a stationary point in Region 2 no matter what value of $u_t \in [0, \gamma]$ is chosen. A stationary point cannot exist in Region 3 either.

2) Case B: $(a + \sqrt{\alpha^2 + \gamma})^2/\gamma^2 \leq \alpha < 1/4a^2$. In this case, a stationary point cannot exist in Region 1 or in Region 3. However, the point $E = (x_*, p_*) = (\sqrt{\alpha}/\gamma, \sqrt{\alpha})$ in Region 2 is a stationary point if $u_t$ is set to $u^* = 2a/\sqrt{\alpha^2 + 1}/\alpha$. From the present assumption that $(a + \alpha \gamma/\sqrt{\alpha^2 + \gamma})^2/\gamma^2 \leq \alpha \leq 1/4a^2$, it can be shown that the value of $u^*$ in $u^* = 2a/\sqrt{\alpha^2 + 1}/\alpha$ satisfies $0 \leq u^* \leq \gamma$. No other point in Region 2 can be a stationary point.

3) Case C: $\alpha < (a + \sqrt{\alpha^2 + \gamma})^2/\gamma^2$. In this case, the phase portrait has a unique stationary point $(x_*, p_*) = \left((a + \sqrt{\alpha^2 + \gamma})/\gamma, \alpha \gamma/\sqrt{\alpha^2 + \gamma}\right)$ in Region 3. No stationary point can exist in Regions 1 and 2.

The vector field in each case is depicted in Fig. 3.

3) **Switching Behavior:** Now we study the behavior of the hybrid system (22) near the switching surface. Noticing that $S$ is a level set of the function $V(x, p) = xp$, this can be studied by checking the signs of the Lie derivatives $L_{f_1} V$, $L_{f_2} V$, and $L_{f_3} V$ evaluated on $S$. Notice the following:

\[ L_{f_1} V = \frac{\partial V}{\partial x} \dot{x} + \frac{\partial V}{\partial p} \dot{p} = px + xp = p - x \]

\[ L_{f_2} V = px + xp = x^2 pu + p - x - \alpha xu \]

\[ L_{f_3} V = px + xp = \gamma x^2 p + p - x - \alpha \gamma x \]
Therefore, on the surface $S$ (i.e., when $px = \alpha$), we have $L f_1 V|_S = L f_2 V|_S = \alpha/x - x$. This indicates that all the vector fields $f_1$, $f_2$, and $f_3$ define a consistent direction with respect to $S$ everywhere on $S$. Namely, they cross $S$ "upward" in the portion where $x < \sqrt{\alpha}$, "downward" where $x > \sqrt{\alpha}$, and are tangential to $S$ at the point $K = (\sqrt{\alpha}, \sqrt{\alpha})$. The point $K$ becomes a stationary point when $(a + \sqrt{\alpha^2 + \gamma})^2/\gamma^2 \leq \alpha \leq 1/4\alpha^2$ (Case B) and $u_t$ is set to be $u^t = 2a/\sqrt{\alpha} + 1/\alpha$. This coincides with $E$ defined by $E = (x_0, p_0) = (\sqrt{\alpha}, \sqrt{\alpha})$. The analysis above has the important implication that the phase portraits in Fig. 3 are free from the "chattering" solutions in cases, and that the solution concept of Carathéodory [24] is sufficient to describe the solutions. However, the uniqueness of the solution is lost in Case B. To see this, consider the family of the trajectories $(x_t, p_t)$ that "stay" on $E = K$ for an arbitrary duration as follows:

1. $(x_t, p_t)$ solves (22a) or (22c) for $0 \leq t \leq t'$ with $(x', p') = (\sqrt{\alpha}, \sqrt{\alpha})$.
2. $(x_t, p_t) = (\sqrt{\alpha}, \sqrt{\alpha})$ for $t' \leq t \leq t''$.
3. $(x_t, p_t)$ solves (22a) or (22c) for $t'' \leq t_1$ with $(x', p') = (\sqrt{\alpha}, \sqrt{\alpha})$.

It is easy to check that all these trajectories are Carathéodory solutions to the canonical equations regardless of the choice of $t'$ and $t''$.

B. Analytical Solution

Using the phase portraits depicted in Fig. 3, we now solve the boundary value problem (20) and (21) with the initial state condition $x_{t_0} = x_0(\geq 0)$ and the terminal costate condition $p_{t_1} = 0$. In what follows, the solution to this boundary value problem is simply referred to as the optimal solution. It is convenient to consider Cases A, B, and C separately.

1) Case A: In this case, the initial coordinate of $(x^*_0, p^*_0)$ of the optimal solution is either in the green or the pink regions illustrated in Fig. 3(a). The boundaries of these regions are defined by the switching surface $S$ and the separatrices converging to the point $E$.

Subcase A-1 ($(x^*_0, p^*_0)$ is in the green region): Consider the particular solution to the vector field $f_1$ satisfying the boundary conditions $x_{t_0} = x_0$ and $p_{t_1} = 0$ as follows:

$$\bar{x} = [(2ax_0 + 1)e^{-at} - 1]/2a =: \tilde{g}(t, x_0) \quad (25a)$$

$$\bar{p}_t = [-2e^{-at} - 1]/2a =: \tilde{h}(t, t_1). \quad (25b)$$

If we compute $(\bar{x}_0, \bar{p}_0)$ from (25) and find it is in Region 1 (i.e., $\bar{x}_0, \bar{p}_0 \leq \alpha$), then Subcase A-1 applies. Since the optimal solution is entirely in Region 1 there is no switching.

Subcase A-2 ($(x^*_0, p^*_0)$ is in the pink region): In this case, there exists a switching time $t' \in (t_0, t_1)$. To compute $t'$, notice that a particular solution (24a) satisfying the initial condition $x_{t_0} = x_0$ is given by $\tilde{x} = \tilde{g}(t, k)$ where $c = \sqrt{\alpha^2 + \gamma}$ and $k_3 = c - a + x_0\gamma/c + a - x_0\gamma e^{2a(t-t_0)}/\alpha$. On the other hand, the particular solution (23) satisfying $p_{t_1} = 0$ is given by (25b). Thus, at the switching time $t'$, it must be that $\tilde{x} \tilde{p}' = \alpha$, or

$$\quad \tilde{g}(t, k_3) \tilde{h}(t', t_1) = \alpha \quad (26)$$

Therefore, $t'$ can be computed by solving (26).

2) Case B: In this case, the initial coordinate $(x^*_0, p^*_0)$ of the optimal solution is in the colored region in Fig. 3(b). Let $x_K$ be the $x$-coordinate at which a particular solution (23) to the vector field $f_1$ that passes through $K = E = (\sqrt{\alpha}, \sqrt{\alpha})$ at a certain time $t'(< t_1)$ reaches at the terminal time $t_1$. It is straightforward to show that $x_K = 2\alpha + 2\sqrt{\alpha}$. The time $t'$ of passing $K = E$ can also be computed as $t_1 - t' = 1/2\ln(2a\sqrt{\alpha} + 1)$.

Subcase B-1 ($(x^*_0, p^*_0)$ is in the green region or on the green curve): Notice that the solution $(\bar{x}_t, \bar{p}_t)$ to $f_1$ with boundary conditions $x_{t_0} = x_0$ and $p_{t_1} = 0$ is still given by (25). If $\tilde{x}_t \leq x_K$ occurs, the optimal solution is entirely in the green region and is characterized by (25). There is no switching.

Subcase B-2 ($(x^*_0, p^*_0)$ is in the pink region): This case occurs when $\tilde{x}_{t_1} > x_K, \tilde{x}_{t_0} \leq \alpha$, and $\alpha > \sqrt{\alpha}$. The optimal solution is entirely in the pink region and is characterized by (25). Hence, there is no switching.

Subcase B-3 ($(x^*_0, p^*_0)$ is on the orange curve): This case occurs when (25) satisfies $\tilde{x}_{t_1} > x_K, \tilde{x}_{t_0} \leq \alpha$, and $\alpha < \sqrt{\alpha}$. In this case, the solution (25) over $t_0 \leq t \leq t_1$ is not contained in Region 1 and thus it is not a valid solution to the boundary problem of our interest. The optimal solution in this case is depicted as orange and green curves in Fig. 3(b). First, it follows the orange curve from $x_{t_0} = x_0$ to $x_{t_1} = \sqrt{\alpha}$, stay on $E = K$ for $t' \leq t \leq t'$, and then follows the green trajectory from $\tilde{x}_t = \sqrt{\alpha}$ to $\tilde{x}_t = x_K$. From (25a), the time $t'$ can be computed from $\sqrt{\alpha} = \tilde{g}(t' \leq t, x_0)$. There are two switches in the optimal control input: a switch from $u = 0$ to $u = u^*$ at $t = t'$, and a switch from $u = u^* \rightarrow u = 0$ at $t = t'$.

Subcase B-4 ($(x^*_0, p^*_0)$ is in the gray region): This case occurs when (25) satisfies $\tilde{x}_{t_1} > x_K$ and $\tilde{x}_{t_0} \leq \alpha$, and (26) has a solution $t'$ in $[t_0, t_1]$. In this case, a simple switching from the gray region to the pink region occurs at $t'$. If (26) does not have a solution $t'$ in $[t_0, t_1]$. In this case, the optimal solution follows the black trajectory from $x_{t_0} = x_0$ to $x_{t_1} = \sqrt{\alpha}$, stays at $E = K$ for $t' \leq t \leq t'$, and then follows the green trajectory from $\tilde{x}_t = \sqrt{\alpha}$ to $\tilde{x}_t = x_K$. From (24a), the time $t'$ can be computed from $\sqrt{\alpha} = \tilde{g}(t' \leq t, k_3)$ where $k_3 = c - a + x_0\gamma/c + a - x_0\gamma e^{2a(t-t_0)}/\alpha$. The optimal control input switches twice: a switch from $u = \gamma$ to $u = u^*$ at $t'$, and a switch from $u = u^* \rightarrow u = 0$ at $t'$. Notice that $\tilde{x}_{t_1} > x_K, \tilde{x}_{t_0} \geq \alpha$, and $\alpha < \sqrt{\alpha}$. The optimal solution is entirely in the pink region and is characterized by (25). Hence, there is no switching.

3) Case C: In this case, the initial state-costate pair $(x^*_0, p^*_0)$ of the optimal solution can belong to four different regions indicated by four different colors in Fig. 3(c). The boundaries of the pink region are defined by the switching surface $S$ and the separatrices converging to the point $E$.

Subcase C-1 ($(x^*_0, p^*_0)$ is in the green region): Let $x_K = 2\alpha + 2\sqrt{\alpha}$ be the $x$-coordinate shown in Fig. 3(c). Consider once again the
trajectory (25) solving $f_1$ with the boundary conditions $x_{t_0} = x_0$ and $p_{t_1} = 0$. If we have $\bar{x}_1 \geq x_K$, then the optimal solution is entirely in the yellow region. No switching occurs in this case.

**Subcase C-2** ($\{x_{t_0}, p_{t_1}\}^*$ is in the blue region): Consider (25) again. If $\bar{x}_1 > x_K$, $x_0 > \sqrt{\alpha}$, and $\bar{x}_0 p_{t_0} \leq \alpha$, then the trajectory (25) is entirely in the blue region. No switching occurs in this case.

**Subcase C-3** ($\{x_{t_0}, p_{t_1}\}^*$ is in the pink region): In this case, a switching occurs once. Let $t' \in (t_0, t_1)$ be the switching time. Then, the optimal solution follows the trajectory of the form $\bar{x}_i = \hat{g}(t, k_3)$ and $\bar{p}_i = \hat{h}(t, k_3, k_4)$ for $t_0 \leq t \leq t'$, and $\bar{p}_i = \hat{h}(t, t_1)$ for $t' \leq t \leq t_1$. Thus, Subcase C-3 applies if the following sequence of events is in terms of unknowns $p_{t_0}$, $k_3$, $k_4$, and $t'$ admits a solution such that $t_0 < t' < t_1$ and $x_{t_0} p_{t_0} > \alpha$:

\begin{align}
\bar{x}_{t_0} &= x_0 = \hat{g}(t_0, k_3), \quad p_{t_0}^* = \hat{h}(t_0, k_3, k_4) \quad (27a) \\
\hat{g}(t', k_3) \hat{h}(t', t_1) &= \alpha \quad (27b) \\
\hat{h}(t', k_3, k_4) &= \hat{h}(t', t_1). \quad (27c)
\end{align}

Condition (27b) ensures that $\bar{x}_{t_1} p_{t_1} = \alpha$, and (27c) ensures that $\bar{p}_{t_1} = \bar{p}_{t_1}$ (i.e., the transition from $\bar{p}_{t_1}$ to $\bar{p}_{t_1}$ continues).

**Subcase C-4** ($\{x_{t_0}, p_{t_1}\}^*$ is in the purple region): Switching occurs twice in this case. Let $t'$ and $t''$ be the first and the second switching times. The optimal solution proceeds the trajectory of the form $\bar{x}_i = (k_1 e^{2 t_0 - 1}/2a)$ and $\bar{p}_i = (k_2 e^{-2 t_0 - 1})/2a$ for $t_0 \leq t \leq t'$ and satisfies $\hat{g}(t, k_3, k_4)$ and $\hat{h}(t, t_1, k_3)$ for $t' \leq t \leq t''$. Therefore, Subcase C-4 applies if the following sequence of events is in terms of unknowns $k_1$, $k_3$, $k_4$, $p_{t_0}$, $t'$, and $t''$ admits a solution such that $x_{t_0} p_{t_0} \leq \alpha$ and $t_0 \leq t' < t'' < t_1$:

\begin{align}
\bar{x}_{t_0} &= x_0 = (k_1 e^{2 t_0 - 1}/2a, \quad p_{t_0}^* = (k_2 e^{-2 t_0 - 1})/2a) \quad (28a) \\
\left[\frac{[k_1 e^{2 t_0 - 1} - 1]}{2a}\right] &\left[\frac{[k_2 e^{-2 t_0 - 1} - 1]}{2a}\right] = \alpha \quad (28b) \\
\left[\frac{\bar{k}_1 e^{-2 t_0} - 1}{2a}\right] &\left[\frac{\bar{k}_2 e^{2 t_0} - 1}{2a}\right] = \hat{g}(t', k_3) \quad (28c) \\
\hat{h}(t', k_3, k_4) &= \hat{h}(t', t_1) \quad (28d) \\
\hat{h}(t', t_1) &= \hat{h}(t', t_1). \quad (28e)
\end{align}

Conditions (28a) and (28d) ensure that $\bar{x}_{t_1} p_{t_1} = \alpha$ and $\bar{x}_{t_1} p_{t_1} = \alpha$ (i.e., switching happens on the switching surface). Conditions (28b), (28c), and (28e) ensure that the trajectory is continuous at switching times.

**V. TIME-INVARIANT SOLUTIONS**

We now turn our attention to the problem of finding the optimal time-invariant channel gain $G \in \mathbb{R}^{m \times n}$ as formulated in (15). Since $(A, B)$ is controllable and $(A, C)$ is detectable (we have assumed that $A$ is Hurwitz for every $C \in \mathbb{R}^{m \times n}$), the algebraic Riccati equation $AX + XA^T - XCX + BB^T = 0$ admits a unique positive semidefinite solution, which is positive definite [25, Th. 13.7, Corollary 13.8]. In this case, it can also be shown that the solution $X_0$ to the Riccati differential equation (16b) with the initial condition $X_0 = X_0, \geq 0$ satisfies $X_0 \rightarrow X$ as $t \rightarrow +\infty$ (e.g., [26, Th. 10.10]), where $X$ is the unique positive definite solution to $AX + XA^T - XCX + BB^T = 0$. Therefore, it follows from the convergence of Cesàro mean that $1/t_j - t_0 \int_{t_0}^{t_j} 1/t_0 \int_{t_0}^{t_j} \text{Tr}(X(t)XC(t))dt \rightarrow \text{Tr}(X(t_0))$ and $1/t_j - t_0 \int_{t_0}^{t_j} \text{Tr}(X(t))dt \rightarrow 0$ as $t_1 \rightarrow +\infty$. Thus, (16) can be simplified as follows:

\[\min_{C \in \mathbb{R}^{m \times n}, X \in \mathbb{S}^n_+} \text{Tr}(X) + \alpha \text{Tr}(XCC^T) \quad (29a)\]

subject to

\[AX + XA^T - XCX + BB^T = 0 \quad (29b)\]

As the main result of this section, we show that (29) can be reformulated as an equivalent semidefinite program with a rank constraint. The result is summarized in the next theorem.

**Theorem 2**: Suppose $A \in \mathbb{R}^{m \times n}$ is Hurwitz. For any given positive constants $\alpha$ and $\gamma$, the following statements hold:

i) The optimal value of the problem (15) coincides with the value of the semidefinite program with a rank constraint:

\[\min_{X, Y \in \mathbb{S}^n_+} \text{Tr}(X) + \alpha \text{Tr}(B^T Y B) + 2\alpha \text{Tr}(A) \quad (30a)\]

subject to

\[\begin{bmatrix} YA + A^T Y & -Y & -Y \end{bmatrix} \begin{bmatrix} B^T Y & -Y \end{bmatrix} \leq 0 \quad (30b)\]

and

\[Z := \begin{bmatrix} X & I \\ I & Y \end{bmatrix} \succeq 0 \quad (30c)\]

ii) An optimal solution $(X^*, Y^*)$ to (30) exists and satisfies $X^* \succ 0$ and $Y^* \succ 0$. Moreover, any matrix $C^* \in \mathbb{R}^{m \times n}$ satisfying

\[C^*^T C^* = Y^* A + A^T Y^* + Y^* B B^T Y^* \quad (31)\]

is an optimal solution to (15).

iii) Suppose that a semidefinite program obtained by removing the rank constraint (30e) from (30) admits an optimal solution satisfying $X^* \succ 0, Y^* \succ 0$, and rank $\begin{bmatrix} X^* & I \\ I & Y^* \end{bmatrix} = n$. Then, any matrix $C^* \in \mathbb{R}^{m \times n}$ satisfying (31) is an optimal solution to (15).

**Remark 3**: In Section V-A below, we prove the equivalence between the original problem (15) and the optimization problem (30). Notice that off-the-shelf solvers are not applicable to (30) because of the nonconvex rank constraint. Remarkably, however, in numerous numerical experiments we have performed (see Section V-B), the convex relaxation obtained by dropping the rank constraint (30e) always admitted a solution satisfying (30e). Thus, as per statement (iii) of Theorem 2, an optimal solution $C^* \in \mathbb{R}^{m \times n}$ to the original problem (15) was always computable by solving the semidefinite program (30a),(30d). Currently, it is not known to us whether the convex relaxation is always exact.

**A. Proof of Theorem 2**

We will show the equivalence between (29) and (30). First, notice that the constraint (29b) implies that $X > 0$, and

\[\text{Tr}(XCC^T) = \text{Tr}(XCX^T)X^{-1} = \text{Tr}(AX + XA^T BB^T)X^{-1} = 2\text{Tr}(A) + \text{Tr}(B^T X^{-1} B). \]

Introducing $Y := X^{-1}$, (29) can be written equivalently as follows:

\[\min_{X \in \mathbb{S}^n_+} \text{Tr}(X) + \alpha \text{Tr}(B^T Y B) + 2\alpha \text{Tr}(A) \quad (32a)\]

subject to

\[AX + XA^T - XCX + BB^T = 0 \quad (32b)\]

\[YA + A^T Y - C^T C + YBB^T Y = 0 \quad (32c)\]

\[X = Y^{-1}, \quad C^T C \preceq \gamma I \quad (32d)\]

with respect to the variables $X \in \mathbb{S}^n_+, Y \in \mathbb{S}^n_+$, and $C \in \mathbb{R}^{m \times n}$. Notice that (32b) and (32c) are redundant conditions under the constraint
$X = Y^{-1}$. Next, we claim that (32) is equivalent to the following optimization problem:

$$\begin{align*}
\min_{X,Y} & \quad \text{Tr}(X) + \alpha \text{Tr}(B^T Y B) + 2\alpha \text{Tr}(A) \\
\text{s.t.} & \quad AX + XA^T + BB^T \succeq 0 \\
& \quad [YA + A^TY - \gamma I \quad YB] \preceq 0 \\
& \quad Z \succeq 0 \\
& \quad \text{rank}(Z) = n. \\
& \quad X \succ 0, \ Y \succ 0.
\end{align*}$$

(34a)

(34b)

(34c)

(34d)

(34e)

It is clear that the constraints (32b)–(32d) imply the constraints (33b)–(33d). Conversely, for any $(X, Y)$ satisfying (33b)–(33d), it is always possible to construct a tuple $(X, Y, C)$ satisfying (32b)–(32d) by choosing $C \in \mathbb{R}^{n \times n}$ to satisfy $C^T C = YA + A^TY + YBB^TY$. Applying the Schur complement formula to (33c), and noticing that $X = Y^{-1} \succ 0$ is equivalent to $Z \succeq 0$ and $\text{rank}(Z) = n$, (33) can be written as follows:

$$\begin{align*}
\min_{X,Y,C} & \quad \text{Tr}(X) + \alpha \text{Tr}(B^T Y B) + 2\alpha \text{Tr}(A) \\
\text{s.t.} & \quad AX + XA^T + BB^T \succeq 0 \\
& \quad [YA + A^TY - \gamma I \quad YB] \succeq 0 \\
& \quad Z \succeq 0 \\
& \quad \text{rank}(Z) = n. \\
& \quad X \succ 0, \ Y \succ 0.
\end{align*}$$

(34a)

(34b)

(34c)

(34d)

(34e)

It is left to show that (30) has an optimal solution $(X^*, Y^*)$ and that $X^* \succ 0, Y^* \succ 0$. The existence of an optimal solution is guaranteed by Weierstrass’ theorem [27, Prop. A.8], since the feasible domain for $(X, Y)$ is closed and the objective function is coercive. To show $X^* \succ 0$ must be the case, consider a sequence of feasible points $(X_n, Y_n)$ such that $X_n \rightarrow X_0$ as $n \rightarrow \infty$ for some singular $X_0 \succeq 0$. Then, from the constraint (34d), it is necessary that the maximum singular value $\sigma(Y_n)$ satisfies $\sigma(Y_n) \rightarrow \infty$ as $n \rightarrow \infty$. This implies $\text{Tr}(B^T Y_n B) \rightarrow \infty$ as $n \rightarrow \infty$ and, thus, the objective function (34a) tends to infinity. Thus, by continuity of the objective function, it cannot be that $X^* = X_0$. The necessity of $Y^* \succ 0$ can be shown similarly. Hence (34e) is a redundant condition and can be removed. Therefore, we have shown the equivalence between (29) and (30), establishing statement (i). Statements (ii) and (iii) also follow from the argument above.

B. Numerical Experiments

To demonstrate the effectiveness of the convex relaxation presented in Theorem 2, we analyzed the numerical solutions to the semidefinite program (30a)–(30d) for a large number of randomly generated matrix pairs $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times n}$. To generate a random stable state space model, we used MATLAB’s $\text{rss}$ function. To solve (30a)–(30d), we used a SDP solver SDPT3 [28]. With $n = 15$, $\alpha = 0.01$, and $\gamma = 100$, we generated 1000 pairs of random matrices $(A, B)$. In each simulation, we observed that the solution to (30a)–(30d) satisfied the rank condition (30e) up to the precision of the SDP solver. Fig. 4(a) shows a typical outcome of the eigenvalues $\lambda_i$ of the matrix $Z \in \mathbb{S}_{++}^n$. It can be seen that the last $n$ eigenvalues are negligible compared with the first $n$.

In Fig. 4(b), we plot a typical outcome of the eigenvalues of $C^T C$ computed by (31). Again, we set $n = 15$, and the upper limit of the channel gain is set to $\gamma = 100$. We solved (30a)–(30d) with different values of $\alpha$. As expected, the optimal channel gain tends to decrease as $\alpha$ increases. It is also noteworthy that both the lower saturation $\lambda_1(C^T C) = 0$ and the upper saturation $\lambda_n(C^T C) = 100$ can occur.

C. Scalar Case

We finally consider a special case of (15) with $n = 1$ to show that: 1) the exactness of the SDP relaxation discussed in Theorem 2 (iii) holds in the scalar case, and 2) stationary points in the phase portrait we used to analyze the finite-horizon optimal channel gain control problem characterize the optimal time-invariant solution to the infinite-horizon channel gain control problem in the scalar case. As in Section IV, we assume $A = a I$, $B = 1$, $\alpha > 0$, and $\gamma > 0$. The next proposition shows that the rank condition (30e) is automatically satisfied even if it is not explicitly included in (30) in the scalar case.

Proposition 2: An optimal solution $(x^*, y^*)$ to the semidefinite program

$$\begin{align*}
\min_{x \geq 0, y \geq 0} & \quad x + ay + 2ax \\
\text{s.t.} & \quad 2ax + 1 \geq 0 \\
& \quad \begin{bmatrix} 2ay - \gamma & y \\ y & -1 \end{bmatrix} \preceq 0, \quad \begin{bmatrix} x & 1 \\ 1 & y \end{bmatrix} \succeq 0
\end{align*}$$

satisfies $x^* \succ 0, y^* \succ 0$ and rank $\begin{bmatrix} x^* & 1 \\ 1 & y^* \end{bmatrix} = 1$.

Proof: Omitted for brevity.

It is also fruitful to obtain an explicit solution to (29) in the scalar case. Setting $u = C^T$, (29) can be written as follows:

$$\begin{align*}
\min_{u,x} & \quad x(1 + \alpha u) \\
\text{s.t.} & \quad 2ax - u^2 + 1 = 0, \quad x \geq 0 \\
& \quad 0 \leq u \leq \gamma.
\end{align*}$$

(36a)

(36b)

(36c)

Equation (36b) can be solved explicitly as $x = 1/\sqrt{\alpha^2 + u - \alpha}$, and the constraint (36c) can be written in terms of $x$ as $x + \sqrt{\alpha^2 + \gamma^2}/\gamma \leq x \leq 1/2a$. Expressing the objective function (36a) in $x$, the problem (36) is simplified to the following:

$$\begin{align*}
\min_{x} & \quad x + 2ax + \alpha/x \\
\text{s.t.} & \quad (x + \sqrt{\alpha^2 + \gamma^2})/\gamma \leq x \leq -1/2a
\end{align*}$$

(37a)

(37b)

which can be solved easily. From this analysis, it can be shown that the optimal solution to (36) is obtained as follows:

1) Case A: $1/4a^2 < \alpha$. In this case, $(x^*, u^*) = (-1/2a, 0)$.
2) Case B: $(x^*, u^*) = (\sqrt{\alpha(2a + 1)/\alpha})$.
3) Case C: $(x^*, u^*) = (\gamma + \sqrt{\alpha^2 + \gamma^2})/\gamma$. In this case, $(x^*, u^*) = (x^*, u^*)$. 

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Notably, these solutions coincide with the stationary points in the phase portrait we obtained in Cases A and C.

VI. CONCLUSION

In this article, we formulated a continuous-time optimal channel gain control problem for minimum-information Kalman–Bucy filtering. Our special focus has been on the optimal time-varying solution to finite-horizon problems for scalar processes and the optimal time-invariant solution to infinite-horizon problems for vector processes. The presented results can be extended to multiple directions in the future.

1) As a natural generalization of Section IV, the optimal time-varying channel gain for multidimensional source processes should be investigated as future work.

2) Although we considered the optimal time-invariant solutions in Section V, it remains to show that there always exists a time-invariant solution that minimizes the average cost (15a) over an infinitely long time horizon. The existence of a time-invariant optimal solution in discrete-time setting has been shown in [29].

3) It remains to find a formal proof of, or a counterexample to disprove, the exactness of the SDP relaxation presented in Theorem 2.

4) The problem formulation should be extended to controlled source processes. Such a problem has been considered in a discrete-time setting [5] where directed information has been used in place of mutual information. Directed information in continuous time has been introduced in [30].

5) A coding-theoretic interpretation (operational meaning) of the problem studied in this article needs to be clarified in the future.

REFERENCES

[1] S. Tatikonda, “Control under communication constraints,” Ph.D. thesis, Dept. of Elect. Eng. and Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, 2000.
[2] D. Teneketzis, “On the structure of optimal real-time encoders and decoders in noisy communication,” IEEE Trans. Inf. Theory, vol. 52, no. 9, pp. 4017–4035, Sep. 2006.
[3] I. Silva, M. S. Derpich, and J. Østergaard, “A framework for control system design subject to average data-rate constraints,” IEEE Trans. Autom. Control, vol. 56, no. 8, pp. 1886–1899, Aug. 2011.
[4] S. Yüksel, “Jointly optimal LQG quantization and control policies for multi-dimensional systems,” IEEE Trans. Autom. Control, vol. 59, no. 6, pp. 1612–1617, Jun. 2014.
[5] T. Tanaka, P. M. Esfahani, and S. K. Mitter, “LQG control with minimum directed information: Semidefinite programming approach,” IEEE Trans. Autom. Control, vol. 63, no. 1, pp. 37–52, Jan. 2018.
[6] V. Kostina and B. Hassibi, “Rate-cost tradeoffs in control,” IEEE Trans. Autom. Control, vol. 64, no. 11, pp. 4525–4540, Nov. 2019.
[7] N. Guo and V. Kostina, “Optimal causal rate-constrained sampling of the Wiener process,” IEEE Trans. Autom. Control, vol. 67, no. 4, pp. 1776–1791, Apr. 2022.
[8] S. Tatikonda, A. Sahai, and S. Mitter, “Stochastic linear control over a communication channel,” IEEE Trans. Autom. Control, vol. 49, no. 9, pp. 1549–1561, Sep. 2004.
[9] M. S. Derpich and J. Østergaard, “Improved upper bounds to the causal quadratic rate–distortion function for Gaussian stationary sources,” IEEE Trans. Inf. Theory, vol. 58, no. 5, pp. 3131–3152, May 2012.
[10] C. D. Charalambous, P. A. Stavrou, and N. U. Ahmed, “Nonanticipative rate distortion function and relations to filtering theory,” IEEE Trans. Autom. Control, vol. 59, no. 4, pp. 937–952, Apr. 2014.
[11] T. Tanaka, K.-K. K. Kim, P. A. Parrilo, and S. K. Mitter, “Semidefinite programming approach to Gaussian sequential rate-distortion trade-offs,” IEEE Trans. Autom. Control, vol. 62, no. 4, pp. 1896–1910, Apr. 2017.
[12] P. A. Stavrou, T. Charalambous, C. D. Charalambous, and S. Loyka, “Optimal estimation via nonanticipative rate distortion function and applications to time-varying Gauss–Markov processes,” SIAM J. Control Optim., vol. 56, no. 5, pp. 3731–3765, 2018.
[13] T. E. Duncan, “On the calculation of mutual information,” SIAM J. Appl. Math., vol. 19, no. 1, pp. 215–220, 1970.
[14] T. Tanaka, M. Skoglund, and V. Ugrinovskii, “Optimal sensor design and zero-delay source coding for continuous-time vector Gauss–Markov processes,” in Proc. IEEE 56th Annu. Conf. Decis. Control, 2017, pp. 3992–3997.
[15] D. Liberzon, Calculus of Variations and Optimal Control Theory: A Concise Introduction. Princeton, NJ, USA: Princeton Univ. Press, 2011.
[16] V. Zinage, T. Tanaka, and V. Ugrinovskii, “Optimal sensor gain control for minimum-information estimation of continuous-time Gauss–Markov processes,” in Proc. Amer. Control Conf., 2022, pp. 60–66.
[17] R. M. Gray, Entropy and Information Theory, 2nd ed. New York, NY, USA: Springer, 2011.
[18] T. M. Cover and J. A. Thomas, Elements of Information Theory. New York, NY, USA: Wiley, 1991.
[19] N. Wan, D. Li, and N. Hovakimyan, “Fundamental limitations of control and filtering in continuous-time systems: An information-theoretic analysis,” 2022, arXiv:2201.00995.
[20] R. Bansal and T. Başar, “Simultaneous design of measurement and control strategies for stochastic systems with feedback,” Automatica, vol. 25, no. 5, pp. 679–694, 1989.
[21] A. Filippov, “On certain questions in the theory of optimal control,” J. Soc. Ind. Appl. Math., Ser. A: Control, vol. 1, no. 1, pp. 76–84, 1962.
[22] M. Athans and P. L. Falb, Optimal Control: An Introduction to The Theory and Its Applications. Mineola, NY, USA: Dover Pub., 2013.
[23] D. Angeli and E. D. Sontag, “Monotone control systems,” IEEE Trans. Autom. Control, vol. 48, no. 10, pp. 1684–1698, Oct. 2003.
[24] J. Cortes, “Discontinuous dynamical systems,” IEEE Control Syst. Mag., vol. 28, no. 3, pp. 36–73, Jun. 2008.
[25] K. Zhou, J. C. Doyle, and K. Glover, Robust and Optimal Control. New Jersey, NJ, USA: Prentice Hall, 1996.
[26] R. R. Bitmead and M. Gevers, “Riccati difference and differential equations: Convergence, monotonicity and stability,” in The Riccati Equation. Berlin, Germany: Springer, 1991, pp. 263–291.
[27] D. Bertsekas, Nonlinear Programming. Belmont, MA, USA: Athena Scientific, 1995.
[28] R. H. Tütüncü, K.-C. Toh, and M. J. Todd, “Solving semidefinite-quadratic-linear programs using SDPT3,” Math. Program., vol. 95, no. 2, pp. 189–217, 2003.
[29] T. Tanaka, “Semidefinite representation of sequential rate-distortion function for stationary Gauss–Markov processes,” in Proc. IEEE Conf. Control Appl., 2015, pp. 1217–1222.
[30] T. Weissman, Y.-H. Kim, and H. H. Permuter, “Directed information, causal estimation, and communication in continuous time,” IEEE Trans. Inf. Theory, vol. 59, no. 3, pp. 1271–1287, Mar. 2013.