A Frequency-Domain Characterization of Optimal Error Covariance for the Kalman–Bucy Filter

Song Fang, Hideaki Ishii, Jie Chen, and Karl Henrik Johansson

Abstract— In this paper, we discover that the trace of the division of the optimal output estimation error covariance over the noise covariance attained by the Kalman–Bucy filter can be explicitly expressed in terms of the plant dynamics and noise statistics in a frequency-domain integral characterization. Towards this end, we examine the algebraic Riccati equation associated with Kalman–Bucy filtering using analytic function theory and relate it to the Bode integral. Our approach features an alternative, frequency-domain framework for analyzing algebraic Riccati equations and reduces to various existing related results.

I. INTRODUCTION

The Kalman–Bucy filter [1] is the optimal state estimator for linear continuous-time systems with white Gaussian noises. It is well known that the optimal state estimation error covariance attained by the Kalman–Bucy filter in steady state satisfies the algebraic Riccati equation [2]–[4], which, however, does not provide an analytical expression of the error covariance. Over the years, various upper bounds and lower bounds have been obtained for the error covariance (see, e.g., [5] and the references therein). Meanwhile, nonrecursive algebraic solutions have been obtained as well (see, e.g., [3], [6], [7] and the references therein).

In this paper, we develop an alternative approach towards this problem from the viewpoint of the Bode integral [8]. Bode integral was the first and arguably the best known result to analyze the fundamental limitations of feedback control systems using analytic function theory [9]–[15]. In particular, the Bode integral is implied by the Poisson–Jensen formula [16]. This line of research has been of continuing interest to the control community [17], [18], and was related to information theory in recent years as well [19]–[27]; such results together with [28] inspired us to reexamine the Kalman–Bucy filter (in steady-state) using analytic function theory, e.g., the Poisson–Jensen formula, which will be seen to play an essential role in our analysis.

In particular, we first develop from the Poisson–Jensen formula a counterpart of Jensen's formula, based on which

Song Fang and Karl Henrik Johansson are with the School of Electrical Engineering and Computer Science, Royal Institute of Technology, sonf@kth.se, kallej@kth.se

Hideaki Ishii is with the Department of Computer Science, Tokyo Institute of Technology, ishii@c.titech.ac.jp

Jie Chen is with the Department of Electronic Engineering, City University of Hong Kong, jichen@cityu.edu.hk

we obtain an analytical expression of the trace of the division of the optimal output estimation error covariance over the noise covariance by the Kalman–Bucy filter; this is given in terms of the plant dynamics, e.g., the unstable poles, and the noise statistics in a frequency-domain integral characterization. The analysis mainly concerns examining the algebraic Riccati equation associated with the Kalman–Bucy filter using analytic function theory.

We compare our result with the existing related results in [29]–[33], and our integral characterization is seen to be the most generic and reduces to all of them. Specifically, the case when the plant is stable and the plant output is a scalar process was considered in [29]. The authors of [30] then analyzed the case when the plant output is scalar but the plant is not necessarily stable. The following results in [31] considered the case when the process noise is relatively small in variance (and in the limit, zero) compared with the observation noise. Most recently, the case when the plant output is a vector process while the observation noise is with an identity covariance matrix was investigated in [33]. In addition, discussions on the dual problem in control can be found in [30], [32].

The remainder of the paper is organized as follows. Section II introduces the technical preliminaries. Section III introduces a counterpart of Jensen's formula. In Section IV, we examine the Kalman–Bucy filter using algebraic function theory and presents an integral characterization of the optimal output estimation error covariance. Relevant discussions and interpretations are also presented. Concluding remarks are given in Section IV.

For the proofs omitted in this conference version, refer to the extended version [34]. For analysis and discussions on the discrete-time Kalman filters, see our parallel work presented in [35].

II. PRELIMINARIES

In this section, we introduce some relevant notions from stochastic processes, and briefly review the basic properties of the Kalman–Bucy filter.

A. Notations and Basic Concepts

In this paper, we consider real-valued continuous zero-mean random variables and vectors, as well as continuous-time stochastic processes. We denote random variables and vectors using boldface letters. The logarithm $\ln$ is defined with base e, and all functions are assumed to be measurable.

A zero-mean stochastic process $\{\mathbf{x}(t)\}, \mathbf{x}(t) \in \mathbb{R}^m$ is said to be (asymptotically) stationary [36] if for any $t \geq 0$,
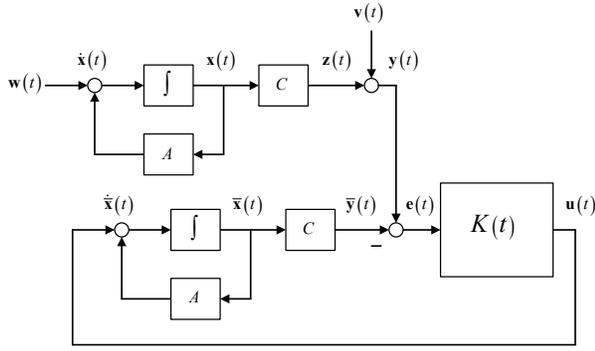
Fig. 1. The Kalman–Bucy filter.

the (asymptotic) correlation matrix

$$R_{\mathbf{x}}(\tau) = \lim_{t \to \infty} \mathrm{E}\left[\mathbf{x}(t+\tau)\mathbf{x}^T(t)\right]$$

exists. The (asymptotic) power spectrum of $\{\mathbf{x}(t)\}$, $\mathbf{x}(t) \in \mathbb{R}^m$ is then defined as

$$\Phi_{\mathbf{x}}(\omega) = \int_{-\infty}^{\infty} R_{\mathbf{x}}(\tau)\,\mathrm{e}^{-\mathrm{j}\omega\tau}\mathrm{d}\tau.$$

It can be verified that $\Phi_{\mathbf{x}}(\omega)$ is positive semidefinite. In the scalar case, we denote $\Phi_{\mathbf{x}}(\omega)$ by $S_{\mathbf{x}}(\omega)$. For the inversion,

$$\begin{aligned} R_{\mathbf{x}}(\tau) &= \lim_{t \to \infty} \mathrm{E}\left[\mathbf{x}(t+\tau)\mathbf{x}^T(t)\right] \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_{\mathbf{x}}(\omega)\,\mathrm{e}^{\mathrm{j}\omega\tau}\mathrm{d}\omega. \end{aligned}$$

Moreover, the (asymptotic) covariance matrix of $\{\mathbf{x}(t)\}$, $\mathbf{x}(t) \in \mathbb{R}^m$ is given by

$$\Sigma_{\mathbf{x}} = \lim_{t \to \infty} \mathrm{E}\left[\mathbf{x}(t)\mathbf{x}^T(t)\right] = R_{\mathbf{x}}(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_{\mathbf{x}}(\omega)\,\mathrm{d}\omega.$$

In the scalar case, $\Sigma_{\mathbf{x}}$ reduces to (asymptotic) variance $\sigma_{\mathbf{x}}^2$. Moreover, if $\{\mathbf{x}(t)\}$, $\mathbf{x}(t) \in \mathbb{R}^m$ is (asymptotically) white, then $\Phi_{\mathbf{x}}(\omega)$ is a constant X for all ω, and

$$R_{\mathbf{x}}(\tau) = \lim_{t \to \infty} \mathrm{E}\left[\mathbf{x}(t+\tau)\mathbf{x}^T(t)\right] = X\delta(\tau),$$

where $\delta(\cdot)$ is the Dirac delta function.

B. The Kalman–Bucy Filter

Consider the continuous-time Kalman–Bucy filtering system [4] depicted in Fig. 1, where the system is linear time-invariant (LTI) with state-space model given by

$$\begin{cases} \dot{\mathbf{x}}(t) &= A\mathbf{x}(t) + \mathbf{w}(t), \\ \mathbf{y}(t) &= C\mathbf{x}(t) + \mathbf{v}(t), \end{cases}$$

where $\mathbf{x}(t) \in \mathbb{R}^m$ is the state to be estimated, $\mathbf{y}(t) \in \mathbb{R}^l$ is the system output, $\mathbf{w}(t) \in \mathbb{R}^m$ is the process noise, and $\mathbf{v}(t) \in \mathbb{R}^l$ is the measurement noise. The system matrices are $A \in \mathbb{R}^{m \times m}$ and $C \in \mathbb{R}^{l \times m}$, and (A, C) is assumed to be detectable. Suppose that $\{\mathbf{w}(t)\}$ and $\{\mathbf{v}(t)\}$ are white Gaussian with covariance matrices

$$\mathrm{E}\left[\mathbf{w}(t+\tau)\mathbf{w}^T(t)\right] = W\delta(\tau),\ W \geq 0,$$

and

$$\mathrm{E}\left[\mathbf{v}(t+\tau)\mathbf{v}^T(t)\right] = V\delta(\tau),\ V > 0,$$

respectively, and that the initial state $\mathbf{x}(0)$ is Gaussian with covariance $\Sigma_{\mathbf{x}(0)}$ satisfying $0 < \det\Sigma_{\mathbf{x}(0)} < \infty$. Furthermore, $\{\mathbf{w}(t)\}$, $\{\mathbf{v}(t)\}$, and $\mathbf{x}(0)$ are assumed to be mutually uncorrelated.

The Kalman–Bucy filter is given by

$$\begin{cases} \dot{\overline{\mathbf{x}}}(t) &= A\overline{\mathbf{x}}(t) + \mathbf{u}(t), \\ \overline{\mathbf{y}}(t) &= C\overline{\mathbf{x}}(t), \\ \mathbf{e}(t) &= \mathbf{y}(t) - \overline{\mathbf{y}}(t), \\ \mathbf{u}(t) &= K(t)\mathbf{e}(t), \end{cases}$$

where $\overline{\mathbf{x}}(t) \in \mathbb{R}^m$, $\overline{\mathbf{y}}(t) \in \mathbb{R}^l$, $\mathbf{e}(t) \in \mathbb{R}^l$, and $\mathbf{u}(t) \in \mathbb{R}^m$. Herein, $K(t)$ denotes the Kalman gain, which is given by

$$K(t) = P(t)C^T V^{-1}, \tag{1}$$

where $P(t)$ denotes the state estimation error covariance matrix as

$$P(t) = \mathrm{E}\left[\left[\mathbf{x}(t) - \overline{\mathbf{x}}(t)\right]\left[\mathbf{x}(t) - \overline{\mathbf{x}}(t)\right]^T\right].$$

Herein, $P(t)$ is obtained using the Riccati equation

$$\dot{P}(t) = AP(t) + P(t)A^T + W - P(t)C^T V^{-1}CP(t),$$

with $P(0) = \mathrm{E}\left[\mathbf{x}(0)\mathbf{x}^T(0)\right]$.

It is known that the Kalman–Bucy filtering system converges, i.e., the estimator is asymptotically stable, the state estimation error $\{\mathbf{x}(t) - \overline{\mathbf{x}}(t)\}$ and the output estimation error $\{\mathbf{e}(t)\}$ are asymptotically stationary, and $\{\mathbf{e}(t)\}$ is asymptotically white, under the assumption that the system (A, C) is detectable. Moreover, in steady state, the optimal state estimation error covariance matrix

$$P = \lim_{t \to \infty} \mathrm{E}\left[\left[\mathbf{x}(t) - \overline{\mathbf{x}}(t)\right]\left[\mathbf{x}(t) - \overline{\mathbf{x}}(t)\right]^T\right]$$

attained by the Kalman–Bucy filter satisfies the algebraic Riccati equation

$$AP + PA^T + W - PC^T V^{-1}CP = 0, \tag{2}$$

whereas the steady-state Kalman gain is given by

$$K = PC^T V^{-1}. \tag{3}$$

In addition, the optimal steady-state output estimation error covariance matrix is found to be

$$\begin{aligned} \Sigma_{\mathbf{z}-\overline{\mathbf{y}}} &= \lim_{t \to \infty} \mathrm{E}\left[\left[\mathbf{z}(t) - \overline{\mathbf{y}}(t)\right]\left[\mathbf{z}(t) - \overline{\mathbf{y}}(t)\right]^T\right] \\ &= CPC^T, \end{aligned} \tag{4}$$

where $\mathbf{z}(t) = C\mathbf{x}(t)$ denotes the true value of the system output (in comparison, $\mathbf{y}(t)$ denotes the measured value of the system output). Moreover, when $l = 1$, (4) reduces to

$$\sigma_{\mathbf{z}-\overline{\mathbf{y}}}^2 = \lim_{t \to \infty} \mathrm{E}\left[\left[\mathbf{z}(t) - \overline{\mathbf{y}}(t)\right]^2\right].$$

It is also worth mentioning that

$$\lim_{t \to \infty} \mathrm{E}\left[\mathbf{e}(t+\tau)\mathbf{e}^T(t)\right] = V\delta(\tau).$$

III. A Counterpart of Jensen's Formula

In this section, we first examine the Poisson–Jensen formula (see, e.g., Appendix C of [16]) for a special class of analytic functions, and obtain a formula that will be essential for the rest of the paper.

Proposition 1: Let

$$f(s) = \frac{p(s)}{q(s)} \tag{5}$$

be a rational transfer function for which the numerator polynomial $p(s)$ and denominator polynomial $q(s)$ are both of order m. Suppose that all the poles of $f(s)$ are stable and that $f(s)$ has no zeros on the imaginary axis. In addition, suppose that

$$\lim_{s \to \infty} f(s) = 1. \tag{6}$$

Then,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \ln |f(j\omega)| \, d\omega$$
$$= \lim_{s \to \infty} \frac{1}{2} s \ln |f(s)| + \sum_{i=1}^{m} \max \{0, \Re[\varphi_i]\}, \tag{7}$$

where φ_i denote the zeros of $f(s)$.

The proof of Proposition 1 is omitted due to lack of space. Note that in the discrete-time case (for the unit disk), Jensen's formula is a consequence of the more general Poisson–Jensen formula [16]. Likewise in the continuous-time case (for the half plane), (7) is also a consequence of the more general Poisson–Jensen formula mandating that for every point $s_0 = \sigma_0 + j\omega_0$, $\sigma_0 > 0$,

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sigma_0}{\sigma_0^2 + (\omega - \omega_0)^2} \ln |f(j\omega)| \, d\omega$$
$$= \ln |f(s_0)| - \sum_i \ln \left| \frac{s_0 - z_i}{s_0 + z_i^*} \right|, \tag{8}$$

where z_i denote the nonminimum-phase zeros of $f(s)$ and z_i^* denote the complex conjugate of z_i. As such, (7) as well as the subsequent (11) may be viewed as the continuous-time (half-plane) counterparts of the discrete-time (unit-disk) Jensen's formula.

More generally, when not necessarily all the poles of $f(s)$ are stable, we can obtain the following result based on Proposition 1. The proof of Proposition 2 is also omitted due to lack of space.

Proposition 2: Let

$$f(s) = \frac{p(s)}{q(s)} \tag{9}$$

be a rational transfer function for which the numerator polynomial $p(s)$ and denominator polynomial $q(s)$ are both of order m. Suppose that $f(s)$ has no zeros on the imaginary axis and that

$$\lim_{s \to \infty} f(s) = 1. \tag{10}$$

Then,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \ln |f(j\omega)| \, d\omega$$
$$= \lim_{s \to \infty} \frac{1}{2} s \ln |f(s)| + \sum_i \max \{0, \Re[\varphi_i]\}$$
$$- \sum_j \max \{0, \Re[\eta_j]\}, \tag{11}$$

where φ_i denote the zeros of $f(s)$ and η_j denote its poles.

IV. Optimal Error Covariance by the Kalman–Bucy Filter

We now study the optimal output estimation error covariance matrix $\Sigma_{\mathbf{z} - \overline{\mathbf{y}}}$ by the continuous-time Kalman–Bucy filter, which is determined by the plant dynamics and the noise statistics in an integral characterization. Due to lack of space, the proof of Theorem 1 is omitted.

Theorem 1: The optimal output estimation error covariance matrix $\Sigma_{\mathbf{z} - \overline{\mathbf{y}}}$ by the Kalman–Bucy filter satisfies

$$\text{tr} \left(\Sigma_{\mathbf{z} - \overline{\mathbf{y}}} V^{-1} \right) = \text{tr} \left(CPC^T V^{-1} \right)$$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \ln \left[\frac{\det \Phi_{\mathbf{y}}(\omega)}{\det V} \right] d\omega$$
$$+ 2 \sum_{i=1}^{m} \max \{0, \Re[\lambda_i(A)]\}, \tag{12}$$

where

$$\Phi_{\mathbf{y}}(\omega) = C(j\omega I - A)^{-1} W(-j\omega I - A)^{-T} C^T + V, \tag{13}$$

and $\lambda_i(A)$ denote the eigenvalues of A.

On the right-hand side of (12), the term

$$\sum_{i=1}^{m} \max \{0, \Re[\lambda_i(A)]\} \tag{14}$$

quantifies the instability of the system [37], [38]. Moreover, when A is stable, i.e., when all the eigenvalues $\lambda_i(A)$ of matrix A satisfy $\Re[\lambda_i(A)] < 0$, $\{\mathbf{y}(t)\}$ is stationary and thus

$$C(j\omega I - A)^{-1} W(-j\omega I - A)^{-T} C^T + V \tag{15}$$

denotes the power spectrum $\Phi_{\mathbf{y}}(\omega)$ of $\{\mathbf{y}(t)\}$; in general, however, A needs not to be stable, and hence the function

$$C(sI - A)^{-1} W(-sI - A)^{-T} C^T + V \tag{16}$$

is a Popov function (see, e.g., [4], for a detailed discussion). One key step in the proof of Theorem 1 (see [34]) is

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \ln \left| \det \left[I + C(j\omega I - A)^{-1} K \right]^{-1} \right| d\omega$$
$$= \lim_{s \to \infty} \frac{1}{2} s \left\{ \ln \left| \det \left[I + C(sI - A)^{-1} K \right]^{-1} \right| \right\}$$
$$+ \sum_{i=1}^{m} \max \{0, \Re[\varphi_i]\}$$
$$= -\frac{1}{2} \text{tr} \left(CPC^T V^{-1} \right) + \sum_{i=1}^{m} \max \{0, \Re[\lambda_i(A)]\}, \tag{17}$$

where φ_i denote the zeros of $\det \left[I + C \left(sI - A\right)^{-1} K\right]^{-1}$. Note also that an alternative proof of (17) is given in [39].

A. Interpretation from the Viewpoint of Bode Integral

As a matter of fact, equation (17), in particular,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \ln \left| \det \left[I + C \left(j\omega I - A\right)^{-1} K\right]^{-1} \right| d\omega$$
$$= -\frac{1}{2} \mathrm{tr} \left(CPC^T V^{-1}\right) + \sum_{i=1}^{m} \max \left\{0, \Re \left[\lambda_i \left(A\right)\right]\right\}, \tag{18}$$

can also be obtained as a consequence of the (continuous-time) Bode integral [16]; it is interesting to discover that such a relation exists between these two top equations [40], namely algebraic Riccati equation and Bode integral, of the control field. Specifically, the estimator may be viewed as a feedback system, and thus the Bode integral can be obtained for its sensitivity from $\{y(t)\}$ to $\{e(t)\}$ (see Fig. 1). Indeed, if we denote

$$L(s) = C \left(sI - A\right)^{-1} K, \tag{19}$$

then (18) can be rewritten as

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \ln \left| \det \left[I + L \left(j\omega\right)\right]^{-1} \right| d\omega$$
$$= -\lim_{s \to \infty} \frac{1}{2} \mathrm{str} L(s) + \sum_{i=1}^{m} \max \left\{0, \Re \left[\lambda_i\right]\right\}, \tag{20}$$

where λ_i denote the poles of $L(s)$. Herein, we have used the fact that

$$\sum_{i=1}^{m} \max \left\{0, \Re \left[\lambda_i\right]\right\} = \sum_{i=1}^{m} \max \left\{0, \Re \left[\lambda_i \left(A\right)\right]\right\}, \tag{21}$$

which holds since the system is detectable (all the unstable modes of the system are observable) and thus the unstable poles of $L(s)$ correspond to the eigenvalues of A with real parts larger than zero. It is also worth mentioning that when it is further assumed that $l = 1$, (20) reduces to

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \ln \left| \frac{1}{1 + L \left(j\omega\right)} \right| d\omega$$
$$= -\lim_{s \to \infty} \frac{1}{2} s L(s) + \sum_{i=1}^{m} \max \left\{0, \Re \left[\lambda_i\right]\right\}. \tag{22}$$

On the other hand, it can be verified that the Bode integrals given in (20) and (22) are implied by (7) by letting $f(s) = \det \left[I + L(s)\right]^{-1}$ and $f(s) = 1/\left[1 + L(s)\right]$, respectively. In addition, note that one common instance for

$$\lim_{s \to \infty} s \left\{\ln \left| \det \left[I + L(s)\right]^{-1} \right| \right\} = 0 \tag{23}$$

to hold is that all the entries of $L(s)$ have at least two more poles than zeros, which is not satisfied in the case of Kalman–Bucy filtering as discussed in this paper.

B. A More Explicit Expression of (12)

Using (11), we can obtain a more explicit form for the formula in (12).

Theorem 2: Equality (12) can be rewritten as

$$\mathrm{tr} \left(\Sigma_{\mathbf{z}-\overline{\mathbf{y}}} V^{-1}\right) = \mathrm{tr} \left(CPC^T V^{-1}\right)$$
$$= \sum_{j} \max \left\{0, \Re \left[\varphi_j\right]\right\} - \sum_{k} \max \left\{0, \Re \left[\eta_k\right]\right\}$$
$$+ 2 \sum_{i=1}^{m} \max \left\{0, \Re \left[\lambda_i \left(A\right)\right]\right\}, \tag{24}$$

where φ_j denote the zeros of $\det \left[\Phi_{\mathbf{y}} \left(s\right) V^{-1}\right]$ and η_k denote its poles.

Proof: By invoking (11), it holds that

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \ln \left[\frac{\det \Phi_{\mathbf{y}} \left(\omega\right)}{\det V}\right] d\omega$$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \ln \det \left[\Phi_{\mathbf{y}} \left(\omega\right) V^{-1}\right] d\omega$$
$$= \lim_{s \to \infty} \frac{1}{2} s \left\{\ln \det \left[\Phi_{\mathbf{y}} \left(s\right) V^{-1}\right]\right\} + \sum_{j} \max \left\{0, \Re \left[\varphi_j\right]\right\}$$
$$- \sum_{k} \max \left\{0, \Re \left[\eta_k\right]\right\}.$$

Since

$$\Phi_{\mathbf{y}} \left(s\right) V^{-1} = C \left(sI - A\right)^{-1} W \left(-sI - A\right)^{-T} C^T V^{-1} + I,$$

it follows from [39] (Appendix A) that

$$\lim_{s \to \infty} \frac{1}{2} s \left\{\ln \det \left[\Phi_{\mathbf{y}} \left(s\right) V^{-1}\right]\right\} = 0,$$

Noting also (12), (24) follows by simple substitution. ∎

Concerning the terms in (24), it holds in general that

$$\sum_{k} \max \left\{0, \Re \left[\eta_k\right]\right\} \leq 2 \sum_{i=1}^{m} \max \left\{0, \Re \left[\lambda_i \left(A\right)\right]\right\}, \tag{25}$$

since the unstable poles of $\det \left[\Phi_{\mathbf{y}} \left(s\right) V^{-1}\right]$ must belong to the set of eigenvalues A with real parts larger than zero, while not all such eigenvalues of A may be present in the set of unstable poles of $\det \left[\Phi_{\mathbf{y}} \left(s\right) V^{-1}\right]$. Note that this is different from the case of $\det \left[I + C \left(sI - A\right)^{-1} K\right]$, as discussed in the proof of Theorem 1 (see [34]); in such a case, since the system (A, C) is detectable (all the unstable modes of the system are observable), the set of unstable poles of $\det \left[I + C \left(sI - A\right)^{-1} K\right]$ is exactly the same as the set of eigenvalues of A with real parts larger than zero.

C. Some Straightforward Corollaries and Relation to Existing Results

Using the fact that

$$\underline{\lambda} \left(C^T V^{-1} C\right) \mathrm{tr} P \leq \mathrm{tr} \left(CPC^T V^{-1}\right)$$
$$= \mathrm{tr} \left(V^{-\frac{1}{2}} CPC^T V^{-\frac{1}{2}}\right)$$
$$\leq \overline{\lambda} \left(C^T V^{-1} C\right) \mathrm{tr} P, \tag{26}$$

we could obtain lower and upper bounds on $\text{tr}P$ based on (12) as

$$\text{tr}P \geq \frac{1}{\overline{\lambda}\left(C^T V^{-1} C\right)} \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} \ln\left[\frac{\det \Phi_{\mathbf{y}}(\omega)}{\det V}\right] d\omega + 2\sum_{i=1}^{m} \max\left\{0, \Re\left[\lambda_i(A)\right]\right\} \right\}, \quad (27)$$

and

$$\text{tr}P \leq \frac{1}{\underline{\lambda}\left(C^T V^{-1} C\right)} \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} \ln\left[\frac{\det \Phi_{\mathbf{y}}(\omega)}{\det V}\right] d\omega + 2\sum_{i=1}^{m} \max\left\{0, \Re\left[\lambda_i(A)\right]\right\} \right\}. \quad (28)$$

One might also compare our bounds with those in, e.g., [5], which, however, goes beyond the scope of this paper.

We next consider some special cases of Theorem 1.

Corollary 1: If $l = 1$, and suppose that $\{\mathbf{w}(t)\}$ and $\{\mathbf{v}(t)\}$ are white Gaussian with covariance matrix W and variance $\sigma_{\mathbf{v}}^2$, respectively, then

$$\sigma_{\mathbf{z}-\overline{\mathbf{y}}}^2 = CPC^T$$
$$= \frac{\sigma_{\mathbf{v}}^2}{2\pi} \int_{-\infty}^{\infty} \ln\left[\frac{S_{\mathbf{y}}(\omega)}{\sigma_{\mathbf{v}}^2}\right] d\omega + 2\sigma_{\mathbf{v}}^2 \sum_{i=1}^{m} \max\left\{0, \Re\left[\lambda_i(A)\right]\right\}, \quad (29)$$

where

$$S_{\mathbf{y}}(\omega) = C\left(j\omega I - A\right)^{-1} W\left(-j\omega I - A\right)^{-T} C^T + \sigma_{\mathbf{v}}^2. \quad (30)$$

Corollary 2: If $l = m = 1$, and suppose that $\{\mathbf{w}(t)\}$ and $\{\mathbf{v}(t)\}$ are white Gaussian with variances $\sigma_{\mathbf{w}}^2$ and $\sigma_{\mathbf{v}}^2$, respectively, then

$$P = \frac{1}{C^2} \left\{ \frac{\sigma_{\mathbf{v}}^2}{2\pi} \int_{-\infty}^{\infty} \ln\left[C^2 \left|\frac{1}{j\omega - A}\right|^2 \frac{\sigma_{\mathbf{w}}^2}{\sigma_{\mathbf{v}}^2} + 1\right] d\omega + 2\sigma_{\mathbf{v}}^2 \max\left\{0, \Re\left[A\right]\right\} \right\}. \quad (31)$$

We now present a list of comparisons with existing results in the related works. It will be seen that our result in (12) is the most generic and reduces to all the listed ones, although they have adopted different approaches.

- When $l = 1$ and A is stable, $\{\mathbf{z}(t)\}$ is stationary and

$$C\left(j\omega I - A\right)^{-1} W\left(-j\omega I - A\right)^{-T} C^T \quad (32)$$

denotes its power spectrum $S_{\mathbf{z}}(\omega)$. Hence, (29) reduces to

$$\sigma_{\mathbf{z}-\overline{\mathbf{y}}}^2 = CPC^T = \frac{\sigma_{\mathbf{v}}^2}{2\pi} \int_{-\infty}^{\infty} \ln\left[\frac{S_{\mathbf{z}}(\omega) + \sigma_{\mathbf{v}}^2}{\sigma_{\mathbf{v}}^2}\right] d\omega. \quad (33)$$

This coincides with the result in [29].

- When $l = 1$, $\sigma_{\mathbf{v}}^2 = 1$, and $W = BB^T$, where $B \in \mathbb{R}^{m \times 1}$, (29) reduces to

$$\sigma_{\mathbf{z}-\overline{\mathbf{y}}}^2 = CPC^T$$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \ln\left[\left|C\left(j\omega I - A\right)^{-1} B\right|^2 + 1\right] d\omega + 2\sum_{i=1}^{m} \max\left\{0, \Re\left[\lambda_i(A)\right]\right\}. \quad (34)$$

This coincides with the conclusion in [30].

- When $l = m = 1$, $\sigma_{\mathbf{v}}^2 = 1$, and $\sigma_{\mathbf{w}}^2 = \varepsilon^2$, (29) reduces to

$$\sigma_{\mathbf{z}-\overline{\mathbf{y}}}^2 = CPC^T$$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \ln\left[\varepsilon^2 C^2 \left|\frac{1}{j\omega - A}\right|^2 + 1\right] d\omega + 2\sum_{i=1}^{m} \max\left\{0, \Re\left[A\right]\right\}. \quad (35)$$

This coincides with the corresponding result in [31].

- When $l = 1$, $\sigma_{\mathbf{v}}^2 = 1$, and $W = 0$, (29) reduces to

$$\sigma_{\mathbf{z}-\overline{\mathbf{y}}}^2 = CPC^T = 2\sum_{i=1}^{m} \max\left\{0, \Re\left[\lambda_i(A)\right]\right\}. \quad (36)$$

This also coincides with conclusion in [31].

- When $V = I$, (12) reduces to

$$\text{tr}\Sigma_{\mathbf{z}-\overline{\mathbf{y}}} = \text{tr}\left(CPC^T\right)$$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \ln\det \Phi_{\mathbf{y}}(\omega) d\omega + 2\sum_{i=1}^{m} \max\left\{0, \Re\left[\lambda_i(A)\right]\right\}, \quad (37)$$

where

$$\Phi_{\mathbf{y}}(\omega) = C\left(j\omega I - A\right)^{-1} W\left(-j\omega I - A\right)^{-T} C^T + I. \quad (38)$$

This coincides with the result in [33].

Discussions on the dual problem in control can be found in, e.g., [30], [32].

Indeed, equation (33) is referred to as the Yovits–Jackson formula in [41]. In a broad sense, the Yovits–Jackson formula can be viewed as the continuous-time counterpart of the Kolmogorov–Szegö formula [4], [36], [42]–[44] in the discrete-time case. For the multiple-input multiple-output (MIMO) case, when $\{\mathbf{y}(t)\}$ is stationary (A is stable), (12) reduces to

$$\text{tr}\left(\Sigma_{\mathbf{z}-\overline{\mathbf{y}}} V^{-1}\right) = \text{tr}\left(CPC^T V^{-1}\right)$$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \ln\left[\frac{\det \Phi_{\mathbf{y}}(\omega)}{\det V}\right] d\omega$$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \ln\left\{\frac{\det\left[\Phi_{\mathbf{z}}(\omega) + V\right]}{\det V}\right\} d\omega, \quad (39)$$

which may be viewed as the as the continuous-time counterpart of the Wiener–Masani formula [42], [44], [45].

As we pointed out in [35], our formula therein generalizes the Kolmogorov–Szegö formula and the Wiener–Masani formula to the non-stationary case, and correspondingly in this paper, we have generalized the Yovits–Jackson formula to the MIMO case as well as the non-stationary case.

V. Conclusion

In this paper, we have shown that the trace of the division of the optimal output estimation error covariance over the noise covariance attained by the Kalman–Bucy filter can be explicitly expressed in terms of the plant dynamics and noise statistics in a frequency-domain integral characterization. We have also discussed the relation of our integral characterization to the Bode integral. Possible future research directions include analysis of other classes of algebraic Riccati equations.

References

[1] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, vol. 83, no. 1, pp. 95–108, 1961.

[2] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Prentice-Hall, 1979.

[3] S. Bittanti, A. J. Laub, and J. C. Willems, *The Riccati Equation*. Springer, 1991.

[4] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall, 2000.

[5] W. H. Kwon, Y. S. Moon, and S. C. Ahn, "Bounds in algebraic Riccati and Lyapunov equations: A survey and some new results," *International Journal of Control*, vol. 64, no. 3, pp. 377–389, 1996.

[6] D. Vaughan, "A negative exponential solution for the matrix Riccati equation," *IEEE Transactions on Automatic Control*, vol. 14, no. 1, pp. 72–75, 1969.

[7] J. L. Crassidis and J. L. Junkins, *Optimal Estimation of Dynamic Systems*. CRC Press, 2011.

[8] H. W. Bode, *Network Analysis and Feedback Amplifier Design*. D.Van Nostrand, 1945.

[9] B. Francis and G. Zames, "On $\mathcal{H}_\infty$-optimal sensitivity theory for SISO feedback systems," *IEEE Transactions on Automatic Control*, vol. 29, no. 1, pp. 9–16, 1984.

[10] J. S. Freudenberg and D. P. Looze, "Right half plane poles and zeros and design tradeoffs in feedback systems," *IEEE Transactions on Automatic Control*, vol. 30, no. 6, pp. 555–565, 1985.

[11] S. Boyd and C. A. Desoer, "Subharmonic functions and performance bounds on linear time-invariant feedback systems," *IMA Journal of Mathematical Control and Information*, vol. 2, no. 2, pp. 153–170, 1985.

[12] R. H. Middleton, "Trade-offs in linear control system design," *Automatica*, vol. 27, no. 2, pp. 281–292, 1991.

[13] J. Chen, "Sensitivity integral relations and design trade-offs in linear multivariable feedback systems," *IEEE Transactions on Automatic Control*, vol. 40, no. 10, pp. 1700–1716, 1995.

[14] ——, "Logarithmic integrals, interpolation bounds, and performance limitations in MIMO feedback systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 6, pp. 1098–1115, 2000.

[15] G. Stein, "Respect the unstable," *IEEE Control Systems Magazine*, vol. 23, no. 8, pp. 12–25, 2003.

[16] G. C. Goodwin, S. F. Graebe, and M. E. Salgado, *Control System Design*. Prentice Hall, 2001.

[17] M. M. Seron, J. H. Braslavsky, and G. C. Goodwin, *Fundamental Limitations in Filtering and Control*. Springer, 1997.

[18] J. Chen, "Fundamental limitation of feedback control," in *Encyclopedia of Systems and Control*. London: Springer, 2014.

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

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

[21] N. C. Martins, M. A. Dahleh, and J. C. Doyle, "Fundamental limitations of disturbance attenuation in the presence of side information," *IEEE Transactions on Automatic Control*, vol. 52, no. 1, pp. 56–66, 2007.

[22] N. C. Martins and M. A. Dahleh, "Feedback control in the presence of noisy channels: "Bode-like" fundamental limitations of performance," *IEEE Transactions on Automatic Control*, vol. 53, no. 7, pp. 1604–1615, 2008.

[23] K. Okano, S. Hara, and H. Ishii, "Characterization of a complementary sensitivity property in feedback control: An information theoretic approach," *Automatica*, vol. 45, no. 2, pp. 504–509, 2009.

[24] H. Ishii, K. Okano, and S. Hara, "Achievable sensitivity bounds for MIMO control systems via an information theoretic approach," *Systems & Control Letters*, vol. 60, no. 2, pp. 111–118, 2011.

[25] S. Fang, H. Ishii, and J. Chen, "Tradeoffs in networked feedback systems: From information-theoretic measures to Bode-type integrals," *IEEE Transactions on Automatic Control*, vol. 62, no. 3, pp. 1046–1061, 2017.

[26] S. Fang, J. Chen, and H. Ishii, "Design constraints and limits of networked feedback in disturbance attenuation: An information-theoretic analysis," *Automatica*, vol. 79, pp. 65–77, 2017.

[27] ——, *Towards Integrating Control and Information Theories: From Information-Theoretic Measures to Control Performance Limitations*. Springer, 2017.

[28] ——, "Fundamental error bounds in state estimation: An information-theoretic analysis," in *Proceedings of the IEEE Conference on Decision and Control*, 2017, pp. 357–362.

[29] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. John Wiley & Sons, 1968.

[30] B. D. O. Anderson and D. L. Mingori, "Use of frequency dependence in linear quadratic control problems to frequency-shape robustness," *Journal of Guidance, Control, and Dynamics*, vol. 8, no. 3, pp. 397–401, 1985.

[31] J. H. Braslavsky, M. M. Seron, D. Q. Mayne, and P. V. Kokotovic, "Limiting performance of optimal linear filters," *Automatica*, vol. 35, no. 2, pp. 189–199, 1999.

[32] J. Freudenberg, J. Braslavsky, and R. Middleton, "Control over signal-to-noise ratio constrained channels: Stabilization and performance," in *Proceedings of the IEEE Conference on Decision and Control and European Control Conference*, 2005, pp. 191–196.

[33] H. Zhang, "On MMSE, mutual information and Popov function in continuous-time multivariable linear systems," in *Proceedings of the IEEE International Conference on Control Applications*, 2012, pp. 370–375.

[34] S. Fang, H. Ishii, J. Chen, and K. H. Johansson, "A frequency-domain characterization of optimal error covariance for the Kalman–Bucy filter," arXiv:1807.08604.

[35] S. Fang, H. Ishii, and J. Chen, "An integral characterization of optimal error covariance by Kalman filtering," in *Proceedings of the American Control Conference*, 2018, pp. 5031–5036.

[36] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 2002.

[37] L. Qiu, "Quantify the unstable," in *Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems*, 2010, pp. 981–986.

[38] W. Chen and L. Qiu, "Stabilization of networked control systems with multirate sampling," *Automatica*, vol. 49, no. 6, pp. 1528–1537, 2013.

[39] D. Mustafa and K. Glover, *Minimum Entropy H_∞ Control*. Springer, 1990.

[40] P. H. Jonathan, "Top equations in the field," *IEEE Control Systems Magzine*, vol. 35, no. 2, pp. 6–8, 2015.

[41] T. Kailath, "A view of three decades of linear filtering theory," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 146–181, 1974.

[42] A. Lindquist and G. Picci, *Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation and Identification*. Springer, 2015.

[43] P. P. Vaidyanathan, *The Theory of Linear Prediction*. Morgan & Claypool Publishers, 2007.

[44] Y. Chen, J. Karlsson, and T. T. Georgiou, "The role of the time-arrow in mean-square estimation of stochastic processes," *IEEE Control Systems Letters*, vol. 2, no. 1, pp. 85–90.

[45] N. Wiener and P. Masani, "The prediction theory of multivariate stochastic processes," *Acta Mathematica*, vol. 98, no. 1-4, pp. 111–150, 1957.