2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Bode-Like Integral for Stochastic Switched Systems in the Presence of Limited Information

Dapeng Li and Naira Hovakimyan

Abstract— **In this paper, we establish a Bode sensitivity integral formula for a class of feedback closed-loop systems with stochastic switched plants and controllers. Using information theory, we study the information conservation law, based on which a log integral theorem is obtained for the closed loops of interest. Furthermore we develop several algebraic conditions to explicitly capture the performance limitations. Application of this theoretical framework to Networked Control Systems (NCS) is used as an illustrative example.**

I. Introduction

Recent results on fundamental limitations of feedback control in the presence communication channels presented a fairly general and complete approach, in discrete-time setting, towards unification of information theory and control theory, [1], [2]. Entropy rate inequalities corresponding to the information flux in a typical causal closed-loop configuration were derived towards obtaining a Bode-like integral formula. Prior to this, extensions of Bode's theorem have been claimed for certain discrete-time nonlinear systems and linear time-varying systems respectively, [3], [4].

In this paper, we extend the framework from [2] to closed loops with stochastic switched plants. While switched control systems have been studied from various perspectives [5], it is still not clear how to characterize their fundamental limitations within an appropriate framework. We address the problem here by using an information theoretic framework towards obtaining a Bode integral formula, under the assumptions that the switching sequence is an ergodic Markov chain. We first derive a closed-loop information conservation law by using information theoretic arguments similar to [6] and [2]. Then under some stationarity assumption, a Bode integral-like theorem is obtained, characterizing a lower bound on the performance limitations. To enable the simplified calculation of the resulting lower bound, some Lie algebraic conditions are developed.

To demonstrate the usefulness of the theoretical result, we propose an application with NCS with random packet dropouts, which has been widely used in control literature to model typical computer network protocols, such as TCP and UDP [7]. We develop a Bode integral to show that the degree of instability of the plants determines the lower bound of the performance limitation.

The paper is organized as follows. In Section II we introduce the closed-loop feedback configuration and some basic definitions and facts from information theory and the theory of stochastic processes. Section III studies a general feedback scheme, within which we develop a mutual information inequality and a Bode-type integral formula. Section IV applies Bode's integral to NCS. The paper is concluded in Section V.

Notation:

- $\mathbb{R}$ denotes the field of real numbers; $\mathbb{C}$ stands for complex plane; $\mathbb{C}^-$ and $\mathbb{C}^+$ stand for the left half and right half of $\mathbb{C}$ respectively.
- Random variables defined in appropriate probability spaces are represented using boldface letters, such as $\mathbf{x}$, $\mathbf{y}$. If not otherwise stated, the random variables take values in $\mathbb{R}$ throughout the paper.
- If $\mathbf{x}(k), k \in \mathbb{N}^+$, is a discrete time stochastic process, we denote its segment $\{\mathbf{x}(k)\}_{k=l}^{u}$ by $\mathbf{x}_l^u$, and use $\mathbf{x}_0^n :=$ $\mathbf{x}^n$ for simplicity.
- $\mathbf{E}[\cdot]$ is the expectation operator of a random variable.
- $(\cdot)^+ = \max\{\cdot, 0\}$ and $(\cdot)^- = \min\{\cdot, 0\}$.
- $\Re(\cdot)$ gives the real part of a complex number.
- $\lambda_j(\cdot)$ gives the eigenvalues of a square matrix.
- $h(\cdot)$ stands for (differential) entropy and $I(\cdot;\cdot|\cdot)$ for conditioned mutual information; $\bar{h}$ and $\bar{I}$ stand for the entropy rate and mutual information rate respectively.
- When A is a finite set, $|A|$ gives the number of elements in A.
- $sp\{\cdot\}$ denotes the spectrum of an operator.

II. Preliminaries & Problem Formulation

A list of useful properties of entropy and mutual information are given here, and are frequently used in the upcoming arguments.

(P1) *Symmetry and nonnegativity*:
$I(\mathbf{x};\mathbf{y}) = I(\mathbf{y};\mathbf{x}) = h(\mathbf{x}) - h(\mathbf{x}|\mathbf{y}) = h(\mathbf{y}) - h(\mathbf{y}|\mathbf{x}) \geq 0\,.$

(P2) *Kolmogorov equality*:

$$I(\mathbf{x};(\mathbf{y},\mathbf{z})) = I(\mathbf{x};\mathbf{z}) + I(\mathbf{x};\mathbf{y}|\mathbf{z})$$

(P3) *Data processing inequality*:

$$I(\mathbf{x};\mathbf{y}|\mathbf{z}) \geq I(\mathbf{x};g(\mathbf{y})|\mathbf{z})$$

The equality holds, if $g(\cdot)$ is invertible.

(P4) *Invariance of mutual information (entropy)*

$$I(\mathbf{x};\mathbf{y}|\mathbf{z}) = I(\mathbf{x}+g(\mathbf{z});\mathbf{y}|\mathbf{z})\,, h(\mathbf{x}|\mathbf{z}) = h(\mathbf{x}+g(\mathbf{z})|\mathbf{z}),$$

where $g(\cdot)$ is a function.

(P5) *Chain rule*:

$$h(\mathbf{x}^n|\mathbf{y}) = \sum_{k=1}^{n} h(\mathbf{x}_k|\mathbf{y},\mathbf{x}^{k-1})$$

978-1-4577-0079-8/11/$26.00 ©2011 AACC

(P6) *Maximum entropy*: Consider $\mathbf{x} \in \mathbb{R}^m$ and the covariance matrix given by $V := \mathbf{E}[\mathbf{x}\mathbf{x}^\top]$. Then we have

$$h(\mathbf{x}) \leq h(\bar{\mathbf{x}}) = \frac{1}{2}\log((2\pi e)^m \det V),$$

where $\bar{\mathbf{x}}$ is a Gaussian process with the same covariance as $\mathbf{x}$. Equality holds, if $\mathbf{x}$ is Gaussian.

Throughout the paper we consider the feedback configuration depicted in Fig. 1.

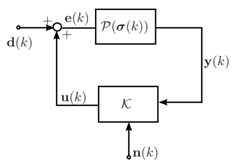

Fig. 1. Basic Feedback Scheme

Several assumptions are made:

- The plant $\mathcal{P}$ is modeled by the following stochastic difference equation

$$\begin{aligned} \mathbf{x}(k+1) &= A(\boldsymbol{\sigma}(k))\mathbf{x}(k) + B(\boldsymbol{\sigma}(k))\mathbf{e}(k), \quad \mathbf{x}(0) = \mathbf{x}_0, \\ \mathbf{y}(k) &= C(\boldsymbol{\sigma}(k))\mathbf{x}(k), \quad k = 0, 1, 2.... \end{aligned} \tag{1}$$

  Here $\mathbf{x}(k) \in \mathbb{R}^m$, and $\mathbf{x}_0$ is assumed to have finite differential entropy or $h(\mathbf{x}_0) < \infty$, and $\boldsymbol{\sigma}(k) \in \{1, 2, ..., N\} =: \mathcal{N}$ is a finite state ergodic Markov process given by

$$P(\boldsymbol{\sigma}(k+1) = j | \sigma(k) = i) := p_{ij} \geq 0,$$

  where p_{ij} is named as transition probability from state i to j, and $\sum_j p_{ij} = 1$ for all $i \in \mathcal{N}$. The stationary distribution of the Markov chain $\boldsymbol{\sigma}$, denoted as $\boldsymbol{\pi} = [\pi_1, ..., \pi_{|\mathcal{N}|}]$, is obtained by solving

$$\boldsymbol{\pi}^\top [p_{ij}]_{i,j\in\mathcal{N}} = \boldsymbol{\pi}^\top, \text{ and } \quad [1, .., 1]\boldsymbol{\pi} = 1.$$

  We also assume $A(n), n \in \mathcal{N}$ is not singular.
- The disturbance $\mathbf{d}(k)$ is a stochastic process, and $\mathbf{n}(k)$ is a stochastic process that models the controller noise. We assume that $\boldsymbol{\sigma}(k)$, $\mathbf{d}(k)$, $\mathbf{n}(k)$ and $\mathbf{x}_0$ are mutually independent.
- The controller $\mathcal{K}$ is given as a deterministic causal map such that

$$\mathcal{K} : (k, \mathbf{y}^{k-1}, \mathbf{n}^k) \mapsto \mathbf{u}(k).$$

Definition 2.1 (Wide Sense Stationary Process): A zero-mean stochastic process $\mathbf{x}(k) \in \mathbb{R}^n$, $t \geq 0$, is stationary, if for all $k \geq 0$ its covariance function, defined by

$$R_{\mathbf{x}}(l) = \mathbf{E}[\mathbf{x}(k+l)\mathbf{x}^\top(k)], \quad l \in \mathbb{N}^+,$$

is independent of l. Throughout this paper, *wide sense stationary* is abbreviated as *stationary* for convenience.

Definition 2.2: The spectral density of a stationary process $\mathbf{v}$ is given as the following Fourier transform

$$f_{\mathbf{v}}(\omega) = \frac{1}{2\pi}\sum_{k=0}^{\infty} \mathbf{v}(k)e^{-j\omega k}$$

Definition 2.3 (Sensitivity-like Function): A sensitivity-like function of the closed loop is defined as

$$S_{\mathbf{d},\mathbf{e}}(\omega) = \sqrt{\frac{f_{\mathbf{e}}(\omega)}{f_{\mathbf{d}}(\omega)}},$$

where $\mathbf{e}$ and $\mathbf{d}$ are stationary and stationarily correlated.

Remark 2.4: The function $S_{\mathbf{d},\mathbf{e}}(\omega)$ is the stochastic analogue of the sensitivity function $|S(j\omega)|$ in Bode's original work [8].

Throughout, we adopt the following stability definition.

Definition 2.5 (Mean-square Stability): The closed loop given in Fig. 1 is said to be mean-square stable, if

$$\sup_{k\geq 0} \mathbf{E}[\mathbf{x}^\top(k)\mathbf{x}(k)] < \infty. \tag{2}$$

Definition 2.6 (Lie Algebra): A Lie algebra is denoted as

$$\mathfrak{g} := \{A(n) : n \in \mathcal{N}\}_{LA},$$

which is generated by the matrices $A(n), n \in \mathcal{N}$, with respect to the standard Lie bracket

$$[A(1), A(2)] := A(1)A(2) - A(2)A(1).$$

We say that the Lie algebra $\mathfrak{g}$ is *solvable* if the following derived series

$$\mathfrak{g} > [\mathfrak{g}, \mathfrak{g}] > [[\mathfrak{g}, \mathfrak{g}], [\mathfrak{g}, \mathfrak{g}]] > ...$$

becomes 0 in finite operations, where $[\cdot, \cdot]$ denotes the algebra generated by Lie bracket and "$>$" denotes the relation of sub-algebra

Theorem 2.7: [Simultaneous triangularization] The matrices $\{A(n) : n \in \mathcal{N}\}$ can be simultaneously triangularized by some linear operator $T \in \mathbb{C}^{m\times m}$, if and only if the Lie algebra $\mathfrak{g}$ is *solvable*.

III. Bode-like Integral Discrete Time Case

In this section we develop the information conservation law of the closed loop system depicted in Fig. 1. In turn, an analogue of Bode's formula is obtained with stationarity assumption.

A. Information conservation law

The following lemma is introduced to characterize the closed loop causality.

Lemma 3.1:

$$I(\mathbf{d}(i); (\mathbf{u}^i, \mathbf{x}_0, \boldsymbol{\sigma}^i)|\mathbf{d}^{i-1}) = 0, \quad \forall\, i \geq 1. \tag{3}$$

Proof:

$$\begin{aligned} & I(\mathbf{d}(i); (\mathbf{u}^i, \mathbf{x}_0, \boldsymbol{\sigma}^i)|\mathbf{d}^{i-1}) \\ & \overset{(a)}{\leq} I(\mathbf{d}(i); (\mathbf{u}^i, \mathbf{n}^i, \boldsymbol{\sigma}^i, \mathbf{x}_0)|\mathbf{d}^{i-1}) \\ & \overset{(b)}{\leq} I(\mathbf{d}(i); (\mathbf{d}^{i-1}, \mathbf{n}^i, \boldsymbol{\sigma}^i, \mathbf{x}_0)|\mathbf{d}^{i-1}) \\ & \overset{(c)}{=} I(\mathbf{d}(i); (\mathbf{n}^i, \boldsymbol{\sigma}^i, \mathbf{x}_0)|\mathbf{d}^{i-1}) \\ & \overset{(d)}{=} 0 \end{aligned}$$

Here, (a) follows from (P3); (b) also follows from (P3), since $\mathbf{u}^i$ is a function of $(\mathbf{d}^{i-1}, \mathbf{n}^i, \boldsymbol{\sigma}^i, \mathbf{x}_0)$; (c) follows from (P4), and (d) is implied because $\mathbf{n}$, $\boldsymbol{\sigma}$, $\mathbf{x}_0$ and $\mathbf{d}$ are mutually independent. ■

In what follows we use the result from Lemma 3.1 to achieve an equality, revealing a key relationship among signals residing in 1.

Lemma 3.2: Consider the closed loop in Fig. 1. The following inequality holds

$$h(\mathbf{e}^k) = h(\mathbf{d}^k)+I((\mathbf{x}_0,\boldsymbol{\sigma}^k);\mathbf{e}^k)+\sum_{i=1}^{k} I(\mathbf{u}^i;\mathbf{e}(i)|\mathbf{e}^{i-1},\mathbf{x}_0,\boldsymbol{\sigma}^k) \tag{4}$$

Proof: We break down the equality (3) by

$$\begin{aligned}
0 &= I(\mathbf{d}(i);(\mathbf{u}^i,\mathbf{x}_0,\boldsymbol{\sigma}^i)|\mathbf{d}^{i-1})\\
&\overset{(a)}{=} I(\mathbf{d}(i);\mathbf{u}^i,\mathbf{x}_0,\boldsymbol{\sigma}^i,\mathbf{d}^{i-1}) - I(\mathbf{d}(i);\mathbf{d}^{i-1})\\
&\overset{(b)}{=} I(\mathbf{d}(i);\mathbf{u}^i,\mathbf{x}_0,\boldsymbol{\sigma}^i,\mathbf{e}^{i-1}) - I(\mathbf{d}(i);\mathbf{d}^{i-1})\\
&\overset{(c)}{=} -h(\mathbf{d}(i)|\mathbf{u}^i,\mathbf{x}_0,\boldsymbol{\sigma}^i,\mathbf{e}^{i-1}) + h(\mathbf{d}(i)|\mathbf{d}^{i-1})\\
&\overset{(d)}{=} -h(\mathbf{e}(i)|\mathbf{u}^i,\mathbf{x}_0,\boldsymbol{\sigma}^i,\mathbf{e}^{i-1}) + h(\mathbf{d}(i)|\mathbf{d}^{i-1})\\
&\overset{(e)}{=} -h(\mathbf{e}(i)|\mathbf{e}^{i-1}) + I((\mathbf{x}_0,\boldsymbol{\sigma}^i));\mathbf{e}(i)|\mathbf{e}^{i-1})+\\
&\qquad I(\mathbf{u}^i;\mathbf{e}(i)|\mathbf{e}^{i-1},\mathbf{x}_0,\boldsymbol{\sigma}^i) + h(\mathbf{d}(i)|\mathbf{d}^{i-1})\,.
\end{aligned}$$

Here (a) follows from (P3), (b) follows from the fact that $\mathbf{e}^{i-1} = \mathbf{u}^{i-1}+\mathbf{d}^{i-1}$, (c) follows from (P1), (d) follows from (P4) and (f) from (P5). Summing up the above equality from 1 to k and using (P5), we have (4). ■

Remark 3.3: The term $\sum_{i=1}^{k} I(\mathbf{u}^i;\mathbf{e}(i)|\mathbf{e}^{i-1},\mathbf{x}_0,\boldsymbol{\sigma}^k)$ is alternatively represented as the directed information from $\mathbf{u}$ to $\mathbf{e}$ conditioned by $(\mathbf{x}_0,\boldsymbol{\sigma}^k)$ [9].

Theorem 3.4: Consider the closed loop shown in Fig. 1. The following entropy rate inequality holds

$$\bar{h}(\mathbf{e}) \geq \bar{h}(\mathbf{d}) + \bar{I}((\mathbf{x}_0,\boldsymbol{\sigma});\mathbf{e})\,. \tag{5}$$

Proof: Considering the nonegativeness of the mutual information, from (4) we have

$$h(\mathbf{e}^k) \geq h(\mathbf{d}^k) + I((\mathbf{x}_0,\boldsymbol{\sigma}^k);\mathbf{e}^k)\,.$$

The proof is completed by dividing both sides of the above equality by k and letting $k \to \infty$. ■

Remark 3.5: The inequality in (5) has been derived in both information theory and control theory literature in different setups and with different generalities. Here we only assume causality of the closed loop.

B. Evaluating an important information rate

As it can be seen in (5), the mutual information rate $\bar{I}((\mathbf{x}_0,\boldsymbol{\sigma});\mathbf{e})$ plays an important role in the conservation law. In this subsection we establish some nontrivial lower bounds for $\bar{I}((\mathbf{x}_0,\boldsymbol{\sigma});\mathbf{e})$ assuming some algebraic conditions.

Theorem 3.6: Consider the closed loop in Fig. 1. The following inequality holds.

$$\bar{I}((\mathbf{x}_0,\boldsymbol{\sigma});\mathbf{e}) \geq \liminf_{k\to\infty} \frac{1}{k}\mathbf{E}\sum_j \Re\left(\log\lambda_j\left(F_k\right)\right)^+\,, \tag{6}$$

where $F_k := A(\boldsymbol{\sigma}(k))A(\boldsymbol{\sigma}(k-1))\cdots A(\boldsymbol{\sigma}(0))$.

Proof: We first consider the dynamics of the plant

$$\mathbf{x}(k+1) = \mathbf{x}(k)A(\boldsymbol{\sigma}(k)) + B(\boldsymbol{\sigma}(k))\mathbf{e}(k)\,,$$

which can be solved as

$$\begin{aligned}
\mathbf{x}(k+1) &= \left(\prod_{i=0}^{k} A(\boldsymbol{\sigma}(i))\right)\mathbf{x}_0+\\
&\quad\sum_{i=0}^{k}\left(\prod_{l=i}^{k} A(\boldsymbol{\sigma}(l))\right)B(\boldsymbol{\sigma}(i))\mathbf{e}(i)\\
&= F_k(\mathbf{x}_0 - \hat{\mathbf{x}}_0(k+1)),
\end{aligned}$$

where

$$\hat{\mathbf{x}}_0(k+1) := -\left(\prod_{i=0}^{k} A(\boldsymbol{\sigma}(i))\right)^{-1}\sum_{i=0}^{k}\left(\prod_{l=i}^{k} A(\boldsymbol{\sigma}(l))\right)B(\boldsymbol{\sigma}(i))\mathbf{e}(i)\,.$$

F_k can be decomposed into the following form by a linear transformation T_k:

$$T_k^{-1}F_kT_k = \begin{bmatrix} F_{ku} & 0\\ 0 & F_{ks}\end{bmatrix},$$

where F_{ku} is unstable and F_{ks} is stable. The same linear transformation can be applied to $\mathbf{x}_0$ and $\hat{\mathbf{x}}_0$ to have

$$T_k\mathbf{x}_0 = \begin{bmatrix}\mathbf{x}_{u0}\\ \mathbf{x}_{s0}\end{bmatrix} \quad\text{and } T_k\hat{\mathbf{x}}_0 = \begin{bmatrix}\hat{\mathbf{x}}_{u0}\\ \hat{\mathbf{x}}_{s0}\end{bmatrix}.$$

We note that $\mathbf{x}_{u0}$ and $\mathbf{x}_{s0}$ are functions of k, however this argument is omitted for notational simplicity.

We then establish the lower bound of $I(\mathbf{x}_0,\boldsymbol{\sigma}^k;\mathbf{e}^k)$ as follows

$$\begin{aligned}
&I(\mathbf{x}_0,\boldsymbol{\sigma}^k;\mathbf{e}^k)\\
&= I(\mathbf{x}_0;\boldsymbol{\sigma}^k) + I(\mathbf{x}_0;\mathbf{e}^k|\boldsymbol{\sigma}^k)\\
&= I(\mathbf{x}_0;\mathbf{e}^k|\boldsymbol{\sigma}^k)\\
&= I(\mathbf{x}_0;\mathbf{e}^k,\boldsymbol{\sigma}^k)\\
&= I(\mathbf{x}_{u0},\mathbf{x}_{s0};\mathbf{e}^k,\boldsymbol{\sigma}^k)\\
&= h(\mathbf{x}_{u0},\mathbf{x}_{s0}) - h(\mathbf{x}_{s0}|\mathbf{e}^k,\boldsymbol{\sigma}^k) - h(\mathbf{x}_{u0}|\mathbf{x}_{s0},\mathbf{e}^k,\boldsymbol{\sigma}^k)\\
&\geq h(\mathbf{x}_{u0},\mathbf{x}_{s0}) - h(\mathbf{x}_{s0}) - h(\mathbf{x}_{u0}|\mathbf{x}_{s0},\mathbf{e}^k,\boldsymbol{\sigma}^k)\,,
\end{aligned}$$

where we have used the assumption that $\mathbf{x}_0$ $\boldsymbol{\sigma}$ are independent. To evaluate the term $h(\mathbf{x}_{u0}|\mathbf{x}_{s0},\mathbf{e}^k,\boldsymbol{\sigma}^k)$, we note that

$$\begin{aligned}
&h(\mathbf{x}_{u0}|\mathbf{x}_{s0},\mathbf{e}^k,\boldsymbol{\sigma}^k)\\
&= h(\mathbf{x}_{u0}-\hat{\mathbf{x}}_{u0}|\mathbf{x}_{s0},\mathbf{e}^k,\boldsymbol{\sigma}^k)\\
&\leq h(\mathbf{x}_{u0}-\hat{\mathbf{x}}_{u0})\\
&\leq \log(2\pi e)^{l_k} - \log\mathbf{E}|\det F_{ku}| + \log\mathbf{E}\det\mathbf{x}_{u0}(k)\mathbf{x}_{u0}^\top(k)\\
&\leq \log(2\pi e)^{l_k} - \mathbf{E}\log|\det F_{ku}| + \log\mathbf{E}\det\mathbf{x}_{u0}(k)\mathbf{x}_{u0}^\top(k),
\end{aligned}$$

where $0 \leq l_k \leq m$ is the dimension of $\mathbf{x}_{u0}$, $\mathbf{x}_{u0}(k)$ is the vector formed by the first l_k elements of $T_k\mathbf{x}(k)$ and the last inequality follows form Jensen's inequality.

Therefore we have

$$I(\mathbf{x}_0,\boldsymbol{\sigma}^k;\mathbf{e}^k) \geq -\log(2\pi e)^{l_k} + \mathbf{E}\log|\det F_{ku}| - \log\mathbf{E}\det\mathbf{x}_u(k)\mathbf{x}_u^\top(k)\,.$$

Note that the stability of the closed loop system implies that $\mathbf{E}\det \mathbf{x}_u(k)\mathbf{x}_u^\top(k) < \infty\,, \forall k$. Then we have

$$\bar{I}(\mathbf{x}_0, \boldsymbol{\sigma}^k; \mathbf{e}^k) \geq \liminf_{k\to\infty} \frac{1}{k}\mathbf{E}\log|\det F_{ku}| = \liminf_{k\to\infty} \frac{1}{k}\mathbf{E}\sum_j \Re\left(\log\lambda_j\left(F_k\right)\right)^+ .$$

■

Remark 3.7: The right hand side of (6) is actually a Lyapunov exponent for the dynamic system (1). For a complete treatment of Lyapunov exponents for stochastic switching systems, one is referred to [10].

To overcome the difficulty of obtaining $\liminf_{k\to\infty}\frac{1}{k}\mathbf{E}\sum_j \Re\left(\log\lambda_j\left(F_k\right)\right)^+$ by using Lyapunov exponent method, we exploit the algebraic structure of the matrices $A(n), n \in \mathcal{N}$. From Theorem 2.7 we know that the solvability of $\mathfrak{g}$ implies that $\{A(n)\}\,, n \in \mathcal{N}$, can be simultaneously triangularizable by some linear transformation $T \in \mathbb{C}^{m\times m}$:

$$T^{-1}A(n)T = \begin{bmatrix} \lambda_1^{(n)} & \star & \star \\ 0 & \ddots & \star \\ 0 & 0 & \lambda_m^{(n)} \end{bmatrix}, \forall n \in \mathcal{N}\,. \quad (7)$$

Now we divide the index set $\{1, ..., m\}$ into two distinct sets $\mathcal{M}_u$ and $\mathcal{M}_s$, defined by

$$\mathcal{M}_u := \left\{ j : \prod_{n\in\mathcal{N}} |\lambda_j^{(n)}|^{\pi_n} > 1\,, j = 1, 2, ..., m \right\},$$

$$\mathcal{M}_s := \{1, ..., m\} \setminus \mathcal{M}_u\,.$$

Corollary 3.8: Suppose that the Lie algebra $\mathfrak{g}$ is *solvable*. Then we have

$$\bar{I}((\mathbf{x}_0, \boldsymbol{\sigma}); \mathbf{e}) \geq \sum_{n\in\mathcal{N}}\sum_{j\in\mathcal{M}_u} \pi_n \log|\lambda_j^{(n)}|$$

Proof: We start with a mutually disjoint partition of the index set $\{1, 2, ..., \boldsymbol{\sigma}(k)\}$, given by

$$\{1, 2, ..., \boldsymbol{\sigma}(k)\} = \bigcup_{n\in\mathcal{N}} \mathcal{K}_n\,,$$

where $\mathcal{K}_n := \{i : \boldsymbol{\sigma}(i) = n, i = 1, 2, .., k\}$. Then we claim that the eigenvalues of F_k take the form $\lambda_j(F_k) = \prod_{n\in\mathcal{N}}\prod_{j=1}^m \left(\lambda_j^{(n)}\right)^{|\mathcal{K}_n|}$, where $\lambda_j^{(n)}$ is the diagonal entry from (7). Indeed it is easy to see that $T^{-1}F_kT = T^{-1}A(\boldsymbol{\sigma}(k))TT^{-1}A(\boldsymbol{\sigma}(k-1))T\cdots T^{-1}A(\boldsymbol{\sigma}(0))T$ is a triangular matrix for all k. Further, the jth diagonal entry of $T^{-1}F_kT$ can be calculated as

$$\lambda_j(F_k) = \prod_{i=0}^k \lambda_j^{(\boldsymbol{\sigma}(i))} = \prod_{n\in\mathcal{N}}\left(\lambda_j^{(n)}\right)^{|\mathcal{K}_n|}$$

Using the above relation and Fatou's Lemma we have

$$\begin{aligned}
&\liminf_{k\to\infty}\frac{1}{k}\mathbf{E}\sum_j \Re\left(\log\lambda_j\left(F_k\right)\right)^+ \\
&= \liminf_{k\to\infty}\mathbf{E}\frac{1}{k}\sum_j \Re\left(\log\lambda_j(F_k)\right)^+ \\
&\geq \mathbf{E}\liminf_{k\to\infty}\frac{1}{k}\sum_j \Re\left(\log\lambda_j(F_k)\right)^+ .
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&\liminf_{k\to\infty}\frac{1}{k}\sum_j \Re\left(\log\lambda_j(F_k)\right)^+ \\
&= \lim_{k\to\infty}\frac{1}{k}\sum_j \Re\left(\log\lambda_j(F_k)\right)^+ \\
&= \sum_j \Re\left(\sum_n \pi_n \log\lambda_j^{(n)}\right)^+ \\
&= \sum_{n\in\mathcal{N}}\sum_{j\in\mathcal{M}_u} \pi_n\log|\lambda_j^{(n)}|\,,
\end{aligned}$$

where the second equality follows from ergodicity of $\boldsymbol{\sigma}(k)$.

■

Remark 3.9: As explained in [5], this modern algebraic approach, though mathematically appealing, shows a significant drawback for its lack of robustness, i.e. even a very small perturbation of system parameters can violate the solvability condition. One may conduct perturbation analysis to relax the algebraic structure requirement, though it is not trivial in general.

Here we propose yet another way to determine the Lyapunov exponent $\liminf_{k\to\infty}\frac{1}{k}\mathbf{E}\sum_i \Re\left(\log\lambda_i\left(F_k\right)\right)^+$ by using operator semigroup theory. To start with, we consider the semigroup generated by matrices $\{A(n)\,, n \in \mathcal{N}\}$ with respect to the matrix multiplication. The following lemma from [11] gives a sufficient condition for the permutability of the spectra of the product of the operators.

Theorem 3.10: If for all $n_1, n_2, n_3 \in \mathcal{N}$,

$$sp(A(n_1)A(n_2)A(n_3)) = sp(A(n_2)A(n_1)A(n_3))\,, \quad (8)$$

then for any sequence $A(n_1), ..., A(n_k)\,,\quad n_1, ..., n_k \in \mathcal{N}$, the following identity holds for any permutation τ of $\{n_1, ..., n_k\}$

$$sp\left\{\prod_i^k A(n_i)\right\} = sp\left\{\prod_{\tau(i)}^{\tau(k)} A(n_{\tau(i)})\right\}.$$

The following corollary is now straightforward to prove.

Corollary 3.11: Suppose that the condition in (8) is satisfied. Then we have

$$\bar{I}((\mathbf{x}_0, \boldsymbol{\sigma}); \mathbf{e}) \geq \sum_j \Re\left(\log\lambda_j\left(\prod_{n\in\mathcal{N}} A(n)^{\pi_n}\right)\right)^+ .$$

Proof:

Theorem 3.10 implies that

$$sp(F_k) = sp\left\{\prod_{n\in\mathcal{N}} A^{|\mathcal{K}_n|}(n)\right\} = sp\left\{\prod_{n\in\mathcal{N}} A^{|\mathcal{K}_{\tau(n)}|}(\tau(n))\right\}$$
$$= \{\hat{\lambda}_1^{(k)}, \ldots, \hat{\lambda}_m^{(k)}\}$$

for any permutation $\tau(\cdot)$, where $\hat{\lambda}_j^{(k)} = \prod_{n\in\mathcal{N}}\left(\lambda_j^{(k)}\right)^{|\mathcal{K}_n|}$. Following the same argument in the proof of Corollary 3.8, we have

$$\liminf_{k\to\infty}\frac{1}{k}\mathbf{E}\sum_j \Re\left(\log\lambda_j\left(F_k\right)\right)^+ \geq$$
$$\mathbf{E}\liminf_{k\to\infty}\frac{1}{k}\sum_j \Re\left(\log\hat{\lambda}_j^{(k)}\right)^+$$

and

$$\liminf_{k\to\infty}\frac{1}{k}\sum_j \Re\left(\log\hat{\lambda}_j^{(k)}\right)^+$$
$$= \lim_{k\to\infty}\frac{1}{k}\sum_j \Re\left(\log\hat{\lambda}_j^{(k)}\right)^+$$
$$= \sum_j \Re\left(\sum_n \pi_n \log\lambda_j^{(n)}\right)^+$$
$$= \sum_j \Re\left(\log\lambda_j\left(\prod_{n\in\mathcal{N}} A(n)^{\pi_n}\right)\right)^+ .$$

The theorem is proved. ■

C. Bode's Integral

Theorem 3.12: Consider the closed loop in Fig. 1. If $\mathbf{d}$ and $\mathbf{e}$ form Gaussian stationary processes, then

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left(S_{\mathbf{d},\mathbf{e}}(\omega)\right)d\omega \geq \liminf_{k\to\infty}\frac{1}{k}\sum_i \Re\left(\log\lambda_i\left(F_k\right)\right)^+ .$$

Proof: This result is evident by considering the following relation, followed by Kolmogrov's formula

$$\bar{h}(\mathbf{d}) = \log(2\pi e) + \frac{1}{4\pi}\int_{-\pi}^{\pi}\log f_{\mathbf{d}}(\omega)d\omega\,,$$
$$\bar{h}(\mathbf{e}) = \log(2\pi e) + \frac{1}{4\pi}\int_{-\pi}^{\pi}\log f_{\mathbf{e}}(\omega)d\omega\,,$$

together with Theorem 3.6. ■

Since we have obtained various lower bounds for $\bar{I}(\mathbf{x}_0, \mathbf{d}, \boldsymbol{\sigma}; \mathbf{e})$ in the previous subsection, the following corollaries can be readily obtained.

Corollary 3.13: Consider the closed loop in Fig. 1. If $\mathbf{d}$ and $\mathbf{e}$ form Gaussian stationary processes, then

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left(S_{\mathbf{d},\mathbf{e}}(\omega)\right)d\omega \geq \log\prod_{n\in\mathcal{N}}|\det A(n)|^{\pi_n} .$$

Corollary 3.14: Consider the closed loop in Fig. 1. If $\mathbf{d}$ and $\mathbf{e}$ form Gaussian stationary processes, and the Lie algebra $\mathfrak{g}$ is *solvable*, then

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left(S_{\mathbf{d},\mathbf{e}}(\omega)\right)d\omega \geq \sum_{n\in\mathcal{N}}\sum_{j\in\mathcal{M}_u}\pi_n\log|\lambda_j^{(n)}| .$$

Corollary 3.15: Consider the closed loop in Fig. 1. If $\mathbf{d}$ and $\mathbf{e}$ form Gaussian stationary processes, and the condition in (8) is satisfied, then

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left(S_{\mathbf{d},\mathbf{e}}(\omega)\right)d\omega \geq \sum_j \Re\left(\log\lambda_j\left(\prod_{n\in\mathcal{N}}A(n)^{\pi_n}\right)\right)^+$$

Remark 3.16: Similar to its deterministic counterpart, Bode's integral in this stochastic setting also captures the performance limitation of a closed loop in frequency domain. The lower bound of the achievable performance is inherent from its open loop plant instability.

Remark 3.17: Though it is hard to determine whether the closed loop in Fig. 1 is stationary in general, some results for LTI systems can be found in [12] and [13].

D. Data Rate Inequality

Another inequality, resulting from the closed loop causality, is developed here. The following lemma provides a lower bound for the mutual information rate $\bar{I}((\mathbf{x}_0, \mathbf{d}); \mathbf{u})$, which accounts for total information rate flow in the loop. Further insight into $\bar{I}((\mathbf{x}_0, \mathbf{d}); \mathbf{u})$ can be found in [2] and [14].

Lemma 3.18: Consider the closed-loop system shown in Fig. 1. We have the following inequality:

$$\bar{I}((\mathbf{x}_0, \mathbf{d}, \boldsymbol{\sigma}); \mathbf{u}) \geq \bar{I}(\mathbf{x}_0, \boldsymbol{\sigma}; \mathbf{e}) + \bar{I}(\mathbf{d}; \mathbf{u})\,.$$

Proof: Using Kolmogorov's formula (P2), we have

$$I((\mathbf{x}_0, \mathbf{d}^k, \boldsymbol{\sigma}^k); \mathbf{u}^k) = I(\mathbf{x}_0, \boldsymbol{\sigma}^k; \mathbf{u}^k|\mathbf{d}^k) + I(\mathbf{u}^k; \mathbf{d}^k)\,,$$

where $k \in \mathbb{N}^+$ is an arbitrary time instance. We can lower bound $I((\mathbf{x}_0, \mathbf{d}^k); \mathbf{u}^k)$ as

$$\begin{aligned}
&I((\mathbf{x}_0, \boldsymbol{\sigma}^k, \mathbf{d}^k); \mathbf{u}^k)\\
&\overset{(a)}{=} I(\mathbf{x}_0, \boldsymbol{\sigma}^k; \mathbf{e}^k|\mathbf{d}^k) + I(\mathbf{u}^k; \mathbf{d}^k)\\
&\overset{(b)}{=} I(\mathbf{x}_0, \boldsymbol{\sigma}^k; \mathbf{e}^k) - I(\mathbf{x}_0, \boldsymbol{\sigma}^k; \mathbf{d}^k) +\\
&\qquad I(\mathbf{x}_0, \boldsymbol{\sigma}^k; \mathbf{d}^k|\mathbf{e}^k) + I(\mathbf{u}^k; \mathbf{d}^k)\\
&\overset{(c)}{=} I(\mathbf{x}_0, \boldsymbol{\sigma}^k; \mathbf{e}^k) + I(\mathbf{x}_0, \boldsymbol{\sigma}^k; \mathbf{d}^k|\mathbf{e}^k) + I(\mathbf{u}^k; \mathbf{d}^k)\\
&\overset{(d)}{\geq} I(\mathbf{x}_0, \boldsymbol{\sigma}^k; \mathbf{e}^k) + I(\mathbf{u}^k; \mathbf{d}^k)\,.
\end{aligned}$$

Here (a) follows from the fact that $I(\mathbf{x}_0; \mathbf{u}^k|\mathbf{d}^k) = I(\mathbf{x}_0; \mathbf{u}^k + \mathbf{d}^k|\mathbf{d}^k) = I(\mathbf{x}_0; \mathbf{e}^k|\mathbf{d}^k)$; (b) follows from (P2); (c) follows from the independence of $\mathbf{d}$ and $\mathbf{x}_0$; and (d) follows from the fact that $I(\mathbf{x}_0, \boldsymbol{\sigma}^k; \mathbf{d}^k|\mathbf{e}^k) \geq 0$. We have obtained the following inequality:

$$I((\mathbf{x}_0, \mathbf{d}^k, \boldsymbol{\sigma}^k); \mathbf{u}^k) \geq I(\mathbf{x}_0, \boldsymbol{\sigma}^k; \mathbf{e}^k) + I(\mathbf{u}^k; \mathbf{d}^k)\,. \qquad (9)$$

The conclusion is readily obtained by dividing the terms on both sides of (9) by k and taking the limit as $k \to \infty$. ■

IV. Example: Networked Control Systems with Random Packet Dropouts

In this section, we apply the framework from the previous section to examine the performance limitation problems in the networked control systems (NCS). To be specific, we only consider the control systems with a lossy communication channel placed between the sensor and the controller, which has been studied in various papers [15] [16] [17]. In this paper we adopt a structure similar to [16], shown in Fig. 2, where a switch is placed after the output to model packet dropouts.

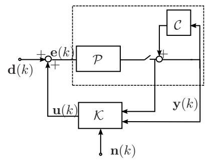

Fig. 2. A networked control system

The packet dropouts are compensated for by an output of an LTI system, which has to be designed. The controller can be represented by any causal map from $\mathbf{y}_0^k$ to $u(k)$. The sequence of *ON's* and *OFF's* of the erasure channel is modeled as a two-state Markov chain with transition probability matrix

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}, \qquad 0 \le p, q \le 1\,.$$

One can calculate the stationary distribution as $\boldsymbol{\pi} = \left[\frac{q}{p+q}, \frac{p}{p+q}\right]$. Let the state space realization of the plant and the channel compensator be

$$\left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] \text{ and } \left[\begin{array}{c|c} A_c & B_c \\ \hline C_c & 0 \end{array}\right]$$

respectively. We can then regard the dashed box in Fig. 2 as a generalized "plant" with state matrices $\tilde{A}(1) = \begin{bmatrix} A & 0 \\ B_cC & A_c \end{bmatrix}$, $\tilde{A}(2) = \begin{bmatrix} A & 0 \\ 0 & A_c + B_cC_c \end{bmatrix}$ for the "ON" and "OFF" of the erasure channel respectively.

To simplify the subsequent analysis, we further assume that the compensator is chosen such that A_c and $A_c + B_cC_c$ are stable. Under these additional conditions and with account of Theorem 3.6 we have the corresponding Bode's integral theorem.

Theorem 4.1: Consider the NCS in Fig. 2, and assume that the signal $\mathbf{u}$ is Gaussian and stationary. The following relation holds for all causal controllers $\mathcal{K}$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left(S_{\mathbf{d},\mathbf{e}}(\omega)\right) d\omega \ge \sum_j \Re\left(\log \lambda_j\left(A\right)\right)^+ \,. \qquad (10)$$

Proof: The proof is a simple application of Theorem 3.12, and is therefore omitted here. ■

Remark 4.2: This theorem characterizes the control design limitation for NCS with random packet dropout. Given the stable compensator, the right hand side in (10) shows that the lower bound of the closed loop performance is determined solely by the degree of instability of A. This observation suggests that, considering the relatively loose definition of stability in (2), packet dropout does not make the system "more" unstable. However, the dropout may add up to the performance limitation in other forms, for which a close scrutiny is required.

V. Conclusions

This paper has developed a relatively complete Bode's integral formula for stochastic switched closed loops. Information theory has been employed as machinery to obtain a relationship among different system variables, which has in turn resulted in Bode's integral for stationary cases. Various algebraic conditions have been proposed to capture tight performance bounds. An example of applying this theoretic framework to the field of NCS illustrates the usefulness of this fundamental result.

References

[1] N. Martins, M. Dahleh, and J. 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.

[2] N. Martins and M. Dahleh, "Feedback control in the presence of noisy channels:Bode-like fundamental limitations of performance," *IEEE Transactions on Automatic Control*, vol. 52, no. 7, pp. 1604–1615, 2008.

[3] P. Iglesias, "An analogue of bode's integral for stable nonlinear systems: relations to entropy," in *Decision and Control, 2001. Proceedings of the 40th IEEE Conference on*, vol. 4, 2001, pp. 3419–3420.

[4] ——, "Tradeoffs in linear time-varying systems: an analogue of Bode's sensitivity integral," *Automatica*, vol. 37, no. 10, pp. 1541–1550, 2001.

[5] D. Liberzon, *Switching in systems and control*. Springer, 2003.

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

[7] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 163–187, 2007.

[8] H. Bode, "Network analysis and feedback amplifier design," *Princeton, NJ*, 1945.

[9] J. Massey, "Causality, feedback and directed information," in *IEEE 1990 International Symposium on Information Theory and its Applications*, 1990, pp. 27–30.

[10] X. Feng, "Lyapunov exponents and stability of linear stochastic systems," Ph.D. dissertation, Case Western Reserve University, 1990.

[11] H. Radjavi and P. Rosenthal, *Simultaneous triangularization*. Springer Verlag, 2000.

[12] O. Costa and M. Fragoso, "Stability results for discrete-time linear systems with Markovian jumping parameters," *Journal of mathematical analysis and applications*, vol. 179, no. 1, pp. 154–178, 1993.

[13] C. Francq and J. Zakoḷan, "Stationarity of multivariate Markov-switching ARMA models," *Journal of Econometrics*, vol. 102, no. 2, pp. 339–364, 2001.

[14] D. Li and N. Hovakimyan, "Bode-like integral for continuous-time closed-loop systems in the presence of limited information," *IEEE Transactions on Automatic Control (Submitted)*, 2010.

[15] V. Gupta, B. Hassibi, and R. Murray, "Optimal LQG control across packet-dropping links," *Systems & Control Letters*, vol. 56, no. 6, pp. 439–446, 2007.

[16] Q. Ling and M. Lemmon, "Power spectral analysis of networked control systems with data dropouts," *IEEE Transactions on Automatic control*, vol. 49, no. 6, pp. 955–959, 2004.

[17] N. Elia, "Remote stabilization over fading channels," *Systems & Control Letters*, vol. 54, no. 3, pp. 237–249, 2005.