Conditions for Tracking in Networked Control Systems

I. Lopez*
Northern New Mexico College
CET Department
921 Paseo de Oñate
Española, NM 87000
ilopez@nnmc.edu

C.T. Abdallah**, S.K. Jayaweera, H. Tanner
Electrical & Computer Engineering Department
MSC01 1100
1 University of New Mexico
Albuquerque, NM 87131-0001
{chaouki,jayaweera}@ece.unm.edu
tanner@unm.edu

***Abstract*— In this paper we obtain information theoretical conditions for tracking in linear time-invariant control systems. We consider the particular case where the closed loop contains a channel in the feedback loop. The mutual information rate between the feedback signal and the reference input signal is used to quantify information about the reference signal that is available for feedback. This mutual information rate must be maximized in order to improve the tracking performance. The mutual information is shown to be upper bounded by a quantity that depends on the unstable eigenvalues of the plant and on the channel capacity. If the channel capacity reaches a lower limit, the feedback signal becomes completely uncorrelated with the reference signal, rendering feedback useless. We also find a lower bound on the expected squared tracking error in terms of the entropy of a random reference signal. Examples and simulations are provided to demonstrate the results.**

I. Introduction

The goal of this work is to find fundamental limitations on feedback tracking systems in terms of information theoretical quantities. This is important since the emerging control applications involve the presence of a constraint communication channel in the feedback loop. Typically, control systems have been understood as signal processing blocks or systems interchanging energy. However, these approaches are not appropriate for the new scenarios. That is why we suggest that an interpretation in terms of information flow may be more suitable for the future design of control algorithms.

Previous related work in [3], [5], [6], [13], [14], [15] and [16] detailed some aspects of performance and limitations of control systems in terms of information theoretic quantities. Specifically, the work in [14] dealt with the tracking issues without a channel in the feedback link, while [3] dealt with disturbance rejection. A result in [14], shows that a necessary condition for efficient tracking is that the information flow from the reference signal to the output should be greater than the information flow between the disturbance and the output. We know that in the absence of noise, and without a communication channel in the feedback loop, the mutual information rate (or information rate) between reference signal and the output is infinite. We know, however, that if the feedback signal is transmitted by means of a finite capacity channel, the mutual information rate is upper bounded by $C_f - \sum_{i=1} \max\{0, \log_2(|\lambda_i(A)|)\}$, [12].

Following the same approach of [4], we expect that the parameters of the plant and feedback channel capacity C_f will be related, and that there will be a trade-off between these parameters. If by some reason this upper bound happens to be zero, then we reach a fundamental limitation where no information of the reference signal is available for feedback. This means that the two signals are independent, therefore, uncorrelated, and this is exactly the condition that implies that tracking is impossible. In other words, the feedback signal does not provide any useful information for the reference to be tracked.

We note that the condition for a non-zero mutual information between the reference and the feedback signal is a necessary condition for tracking, but not a sufficient one. A large mutual information between the reference signal and the feedback signal does not necessarily imply that tracking is possible (it only implies that the signals are highly correlated). This is expected because even in the case of a perfect infinite capacity channel, the tracking issue requires additional conditions to be satisfied.

These results are fundamental limitations in terms of information quantities that any control system designer must be aware of before trying to design a new control system.

II. Notation

We present next the notation used in the rest of this work.

- Let $\mathbf{x}^k = \{\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(k)\}$ and $\mathbf{y}^k = \{\mathbf{y}(1), \mathbf{y}(2), \ldots, \mathbf{y}(k)\}$ be sets of observations of stochastic processes $\mathbf{x}$ and $\mathbf{y}$. We follow the notation in [11] where bold letters represent stochastic processes.
- Let $\mathbf{x}(k)$ be a time sample of the stochastic process $\mathbf{x}$.
- Let $\mathbf{x_j}$ be the "j-th" state component. For example, if $\mathbf{x}$ has dimension $n = 3$, then $\mathbf{x_j}$ will denote any of the state components $\mathbf{x_1}$, $\mathbf{x_2}$ or $\mathbf{x_3}$.
- Let $\mathbf{x_J}$ denote the set of state components, $\mathbf{x_j}$, such that $\mathrm{j} \in \mathrm{J}$. For example, if $\mathrm{J} = \{1, 3\}$, then $\mathbf{x_J}$ is the set $\{\mathbf{x_1}, \mathbf{x_3}\}$.
- Let $|.|$ denote the absolute value and **det(.)** denotes the absolute value of the determinant of a matrix.

* The research of this author is partially supported by CONACYT.
** The research of this author is partially supported by NSF award CNS 0626380 under the FIND initiative.

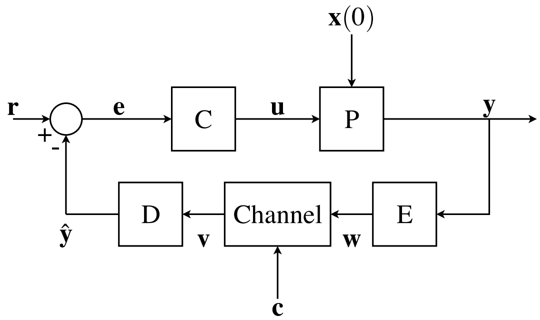

Fig. 1. Closed-Loop System with Communication Channel in Feedback Link.

We also define the blocks in Figure 1:

- C is the controller, which does not have any constraints (it could be time-invariant, nonlinear, etc.).
- P is the plant to be controlled and is assumed to be discrete, linear, time-invariant, with state-space realization

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + B\mathbf{u}(k); \quad (1)$$
$$\mathbf{y}(k) = \mathbb{C}\mathbf{x}(k). \quad (2)$$

- E is the encoder assumed to be a causal operator well defined in the input alphabet of the channel.
- D is the decoder assumed to be well defined and conserving equimemory with the encoder.
- The Channel block is any type of communication channel with finite capacity.
- c is the channel noise.

III. Information Theory Preliminaries

Before proceeding, we enumerate some well-known information-theoretical properties that will be very useful later on.

Properties *3.1:* Assume that $\mathbf{z}, \mathbf{w}, \mathbf{u} \in \mathbb{R}$ are random variables and $f(\mathbf{z})$, $g(\mathbf{z})$ are real functions. All of the following may be found in several references as [2], [4] and [9].

(a) $h(\mathbf{z}|\mathbf{w}) \leq \mathbf{h}(\mathbf{z})$ with equality if $\mathbf{z}$ and $\mathbf{w}$ are independent.

(b) Let $\mathbf{z}$ have mean μ and covariance $\mathrm{Cov}\{\mathbf{z}^n\}$. Then
$$h(\mathbf{z}^n) \leq \frac{1}{2}\log_2\left((2\pi e)^n \mathbf{det}(\mathrm{Cov}\{\mathbf{z}^n\})\right)$$
with equality if $\mathbf{z}$ has a multivariate normal distribution.

(c) $h(a\mathbf{z}) = h(\mathbf{z}) + \log_2\left(|a|\right)$ for nonzero constant a.

(d) $h(A\mathbf{z}) = h(\mathbf{z}) + \log_2\left(\mathbf{det}(A)\right)$ for nonsingular A matrix.

(e) $h(\mathbf{z}|\mathbf{w}) = h(\mathbf{z} - g(\mathbf{w})|\mathbf{w})$.

(f) $I(\mathbf{z};\mathbf{w}) = I(\mathbf{w};\mathbf{z}) \geq 0$.

(g) $I(\mathbf{z};\mathbf{w}) \geq I(g(\mathbf{z}); f(\mathbf{w}))$.

(h) $I(\mathbf{z};\mathbf{w}|\mathbf{u}) = I((\mathbf{u},\mathbf{z});\mathbf{w}) - I(\mathbf{u};\mathbf{w}) = h(\mathbf{z}|\mathbf{u}) - h(\mathbf{z}|\mathbf{w},\mathbf{u}) = h(\mathbf{w}|\mathbf{u}) - h(\mathbf{w}|\mathbf{z},\mathbf{u})$.

(i) For any random variable $\mathbf{z}$ and estimate $\hat{\mathbf{z}}$: $E\{(\mathbf{z} - \hat{\mathbf{z}})^2\} \geq \frac{1}{2\pi e}2^{2h(\mathbf{z})}$, with equality if and only if $\mathbf{z}$ is Gaussian and $\hat{\mathbf{z}}$ is the mean of $\mathbf{z}$.

(j) The variance of the error in the estimate $\hat{\mathbf{z}}$ of $\mathbf{z}$ given the infinite past is lower bounded as $\sigma_\infty^2(\mathbf{z}) = \lim_{k\to\infty} E\{(\mathbf{z}-\hat{\mathbf{z}})^2(k)|(\mathbf{z}-\hat{\mathbf{z}})(k-1)\} \geq \frac{1}{2\pi e}2^{2h_\infty(\mathbf{z})}$ with equality if $\mathbf{z}$ is Gaussian.

(k) If z is an asymptotically stationary process, then
$$h_\infty(z) \leq \frac{1}{4\pi}\int_{\pi}^{\pi}\log_2\left(2\pi e\widehat{\Phi}_z(\omega)\right)d\omega$$
where $\widehat{\Phi}_z$ is the asymptotic power spectral density of z and equality holds if, in addition, z is Gaussian autoregressive.

IV. Signal Analysis

The functional dependencies among the signals involved in the closed-loop shown in Figure 1 are the following:

$$\begin{aligned}
\mathbf{y}(k) &= f_1(\mathbf{r}^{k-1}, \mathbf{c}^{k-1}, \mathbf{x}(0)); \\
\mathbf{e}(k) &= f_2(\mathbf{r}^k, \hat{\mathbf{y}}^k) = \mathbf{r}(k) - \hat{\mathbf{y}}(k); \\
\mathbf{u}(k) &= f_3(\mathbf{e}^k); \\
\hat{\mathbf{y}}(k) &= f_4(\mathbf{y}^k, \mathbf{c}^k).
\end{aligned}$$

V. Assumptions

The matrix A in block P in Figure 1 is assumed to be diagonal with only unstable eigenvalues ($|\lambda_i(A)| > 1$) and therefore, A^k is invertible $\forall k$. We assume that A has unstable eigenvalues since it is the worst case. Since we are considering the tracking problem, the control law is a function of the error $\mathbf{e}^k = \mathbf{r}^k - \hat{\mathbf{y}}^k$, $\mathbf{u}(k) = f_3(\mathbf{e}^k)$. We note for now that the output is an n-dimensional vector, but this will be relaxed later on. In our setup f_3 is not limited to be a linear or time-invariant control law. We note that the solution of the difference equation (1) may be written as $\mathbf{x}(k) = A^k\mathbf{x}(0) + \sum_{i=0}^{k-1} A^{k-i-1}Bf_3(\mathbf{e}^i)$. If $\mathbb{C} = \mathbb{I}$, then from the tracking error, defined by $\boldsymbol{\varepsilon}(k) = \mathbf{r}(k) - \mathbf{y}(k)$, we have

$$\mathbf{r}(k) - \boldsymbol{\varepsilon}(k) = \mathbf{y}(k) = \mathbf{x}(k) = A^k\mathbf{x}(0) + \sum_{i=0}^{k-1} A^{k-i-1}Bf_3(\mathbf{e}^i). \quad (3)$$

We rearrange the terms as

$$\mathbf{x}(0) + A^{-k}\sum_{i=0}^{k-1} A^{k-i-1}Bf_3(\mathbf{e}^i) = -A^{-k}(\boldsymbol{\varepsilon}(k) - \mathbf{r}(k)). \quad (4)$$

In a tracking problem, we do not necessarily assume that the state is bounded, since for unbounded reference signals, the state may grow unbounded. Instead, we assume that the closed-loop is such that the error is bounded, i.e.,

$$E\{\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}\} < \infty.$$

Since this implies that $\boldsymbol{\varepsilon}$ is a second-order process, the mean $E\{\boldsymbol{\varepsilon}\}$ and the covariance $\mathrm{Cov}\{\boldsymbol{\varepsilon}\} = E\{(\boldsymbol{\varepsilon} + E\{\boldsymbol{\varepsilon}\})(\boldsymbol{\varepsilon} + E\{\boldsymbol{\varepsilon}\})^T\}$ must be finite. For bounded reference signals, the condition $E\{\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}\} < \infty$ guarantees stability since by the triangle inequality [8] we know that

$$\sqrt{E\{\mathbf{x}^2(k)\}} \leq \sqrt{E\{\mathbf{r}^2(k)\}} + \sqrt{E\{\boldsymbol{\varepsilon}^2(k)\}}. \quad (5)$$

Since the two terms on the right side of equation (5) are finite, then we also get that $\sqrt{E\{\mathbf{x}^2(k)\}} < \infty$ and, therefore, the system remains stable.

VI. Auxiliary Results

We first introduce some results that will later be used to obtain the limitations on tracking systems. Specifically, the following result will be used to prove Lemma 6.3. Let us consider the set P_j defined as $P_j = \{i \in \mathbb{N}, j \leq n : i \in \{1,2,\dots,n\} - \{j\}\}$. The following lemma holds for stabilization and is a slight modification of the result presented in [4].

***Lemma** 6.1:* Consider the closed-loop system in Figure 1, where the plant is a DTLI system described by equations (1) and (2), with $\mathbb{C} = \mathbb{I}$, and A diagonal in equation (2). If $\mathrm{E}\{\mathbf{x}_{P_j}(k)\mathbf{x}_{P_j}^T(k)\} < \infty$, then

$$\lim_{k\to\infty} \frac{I(\mathbf{x}_{P_j}(0);\mathbf{e}^k|\mathbf{r}^k,\mathbf{x}_j(0))}{k} \geq \sum_{i\neq j} \log_2\left(|\lambda_i(A)|\right).$$

Proof: By the chain rule expressed in Property 3.1.(h) we expand the expression given by $I(\mathbf{x}_{P_j}(0);\mathbf{e}^k|\mathbf{r}^k,\mathbf{x}_j)$ as

$$\begin{aligned} &I(\mathbf{x}_{P_j}(0);\mathbf{e}^k|\mathbf{r}^k,\mathbf{x}_j(0)) \\ &\quad= \sum_{i\neq j}^{n} I(\mathbf{x}_i(0);\mathbf{e}^k|\mathbf{r}^k,\mathbf{x}_j(0),\mathbf{x}_1(0),\dots,\mathbf{x}_{i-1}(0)). \end{aligned} \tag{6}$$

Each state component may be expressed as

$$\mathbf{x}_i(k) = \lambda_i^k \mathbf{x}_i(0) + g_i(\mathbf{e}^k); \tag{7}$$

for some function g_i. Therefore, each initial state component is given by $\mathbf{x}_i(0) = \lambda_i^{-k}(\mathbf{x}_i(k) - g_i(\mathbf{e}^k))$. From the definition of mutual information we expand the "i-th" additive term in equation (6).

$$\begin{aligned} &I(\mathbf{x}_i(0);\mathbf{e}^k|\mathbf{r}^k,\mathbf{x}_j(0),\mathbf{x}_1(0),\dots,\mathbf{x}_{i-1}(0)) \\ &\quad= h(\mathbf{x}_i(0)|\mathbf{r}^k,\mathbf{x}_j(0),\mathbf{x}_1(0),\dots,\mathbf{x}_{i-1}(0)) \\ &\qquad -h(\mathbf{x}_i(0)|\mathbf{e}^k,\mathbf{r}^k,\mathbf{x}_j(0),\mathbf{x}_1(0),\dots,\mathbf{x}_{i-1}(0)). \end{aligned}$$

From the independence between $\mathbf{x}(0)$ and $\mathbf{r}^k$, $\forall i \in P_j$, the term $\mathbf{r}^k$ may be eliminated in the first entropy term

$$\begin{aligned} &I(\mathbf{x}_i(0);\mathbf{e}^k|\mathbf{r}^k,\mathbf{x}_j(0),\mathbf{x}_1(0),\mathbf{x}_2(0),\dots,\mathbf{x}_{i-1}(0)) \\ &\quad= h(\mathbf{x}_i(0)|\mathbf{x}_j(0),\mathbf{x}_1(0),\mathbf{x}_2(0),\dots,\mathbf{x}_{i-1}(0)) \\ &\qquad - h(\mathbf{x}_i(0)|\mathbf{e}^k,\mathbf{r}^k,\mathbf{x}_j(0),\mathbf{x}_1(0),\mathbf{x}_2(0),\dots,\mathbf{x}_{i-1}(0)). \end{aligned}$$

From equation (7), the term

$$h(\mathbf{x}_i(0)|\mathbf{e}^k,\mathbf{r}^k,\mathbf{x}_1(0),\mathbf{x}_2(0),\dots,\mathbf{x}_{i-1}(0))$$

may be rewritten as

$$\begin{aligned} &h(\mathbf{x}_i(0)|\mathbf{e}^k,\mathbf{r}^k,\mathbf{x}_1(0),\mathbf{x}_2(0),\dots\mathbf{x}_{i-1}(0)) \\ &\quad= h(\lambda_i^{-k}(\mathbf{x}_i(k)-g_i(\mathbf{e}^k))|\mathbf{e}^k,\mathbf{r}^k,\mathbf{x}_1(0),\mathbf{x}_2(0),\dots\mathbf{x}_{i-1}(0)). \end{aligned}$$

By Properties 3.1.(c), 3.1.(b) and 3.1.(a) we have that

$$\begin{aligned} &h(\lambda_i^{-k}\mathbf{x}_i(k)|\mathbf{e}^k,\mathbf{r}^k,\mathbf{x}_1(0),\mathbf{x}_2(0),\dots\mathbf{x}_{i-1}(0)) \\ &\quad= h(\lambda_i^{-k}\mathbf{x}_i(k)|\mathbf{e}^k,\mathbf{r}^k,\mathbf{x}_1(0),\mathbf{x}_2(0),\dots\mathbf{x}_{i-1}(0)); \\ &\quad= -k\log_2(|\lambda_i|) + h(\mathbf{x}_i(k)|\mathbf{e}^k,\mathbf{r}^k,\mathbf{x}_1(0),\mathbf{x}_2(0),\dots\mathbf{x}_{i-1}(0)); \\ &\quad\leq -k\log_2(|\lambda_i|) + h(\mathbf{x}_i(k)); \\ &\quad\leq -k\log_2(|\lambda_i|) + \frac{1}{2}\log_2\Big((2\pi e)\mathbf{det}(\mathrm{Cov}\{\mathbf{x}_i\})\Big). \end{aligned}$$

Then

$$\begin{aligned} &I(\mathbf{x}_i(0);\mathbf{e}^k|\mathbf{r}^k,\mathbf{x}_j(0),\mathbf{x}_1(0),\mathbf{x}_2(0),\dots,\mathbf{x}_{i-1}) \\ &\quad\geq h(\mathbf{x}_i(0)|\mathbf{x}_j(0),\mathbf{x}_1(0),\mathbf{x}_2(0),\dots,\mathbf{x}_{i-1}(0)) \\ &\qquad +k\log_2(|\lambda_i|) - \frac{1}{2}\log_2\Big((2\pi e)\mathbf{det}(\mathrm{Cov}\{\mathbf{x}_i\})\Big). \end{aligned}$$

Dividing by k and taking the limit to infinity we obtain

$$\lim_{k\to\infty} \frac{I(\mathbf{x}_i(0);\mathbf{e}^k|\mathbf{r}^k,\mathbf{x}_j(0),\mathbf{x}_1(0),\dots,\mathbf{x}_{i-1}(0))}{k} \geq \log_2(|\lambda_i|). \tag{8}$$

From equations (6) and (8) we have

$$\lim_{k\to\infty} \frac{I(\mathbf{x}_{P_j}(0);\mathbf{e}^k|\mathbf{r}^k,\mathbf{x}_j(0))}{k} \geq \sum_{i\neq j} \log_2\left(|\lambda_i(A)|\right).$$

■

We next focus on the tracking problem which is different from the stabilization one treated in previous works. We first consider the following two lemmas.

***Lemma** 6.2:* Consider the closed-loop system in Figure 1, where the plant is a DLTI system described by equations (1) and (2), $\mathbb{C} = \mathbb{I}$. If $\mathrm{E}\{\boldsymbol{\varepsilon}(k)\boldsymbol{\varepsilon}^T(k)\} < \infty$, then

$$\lim_{k\to\infty} \frac{I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{r}^k)}{k} \geq \sum_{i} \log_2\left(|\lambda_i(A)|\right).$$

Proof: The mutual information $I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{r}^k)$ may be expanded as:

$$\begin{aligned} I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{r}^k) &= h(\mathbf{x}(0)|\mathbf{r}^k) - h(\mathbf{x}(0)|\mathbf{e}^k,\mathbf{r}^k); \\ &= h(\mathbf{x}(0)) - h(\mathbf{x}(0)|\mathbf{e}^k,\mathbf{r}^k); \end{aligned}$$

where we have used the fact that $\mathbf{x}(0)$ and $\mathbf{r}^k$ are independent. If we focus on the quantity $h(\mathbf{x}(0)|\mathbf{e}^k,\mathbf{r}^k)$ and using the properties of entropy we obtain:

$$\begin{aligned} &h(\mathbf{x}(0)|\mathbf{e}^k,\mathbf{r}^k) \\ &\quad= h(\mathbf{x}(0) + A^{-k}\sum_{i=0}^{k} A^{k-i-1}Bf_3(\mathbf{e}^i)|\mathbf{e}^k,\mathbf{r}^k); \\ &\quad= h(-A^{-k}(\boldsymbol{\varepsilon}(k) - \mathbf{r}(k))|\mathbf{e}^k,\mathbf{r}^k); && (9) \\ &\quad= h(-A^{-k}\boldsymbol{\varepsilon}(k)|\mathbf{e}^k,\mathbf{r}^k); && (10) \\ &\quad\leq h(-A^{-k}\boldsymbol{\varepsilon}(k)); && (11) \\ &\quad\leq \frac{1}{2}\log_2\Big((2\pi e)^n\mathbf{det}(\mathrm{Cov}\{-A^{-k}\boldsymbol{\varepsilon}\})\Big); && (12) \\ &\quad= \frac{n}{2}\log_2\Big(2\pi e\Big) + \frac{1}{2}\log_2\Big(\mathbf{det}(-A^{-k}\mathrm{Cov}\{\boldsymbol{\varepsilon}\}(-A^{-k})^T)\Big); \\ &\quad= \frac{n}{2}\log_2\Big(2\pi e\Big) + \frac{1}{2}\log_2\Big(\mathbf{det}(A^{-k}(A^{-k})^T)\Big) \\ &\qquad + \frac{1}{2}\log_2\Big(\mathbf{det}(\mathrm{Cov}\{\boldsymbol{\varepsilon}\})\Big); \\ &\quad= \frac{n}{2}\log_2\Big(2\pi e\Big) - k\sum_i \log_2\Big(|\lambda_i(A)|\Big) \\ &\qquad + \frac{1}{2}\log_2\Big(\mathbf{det}(\mathrm{Cov}\{\boldsymbol{\varepsilon}\})\Big). \end{aligned}$$

Where equation (9) is due to equation (4), equation (10) is due to Property 3.1.(e), equation (11) is due to Property

3.1.(a), and equation (12) is due to Property 3.1.(b). From these simplifications we obtain

$$
\begin{aligned}
& I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{r}^k) \\
& \geq \quad h(\mathbf{x}(0)) - \frac{n}{2}\log_2(2\pi e) + k\sum_i \log_2\Big(|\lambda_i(A)|\Big) \\
& \qquad - \frac{1}{2}\log_2\Big(\mathbf{det}(\mathrm{Cov}\{\ \boldsymbol{\varepsilon}\ \})\Big).
\end{aligned}
$$

Finally, if we divide by k and take the limit as $k \to \infty$, we obtain:

$$
\lim_{k\to\infty} \frac{I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{r}^k)}{k} \geq \sum_i \log_2\Big(|\lambda_i(A)|\Big).
$$

since $\boldsymbol{\varepsilon}$ is a second order process.

■

We note that for Lemma 6.2, we have assumed that $\mathbf{y}(k) = \mathbf{x}(k)$, i.e. the entire state is available for measurement. However, the lemma still holds when the output is only one component of the state vector (single output), e.g. $\mathbf{y}(k) = \mathbf{x}_1(k)$. In that case, we need to guarantee that the components of the state that do not appear in the output remain bounded. The only component that can grow unbounded is the one that appears in the output (in the case of an unbounded reference signal). For example, if the plant is a third order system ($n = 3$) and $\mathbb{C} = \begin{bmatrix}1 & 0 & 0\end{bmatrix}$, we have to guarantee that the difference between the reference signal and the output $\mathbf{y} = \mathbf{x}_1$ must remain bounded; and that the state components that do not appear in the output $\{\mathbf{x}_2, \mathbf{x}_3\}$ remain bounded, i.e., $E\{\mathbf{x}_j(k)\mathbf{x}_j(k)^T\} < \infty, \forall j \in \{2,4\}$. Before generalizing Theorem 6.2 we introduce the following notation:

- $\mathbf{y} = \mathbf{x_j}$
- Let $\mathbf{x}_{\bar{Y}}$ be the vector of state components that do not appear in output $\mathbf{y}$.

For example, if $\mathbb{C} = \begin{bmatrix}1 & 0 & 0\end{bmatrix}$, then $\mathbf{x}_j = \{\mathbf{x}_1\}$ whereas $\mathbf{x}_{\bar{Y}} = \{\mathbf{x}_2, \mathbf{x}_3\}$. We then prove the following.

Lemma 6.3: Consider closed-loop system given in Figure 1, where the plant is a DLTI system described by equation (1) and $\mathbf{y} = q\mathbf{x}_j$ for some $j \in \{1,\ldots,n\}$, q a non-zero constant. If $\mathrm{E}\{\boldsymbol{\varepsilon}(k)\boldsymbol{\varepsilon}^T(k)\} < \infty$ and $\mathrm{E}\{\mathbf{x}_{\bar{Y}}(k)\mathbf{x}_{\bar{Y}}^T(k)\} < \infty$, then

$$
\lim_{k\to\infty} \frac{I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{r}^k)}{k} \geq \sum_i \log_2\Big(|\lambda_i(A)|\Big).
$$

Proof: $I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{r}^k)$ may be expanded as

$$
I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{r}^k) = I(\mathbf{x}_j(0);\mathbf{e}^k|\mathbf{r}^k) + I(\mathbf{x}_{\bar{Y}}(0);\mathbf{e}^k|\mathbf{r}^k,\mathbf{x}_j(0)); \quad (13)
$$

where $\mathbf{x}_j = \mathbf{y}$, and $\mathbf{x}_{\bar{Y}}$ are the states that do not appear in $\mathbf{y}$. We also know that $\mathbf{y}(k) = \mathbf{r}(k) - \boldsymbol{\varepsilon}(k) = q\mathbf{x}_j(k)$, then $\mathbf{y}(k)$ may be expressed as $\mathbf{y}(k) = q\lambda_{x_j}^k \mathbf{x}_j(0) + G(\mathbf{e}^k)$. Where $G(\mathbf{e}^k)$ is a function of the error $\mathbf{e}^k$ and $\lambda_{x_j}^k$ is the eigenvalue corresponding to x_j. We may also expand $I(\mathbf{x}_j(0);\mathbf{e}^k|\mathbf{r}^k)$ using the definition of mutual information:

$$
\begin{aligned}
I(\mathbf{x}_j(0);\mathbf{e}^k|\mathbf{r}^k) &= h(\mathbf{x}_j(0)|\mathbf{r}^k) - h(\mathbf{x}_j(0)|\mathbf{e}^k,\mathbf{r}^k); \\
&= h(\mathbf{x}_j(0)) - h(\mathbf{x}_j(0)|\mathbf{e}^k,\mathbf{r}^k); \\
&= h(\mathbf{x}_j(0)) \\
&\quad - h\Big(\lambda_{x_j}^{-k}(\boldsymbol{\varepsilon}(k) - \mathbf{r}(k) - G(\mathbf{e}^k))|\mathbf{e}^k,\mathbf{r}^k\Big).
\end{aligned}
$$

Consider the term $h\Big(q^{-1}\lambda_{x_j}^{-k}(\mathbf{r}(k) - \boldsymbol{\varepsilon}(k) - G(\mathbf{e}^k))|\mathbf{e^k},\mathbf{r^k}\Big)$ which may be simplified to

$$
\begin{aligned}
& h\Big(q^{-1}\lambda_{x_j}^{-k}(\mathbf{r}(k) - \boldsymbol{\varepsilon}(k) - G(\mathbf{e}^k))|\mathbf{e}^k,\mathbf{r}^k\Big) \\
&= -k\log_2\Big(|\lambda_{x_j}|\Big) - \log_2\Big(|q|\Big) \\
&\quad + h\Big((\mathbf{r}(k) - \boldsymbol{\varepsilon}(k) - G(\mathbf{e}^k))|\mathbf{e}^k,\mathbf{r}^k\Big); \\
&= -k\log_2\Big(|\lambda_{x_j}|\Big) - \log_2\Big(|q|\Big) + h\Big(\boldsymbol{\varepsilon}(k)|\mathbf{e}^k,\mathbf{r}^k\Big); \quad (14) \\
&\leq -k\log_2\Big(|\lambda_{x_j}|\Big) - \log_2\Big(|q|\Big) + h\Big(\boldsymbol{\varepsilon}(k)\Big); \quad (15) \\
&\leq -k\log_2\Big(|\lambda_{x_j}|\Big) - \log_2\Big(|q|\Big) \\
&\quad + \frac{1}{2}\log_2\Big((2\pi e)\mathbf{det}(\mathrm{Cov}\{\boldsymbol{\varepsilon}\})\Big); \quad (16)
\end{aligned}
$$

where equation (14) is due to Property 3.1.(e), equation (15) is due to Property 3.1.(a) and equation (16) is due to Property 3.1.(b). We then have

$$
\begin{aligned}
I(\mathbf{x}_j(0);\mathbf{e}^k|\mathbf{r}^k) \quad \geq \quad & h(\mathbf{x}_j(0)) + k\log_2\Big(|\lambda_{x_j}|\Big) \\
& - \frac{1}{2}\log_2\Big(|q^{-2}2\pi e|\Big) \\
& - \frac{1}{2}\log_2\Big(\mathbf{det}(\mathrm{Cov}\{\boldsymbol{\varepsilon}\})\Big).
\end{aligned}
$$

Dividing by k and taking the limit when k tends to infinity we obtain

$$
\lim_{k\to\infty} \frac{I(\mathbf{x}_j(0);\mathbf{e}^k|\mathbf{r}^k)}{k} \geq \log_2\Big(|\lambda_{x_Y}|\Big). \quad (17)
$$

We note that the right side of equation (13) contains the term $I(\mathbf{x}_{\bar{Y}}(0);\mathbf{e}^k|\mathbf{r}^k,\mathbf{x}_j(0))$ but this term is bounded from below (Lemma 6.1) since $\mathrm{E}\{\mathbf{x}_{\bar{Y}}(k)\mathbf{x}_{\bar{Y}}^T(k)\} < \infty$ is required. Therefore

$$
\lim_{k\to\infty} \frac{I(\mathbf{x}_{\bar{Y}}(0);\mathbf{e}^k|\mathbf{r}^k,\mathbf{x}_j(0))}{k} \geq \sum_{i\neq j} \log_2\Big(|\lambda_i(A)|\Big). \quad (18)
$$

From inequalities (17) and (18) we have

$$
\lim_{k\to\infty} \frac{I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{r}^k)}{k} \geq \sum_i \log_2\Big(|\lambda_i(A)|\Big).
$$

■

VII. Results

Using the results in Section VI, we find limitations on tracking systems that are imposed by the presence of a finite capacity channel. We consider the expression $I(\mathbf{r}^k;\hat{\mathbf{y}}^k)$ instead of $I(\mathbf{r}^k;\mathbf{y}^k)$. Although $I(\mathbf{r}^k;\mathbf{y}^k)$ provides the actual information between the output and the reference signals, the former is easier to calculate than the later. The mutual information $I(\mathbf{r}^k;\hat{\mathbf{y}}^k)$ represents the information between the transmitted feedback, i.e., $\hat{\mathbf{y}}^k$, and the reference signal. If this mutual information happens to be zero, all information contained in the feedback signal about the reference signal was lost and the error $\mathbf{e}$ used to generate the control signal is useless. In fact, $I(\mathbf{r};\hat{\mathbf{y}})$ measures the usefulness of feedback. By the properties of mutual information, we have

$$I((\mathbf{r}^k,\mathbf{x}(0));\hat{\mathbf{y}}^k) = I(\mathbf{r}^k;\hat{\mathbf{y}}^k) + I(\mathbf{x}(0);\hat{\mathbf{y}}^k|\mathbf{r}^k). \tag{19}$$

From the definition of mutual information, Property 3.1.(e), and from the fact that $\mathbf{e}^k = \mathbf{r}^k - \hat{\mathbf{y}}^k$, we have

$$\begin{aligned} I(\mathbf{x}(0);\hat{\mathbf{y}}^k|\mathbf{r}^k) &= h(\hat{\mathbf{y}}^k|\mathbf{r}^k) - h(\hat{\mathbf{y}}^k|\mathbf{x}(0),\mathbf{r}^k); \\ &= h(\mathbf{e}^k|\mathbf{r}^k) - h(\mathbf{e}^k|\mathbf{x}(0),\mathbf{r}^k); \\ &= I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{r}^k). \end{aligned} \tag{20}$$

From equation (19) and (20) we have

$$I((\mathbf{r}^k,\mathbf{x}(0));\hat{\mathbf{y}}^k) = I(\mathbf{r}^k;\hat{\mathbf{y}}^k) + I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{r}^k). \tag{21}$$

From equation (21) and knowing that $kC_f \geq I((\mathbf{r}^k,\mathbf{x}(0));\hat{\mathbf{y}}^k)$ we obtain

$$I(\mathbf{r}^k;\hat{\mathbf{y}}^k) \leq kC_f - I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{r}^k). \tag{22}$$

By Lemma 6.2, and dividing equation (22) by k and taking the limit as $k \to \infty$, we finally have

$$I_\infty(\mathbf{r};\hat{\mathbf{y}}) \leq C_f - \sum_i \log_2\left(|\lambda_i(A)|\right).$$

We summarize this result in the following lemma:

***Lemma** 7.1:* Consider the closed-loop system given in Figure 1, where the plant is a DLTI system described by equations (1) and (2), a feedback capacity C_f in the channel. If $\mathrm{E}\{\boldsymbol{\varepsilon}(k)\boldsymbol{\varepsilon}(k)^T\} < \infty$, then

$$I_\infty(\mathbf{r};\hat{\mathbf{y}}) \leq C_f - \sum_i \log_2\left(|\lambda_i(A)|\right).$$

We note from Lemma 7.1 that if the channel does not have a minimum capacity of $\sum_i \log_2(|\lambda_i(A)|)$, the feedback signal does not provide any information of the reference signal. In other words, the advantages of feedback are completely lost and it makes no sense to use it. Lemma 6.2 is one of the main contributions of this work. We note that Lemma 6.3 is needed when the output is only one of the state components and not the whole state.

A. Limitations on the reference signals

The results of the previous sections deal with the idea of bounding the error signal, $\boldsymbol{\varepsilon}(k) = \mathbf{r}(k) - \mathbf{y}(k)$. However, it is well known that given a plant and a particular controller, there will be limitations on the type of signals that may be tracked. We show next that a tracking system may be thought of as a channel where the reference signal is the input message, the closed-loop is a feedback channel (with the encoder-decoder embedded) and the system output is the received message. Under this scenario good message estimation is synonymous with good tracking. We consider $\boldsymbol{\varepsilon} = \mathbf{r} - \mathbf{y}$ as the error estimate of the message. Note from Property 3.1.(i), that

$$E\{(\mathbf{r}-\mathbf{y})^2\} \geq \frac{1}{2\pi e} 2^{2h(\mathbf{r})}.$$

This inequality captures the idea that the greater is the entropy of the reference signal, the larger is the error signal, $\boldsymbol{\varepsilon}$. Moreover, since $E\{(\mathbf{r}-\mathbf{y})^2\}$ is a nonnegative number, we note that the error between the output and the reference cannot reach zero unless the reference signal is deterministic ($h(\mathbf{r}) = -\infty$). In other words, perfect tracking is not possible and tracking gets worse for high entropy reference signals regardless of the type or quality of the channel and the controller. Moreover, the following result holds regardless of the plant. Let us consider that the expected value of $(\boldsymbol{\varepsilon}^k)^2$ given the entire past $\boldsymbol{\varepsilon}_0^{k-1}$ as k tends to infinity given by

$$\sigma_\infty^2(\mathbf{r}) = \lim_{k\to\infty} E\{\boldsymbol{\varepsilon}^2(k)|\boldsymbol{\varepsilon}(k-1)\}.$$

From information theory, the entropy rate lower-bounds the variance $\sigma_\infty^2(\mathbf{r})$:

$$\sigma_\infty^2(\mathbf{r}) \geq \frac{1}{2\pi e} 2^{2h_\infty(\mathbf{r})}.$$

We then obtain the following lemma.

***Lemma** 7.2:* Consider the closed-loop system given in Figure 1, where the plant is a DLTI system described by equations (1) and (2). Then the best estimator $\mathbf{y}$ for $\mathbf{r}$ is bounded as

$$E\{(\mathbf{r}-\mathbf{y})^2\} \geq \frac{1}{2\pi e} 2^{2h(\mathbf{r})}. \tag{23}$$

Moreover, the variance of the best reference estimator, $\sigma_\infty^2(\mathbf{r})$, is bounded from below as follows

$$\sigma_\infty^2(\mathbf{r}) \geq \frac{1}{2\pi e} 2^{2h_\infty(\mathbf{r})}. \tag{24}$$

VIII. EXAMPLES

The results derived in so far are necessary conditions but not sufficient. Since the quantity $I_\infty(\mathbf{r};\hat{\mathbf{y}})$ implies correlation of signals and not necessarily that $\mathbf{y}$ is tracking $\mathbf{r}$. The following examples capture how conservative the results of this work are.

A. Example 1: Erasure Channel

We consider the tracking problem shown in Figure 1 for the reference signal, $\mathbf{r}(k)$. The reference signal is assumed to be a white Gaussian sequence, with zero-mean and with $\sigma_r^2 = 1$. We consider a memoryless erasure channel as shown in Figure 2 in the feedback link with limited rate and a probability of receiving the state measurement of $p_\gamma = 0.70479$. The probability of dropping a packet is therefore $1 - p_\gamma$. We consider a two-part encoder-decoder scheme: First, the encoder converts the real state-vector measured, $\mathbf{x}(k)$, to its binary form, truncates the binary representation to its R most significant bits, then encapsulates the bits in a packet and send the packet through the channel. If the packet is not dropped, the decoder on the receiver site receives

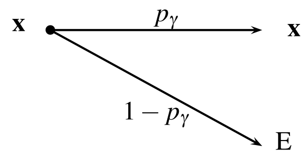

Fig. 2. Erasure Channel Scheme.

the packet, extracts the bits and converts them to its real number representation. If the receiver does not receive a packet, the decoder will assume that a zero was sent and the controller does not apply any control signal. In [10] it is shown that for the scalar case, this scheme guarantees that the error between the actual measurement signal and the decoded signal, $\boldsymbol{\varepsilon}(k) = \mathbf{x}(k) - \bar{\mathbf{x}}(k)$, is bounded and that the feedback channel capacity $C_f = \log_2(a)/p_\gamma$ is achieved. The scheme also assumes that the decoder knows exactly the operation of the encoder and that both have access to the control signal. Consider the following plant: $\mathbf{x}(k+1) = 4.33\mathbf{x}(k) + \mathbf{u}(k);\mathbf{y}(k) = \mathbf{x}(k)$; and $\mathbf{u}(k) = 4.33(\mathbf{r}(k) - \bar{\mathbf{y}}(k))$. One limitation of our result is that it is given in terms of the mutual information rate, which is difficult to compute for this type of problems. However, we know that it imposes a limit to guarantee that $\mathrm{E}\{\boldsymbol{\varepsilon}(k)\boldsymbol{\varepsilon}^T(k)\} < \infty$. In order to explore what happens to $\mathrm{E}\{\boldsymbol{\varepsilon}(k)\boldsymbol{\varepsilon}^T(k)\}$, we plot the power spectrum of $\boldsymbol{\varepsilon}$, $S_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}(\omega)$ whose enclosed area from $[-\pi,\pi]$ is equivalent to the squared output average of $\boldsymbol{\varepsilon}$, i.e.,

$$E\{\boldsymbol{\varepsilon}^2\} = \int_{-\pi}^{\pi} S_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}(\omega)d\omega. \tag{25}$$

According to Theorem 7.1, the minimum feedback channel capacity for stabilization needed is 3 bits/time-step. The power spectrum density is shown in Figure 3, where we notice that the power spectrum is bounded and, therefore, $E\{\boldsymbol{\varepsilon}^2(k)\}$ is finite. If, instead of using 3 bits/time-step, we use 2 bits/time-step, we obtain the new power spectrum of the error in Figure 4. Note that the power spectrum is becoming

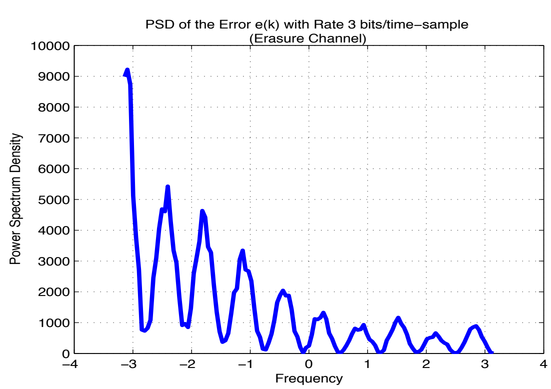

Fig. 3. Example with Erasure Channel and Bit Rate of 3 bits/time-step.

unbounded and so the area below the curve, i.e., $E\{\boldsymbol{\varepsilon}^2(k)\}$ is no longer finite.

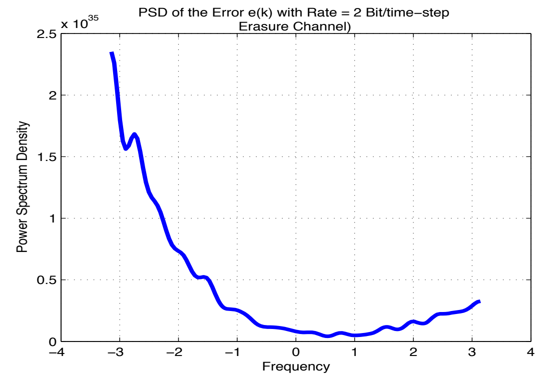

Fig. 4. Example with Erasure Channel and Bit Rate of 2 bits/time-step.

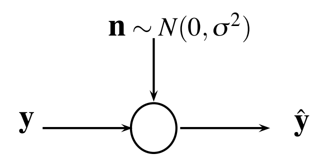

Fig. 5. AGWN Channel Scheme.

B. Example 2: AWGN Channel

We consider the problem of tracking (see Figure 1) a reference signal, $\mathbf{r}(k)$, which is assumed to be a white Gaussian sequence with zero-mean and $\sigma_r^2 = 5000$. We consider a memoryless AWGN channel (Figure 5) in the feedback link with feedback channel capacity, $C_f = (1/2)\log_2(1+P/\Phi)$, where Φ is the noise variance and P is the power constraint such that $E\{\hat{\mathbf{y}}^2\} \leq P$. The variance Φ is varied in the range $[1000; 200000]$, i.e, the SNR from the reference signal to the noise signal changes between 0.025 and 5. Let the plant be: $\mathbf{x}(k+1) = 2\mathbf{x}(k) + \mathbf{u}(k)$; $\mathbf{y}(k) = \mathbf{x}(k)$; and $\mathbf{u}(k) = 2(\mathbf{r}(k) - \hat{\mathbf{y}}(k))$. In this example, we can actually measure the mutual information rate between the reference and the feedback signal for different SNR values, and monitor the upperbound $C_f - \log_2(a)$ given in Lemma 7.1. We use previous results from [9] to measure the mutual information rate, $I_\infty(\mathbf{r};\hat{\mathbf{y}})$, and results from [1] to design a controller. Since the system is linear and all inputs are white Gaussian processes, the output $\hat{\mathbf{y}}$ is also a Gaussian process. From [9], we know that if $\mathbf{r}$ and $\hat{\mathbf{y}}$ are two jointly-Gaussian stationary processes, with spectral densities $\Phi_r(\omega)$ and $\Phi_{\hat{\mathbf{y}}}(\omega)$, and if we define $w = \begin{bmatrix}\mathbf{r}\\ \hat{\mathbf{y}}\end{bmatrix}$, with spectral density $\Phi_w(\omega)$, the mutual information rate of $\mathbf{r}$ and $\hat{\mathbf{y}}$ is given by

$$I_\infty(\mathbf{r};\hat{\mathbf{y}}) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{\det(\Phi_r(\omega))\det(\Phi_{\hat{y}}(\omega))}{\det(\Phi_w(\omega))}d\omega. \tag{26}$$

Figure 6 illustrates that we obtain the expected result. The mutual information rate tends to zero for low SNR and, for this particular case reaches its upper bound, i.e. $C_f - \log_2(a)$, for high SNR. We see that this upper bound never reaches a value of zero (actually, for a SNR of 0, its value is 0.61 bits/time). We conclude, however, that the bound for good tracking, as measured by $I_\infty(\mathbf{r};\hat{\mathbf{y}})$, is higher that the one for stabilization.

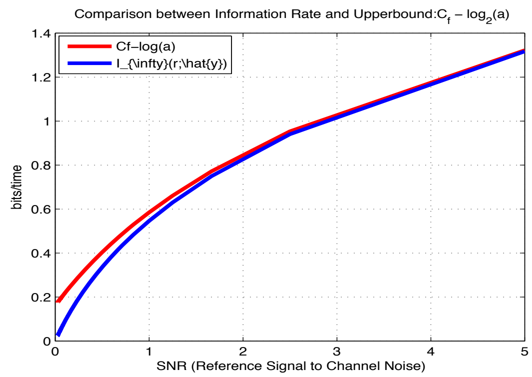

Fig. 6. Example with AWGN Channel for different SNR levels.

C. Example 3: Limitations Due to the Entropy of the Reference

This example is presented to illustrate the results in Subsection VII-A. Let us consider the following system/controller: $\mathbf{x}(k+1) = 1.2\mathbf{x}(k) + \mathbf{u}(k)$; $\mathbf{y}(k) = \mathbf{x}(k)$; and $\mathbf{u}(k) = \mathbf{r}(k) - 0.3\bar{\mathbf{y}}(k)$. Assume that the reference signal is given by

$$\mathbf{r}(k) = 2 + \mathbf{n}_r(k);$$

where $\mathbf{n}_r$ has a Gaussian distribution with zero mean and variance σ^2 with $\sigma = 2$. Moreover, we assume a perfect feedback of the output. Since the reference is a Gaussian signal, by substitution in Property 3.1.(b), the differential entropy of the reference signal is given by $h(\mathbf{r}) = (1/2)\log_2\left(2\pi e\sigma^2\right)$. According to Lemma 7.2, the lower bound in the right side of equation (23) is σ^2. In order to plot $E\{(\mathbf{r}-\mathbf{y})^2\} = E\{(\boldsymbol{\varepsilon})^2\}$, we ran 1000 simulations and averaged then. The average result of these simulations is shown in Figure 7, which clearly illustrates the result of Lemma 7.2.

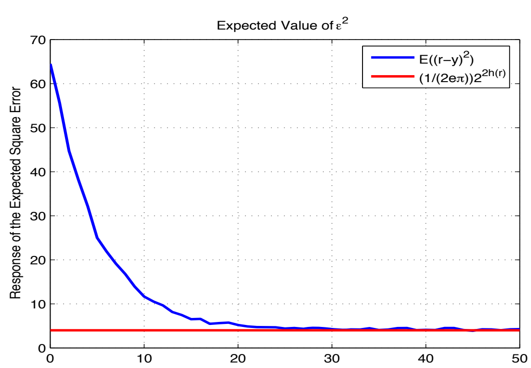

Fig. 7. Example with Gaussian Reference Signal.

IX. Conclusions and Future Work

This paper has provided information theoretical conditions for tracking problems in feedback control systems. The results were obtained in terms of the mutual information rate between the feedback and the reference signals, as well as the channel capacity and the unstable eigenvalues of the LTI system. The lower bound for the channel capacity obtained was expected since it corresponds to the one obtained in previous literature as [12] and [7]. We also obtained a lower bound in terms of the entropy of the reference signal for the maximum achievable accuracy in a tracking system, in the absence of constraint channel. Our results were verified with several examples and simulations. We plan to extend these ideas by exploring the impact of non-minimum phase zeros of the plant. We are also investigating various frequency domain interpretations of our results. Finally, we are in the process of studying steady-state conditions in tracking deterministic reference signals, which do not have a clear interpretation in terms of information theory concepts.

References

[1] J. Braslavsky, R. Middleton, and J. Freundenber. Feedback stabilization over signal-to-noise ratio constrained channels. *IEEE Transactions on Automatic Control*, 52(8):1391–1403, August 2007.

[2] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2006.

[3] N. Martins. *Information Theoretic Aspects of the Control and Mode Estimation of Stochastic Systems*. PhD thesis, Massachusetts Institute of Technology, 2004.

[4] N. Martins and M. Dahleh. Feedback control in the presence of noisy channels: "Bode-like" fundamental limitations of performance. *To appear in IEEE Transactions of Automatic Control*, 2007.

[5] N. Martins and M. Dahleh. Fundamental limitations of disturbance attenuation in the presence of side information. *IEEE Transactions on Automatic Control*, 52(1):56–66, January 2007.

[6] N. Martins, M. Dahleh, and N. Elia. Feedback stabilization of uncertain systems using a stochastic digital link. *Proceedings of the IEEE Conference on Decision and Control*, 2:1889–1895, December 2004.

[7] G. Nair and J. Evans. Stabilizability of stochastic linear systems with finite feedback data rates. *SIAM Journal on Control and Optimization*, 43:413–436, September 2004.

[8] A. Papoulis. *Probability Random Variables, and Stochastic Processes*. Mc Graw Hill, New York, New York, 1991.

[9] M. Pinsker. *Information and Information Stability of Random Variables and Processes*. Holden Day, San Francisco, California, 1964.

[10] L. Shi and R. Murray. Towards a packet-based control theory - Part II: Rates issues. *Proceedings of the American Control Conference*, pages 3482–3487, June 2006.

[11] H. Stark and J. Woods. *Probability and Random Processes with Applications to Signal Processing*. Pearson Education, Upper Saddle River, New Jersey, 2002.

[12] S. Tatikonda and S. Mitter. Control under communication constraints. *IEEE Transactions on Automatic Control*, 49(7):1056–1068, July 2004.

[13] H. Touchette and S. Lloyd. Information-theoretic limits of control. *Physical Review Letters*, 6(84):1156–1159, February 2000.

[14] H. Zhang and Y. Sun. Bode integrals and laws of variety in linear control systems. In *Proceedings of the American Control Conference*, volume 1, pages 66–70, June 2003.

[15] H. Zhang and Y. Sun. Information theoretic limit and bound of disturbance rejection in LTI systems: Shannon entropy and H_∞ entropy. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, volume 2, pages 1378–1383, October 2003.

[16] H. Zhang and Y. Sun. Directed information and mutual information in linear feedback tracking systems. In *Proceedings of the IEEE World Congress on Intelligent Control and Automation*, volume 1, pages 723–727, June 2006.