Feedback Control in the Presence of Noisy Channels: "Bode-Like" Fundamental Limitations of Performance

Nuno C. Martins, *Member, IEEE*, and Munther A. Dahleh, *Fellow, IEEE*

***Abstract*—This paper addresses fundamental limitations of feedback using information theoretic conservation laws and flux arguments. The paper has two parts. In the first part, we derive a conservation law dictating that causal feedback cannot reduce the differential entropy inserted in the loop by external sources. An interpretation of this result is that the total randomness induced by disturbances, as measured by differential entropy, cannot be reduced by causal feedback; it can only be re-allocated in time or in frequency (if well defined). Under asymptotic stationarity assumptions, this result has a spectral representation which constitutes an extension of Bode's inequality for arbitrary feedback. Our proofs make clear the role of causality, as well as how stability assumptions impact the final result. In the second part, we derive an inequality unveiling that the feedback loop must be able to convey information originating from two independent sources: 1) initial states of the physical plant; 2) exogenous disturbance signals. By using such principle, we construct a variety of information rate (information flux) inequalities. Furthermore, we derive a universal performance bound which is parameterized solely by the feedback capacity and the parameters of the plant. The latter is a new fundamental limitation, which is different from Bode's classical result, indicating that finite feedback capacity brings a new type of performance bound.**

***Index Terms*—Differential entropy, feedback capacity.**

I. Introduction

FUNDAMENTAL limitations play a pivotal role in most branches of science. From an engineering perspective, fundamental limits are used not only to discard impossible specifications, but they are also important to prove the optimality of certain policies. A particular instance of such a strategy is Shannon's converse Theorem [3], which is used to certify the quality of coding strategies.

Most converse theorems in Information Theory are derived in great generality. Consequently, it is not surprising that Information Theory is being used in other fields as a powerful conceptual aid for deriving fundamental limits [33]. The contribution of this paper is to use Information Theory not only in extending existing fundamental limitations of feedback systems but also in deriving new bounds and applications.

The interplay between Control and Information Theory is happening along three main directions. One avenue is the study of control under information constraints, which already has a vast collection of insightful results in stabilization [9]–[21], [39], [40]. The basic framework comprises a plant, a channel, an encoder and a decoder, which implicitly embeds a controller. Performance limits have also been addressed for specific channels, such as the deterministic bit-rate channel [14] and the Gaussian channel [37], where, in both cases, the measure of performance is of the expected (average) power type. Effective coding paradigms for the control over noisy channels have been pioneered by the authors of [44], [45] and further design issues have been addressed in [28], [1], [27].

The second avenue in which Information Theory may help Control Theory is in the derivation of fundamental limitations of feedback in a general setting, i.e., not restricted to dealing with information constraints. The article [23] was one of the first to point in that direction, culminating with an extension of Bode's integral formula [2] for a class of differentiable non-linear systems. Prior to that, the authors of [48] have established a conservation principle for linear feedback systems in terms of Kolmogorov-Sinai entropy. The work by [26] gives an entropic interpretation to optimal control, which is advantageous in adaptive control as well as in hierarchical control. More recently, the authors of [36] were able to characterize the information flow in the Kalman filter, which has lead to a new energy related interpretation. Information theoretic techniques have also been used in [12] to study the fundamental limits of disturbance attenuation in feedback systems, assuming that a remote preview of the disturbance is available.

Yet a third avenue of interaction is the use of Control Theory in Information Theory, such as in the work of [5], [31], [32].

A. Contributions of the Paper

Understanding the fundamental limitations of performance in a feedback system is critical for effective control design. One of the most well known tradeoffs is the water-bed effect for discrete-time, linear and time-invariant feedback systems, which results from Bode's integral formula [2]. In such a classical theory, the transfer function between the disturbance $\mathbf{d}$ and $\mathbf{e} = \mathbf{u} + \mathbf{d}$ (see Fig. 1) is called the sensitivity [4] and is represented by $S(z)$. Bode's result, for a strictly proper loop gain, is expressed as [49]

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log |S(e^{j\omega})| d\omega = \sum_{\lambda \in \mathcal{UP}} \log |\lambda| \tag{1}$$

Manuscript received December 20, 2005; revised August 3, 2007. Current version published September 10, 2008. This paper was presented in part at the Allerton Conference (ACC), 2005. This work was supported by the University of California at Los Angeles, MURI Project Title: Cooperative Control of Distributed Autonomous Vehicles in Adversarial Environments" Award: 0205-G-CB222 and by the Portuguese Foundation for Science and Technology, and the European Social Fund, PRAXIS BD19630/99. Recommended by Associate Editor S. Dey.

N. C. Martins is with the ECE Department and the Institute for Systems Research, University of Maryland, College Park, MD 20742 USA (e-mail: nmartins@isr.umd.edu).

A. Dahleh is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Digital Object Identifier 10.1109/TAC.2008.929361

0018-9286/$25.00 © 2008 IEEE

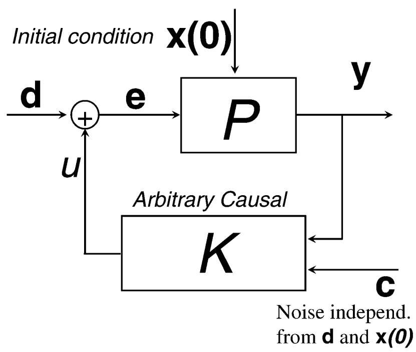

Fig. 1. Basic framework for control by causal feedback.

where $\mathcal{UP}$ are the unstable poles of the open loop system[1] [4], which is assumed to be rational and strictly proper. By using feedback, one would expect that sensitivity can be decreased. On the other hand, (1) quantifies a fundamental limitation which indicates that sensitivity can be, at most, *shaped* in frequency. Equivalently, $|S(e^{j\omega})|$ cannot be made small at all frequencies.

In the first part of this paper, we show that causality is the only requirement for deriving an inequality involving the differential entropy of $(\mathbf{e}(0),\dots,\mathbf{e}(k))$ and $(\mathbf{d}(0),\dots,\mathbf{d}(k))$, which holds regardless of the feedback and time horizon k. In addition, we show that, under deterministic feedback, the aforementioned quantities are related via an equation that also involves the mutual information between the initial state of the plant and $(\mathbf{e}(0),\dots,\mathbf{e}(k))$. An extension of these results, in the limit when the time horizon goes to infinity, can be found in [12]. In particular, we show that under asymptotic stationary assumptions, techniques already found in [12] can be used to obtain an integral inequality that extends Bode's result. We also show that the constant in the right hand side of (1) has an information theoretic origin. In the second part of the paper, we derive an information rate separation principle for feedback systems in the presence of an external disturbance $\mathbf{d}$. According to this principle, the feedback loop must be able to convey information originating from two independent sources: 1) initial states of the physical plant; 2) exogenous disturbance signals. From such a principle, we derive several information rate inequalities, which we use in deriving a universal bound to performance attenuation as a function of the feedback capacity.

In addition to their importance in the classical engineering context, our results have also been used to explain the behavior of certain biological feedback systems [38].

The following notation is adopted:

- Finite segments of sequences are indicated as $a_{k_{\min}}^{k_{\max}} \overset{\text{def}}{=} (a(k_{\min}),\dots,a(k_{\max}))$. We also adopt $a^k = a_0^k$. If $k_{\max} < k_{\min}$ then $a_{k_{\min}}^{k_{\max}} = \emptyset$. Whenever it is clear from the context, we refer to an infinite sequence a^∞ of elements in $\mathbb{R}^n$ as a.
- Random variables are represented using boldface letters, such as $\mathbf{a}$.
- If $\mathbf{a}(k)$ is a time sample of a stochastic process, then we use $a(k)$ to indicate a time sample of a specific realization. Similar to the convention used for deterministic sequences, we may denote the whole stochastic process $\mathbf{a}^\infty$ just as $\mathbf{a}$. Finite segments of stochastic processes may also be indicated as $\mathbf{a}_{k_{\min}}^{k_{\max}}$ or as $\mathbf{a}^k$ if the time index starts at $k_{\min} = 0$.
- The probability density of a random variable $\mathbf{a}$, if it exists, is denoted as p_a. The joint probability density of $\mathbf{a}$ and $\mathbf{b}$ is denoted by $p_{a,b}$ and the conditional probability density of $\mathbf{a}$, given $\mathbf{b}$, is indicated as $p_{a|b}$.
- The expectation operator over $\mathbf{a}$ is written as $\mathcal{E}[\mathbf{a}]$.
- We write $\log_2(\cdot)$ simply as $\log(\cdot)$ and we adopt the convention $0 \log 0 = 0$.

[1]The poles of the open loop system comprise the poles of the plant and of the controller.

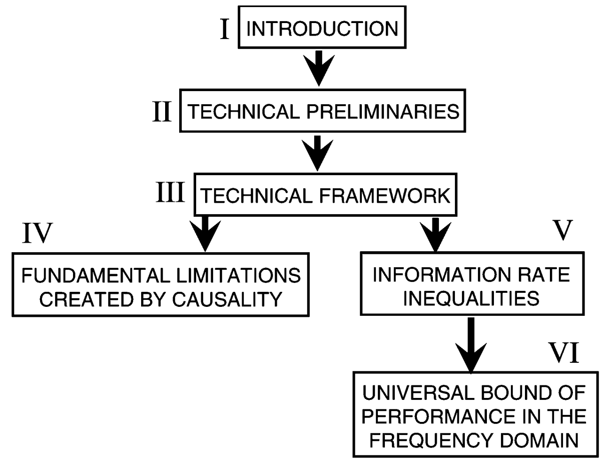

Fig. 2. Section precedence diagram.

The paper is organized into seven sections, which can be read in the order indicated by the diagram of Fig. 2. Section II provides the technical preliminaries, while Section III lays down the technical framework. The extension of Bode's result is derived in Section IV and Section V develops an information rate separation principle. The results of Section V are used in Section VI to prove a universal performance bound parameterized by feedback capacity. Section VII ends the paper with conclusions.

II. Technical Preliminaries

We start this section by summarizing the main definitions of Information Theory, which are used throughout the paper. We adopt [16], as a primary reference, because it addresses general probabilistic spaces in a unified framework. Most of the concepts can also be found in [3].

Definition 2.1: Let $\mathbf{a}$ and $\mathbf{b}$ be given random variables taking values in $\mathbb{R}^m$ with a well defined joint probability density. The mutual information, between $\mathbf{a}$ and $\mathbf{b}$, specified by $I : (\mathbf{a};\mathbf{b}) \rightarrow \mathbf{R}_+ \bigcup \{\infty\}$, is given by

$$I(\mathbf{a};\mathbf{b}) = \int_{\mathbb{R}^{2m}} p_{a,b}(\gamma_a,\gamma_b) \log\left(\frac{p_{a|b}(\gamma_a,\gamma_b)}{p_a(\gamma_a)}\right) d\gamma_a d\gamma_b.$$

The definition of mutual information for arbitrary alphabets can be found in [16, p.9]. The conditional mutual information between $\mathbf{a}$ and $\mathbf{b}$, given $\mathbf{c}$, is indicated as $I(\mathbf{a};\mathbf{b}|\mathbf{c})$. A rigorous definition of conditional mutual information is given in [16, p. 37].

The following gives a formula for computing $I(\mathbf{a};\mathbf{b}|\mathbf{c})$, which we could have used as an alternative definition:

$$I(\mathbf{a};\mathbf{b}|\mathbf{c}) = I((\mathbf{a},\mathbf{c});\mathbf{b}) - I(\mathbf{c};\mathbf{b}).$$

Definition 2.2: If $\mathbf{a}$ is a random variable, with alphabet $\mathbb{R}^q$, along with a probability density function $p_a(\cdot)$ then we define the differential entropy of $\mathbf{a}$ as

$$h(\mathbf{a}) = -\int_{\mathbb{R}^q} p_a(\gamma)\log p_a(\gamma)d\gamma.$$

If $\mathbf{b}$ is another random variable and $I(\mathbf{a};\mathbf{b}) < \infty$ then the conditional differential entropy of $\mathbf{a}$ given $\mathbf{b}$ is defined by

$$h(\mathbf{a}|\mathbf{b}) = h(\mathbf{a}) - I(\mathbf{a};\mathbf{b}). \tag{2}$$

Definition 2.3: (**Information Rate**) Let $\mathbf{a}$ and $\mathbf{b}$ be stochastic processes. The following is the definition of (mutual) information rate[2]

$$I_\infty(\mathbf{a};\mathbf{b}) = \limsup_{N\to\infty}\frac{I(\mathbf{a}^{N-1};\mathbf{b}^{N-1})}{N}.$$

The use of the information rate is motivated by its universality [3], i.e., it quantifies the rate at which information can be reliably transmitted through an arbitrary communication medium.

Definition 2.4: (**Entropy Rate**) For a given stochastic process, we also define entropy rate as

$$h_\infty(\mathbf{a}) = \limsup_{N\to\infty}\frac{h(\mathbf{a}^{N-1})}{N}. \tag{3}$$

A. Basic Properties of Differential Entropy and Mutual Information

The following is a list of properties used throughout the paper. Whenever appropriate, e.g., in proofs, we will refer to these properties by their numbers (P1)–(P5).

The proof of such properties may be found in [16] and, in some cases, in [3]:

- **(P1) Symmetry and positivity properties :** $I(\mathbf{a};\mathbf{b}) = I(\mathbf{b};\mathbf{a}) \geq 0$ and $I(\mathbf{a};\mathbf{b}|\mathbf{c}) \geq 0$.
- **(P2) Kolmogorov's formula** [3] (3.6.6 in [16]):

$$I((\mathbf{a},\mathbf{b});\mathbf{c}|\mathbf{d}) = I(\mathbf{b};\mathbf{c}|\mathbf{d}) + I(\mathbf{a};\mathbf{c}|(\mathbf{b},\mathbf{d})).$$

- **(P3) Data Processing inequality :** If ϕ and θ are measurable functions in the appropriate probability spaces then $I(\phi(\mathbf{a},\mathbf{c});\theta(\mathbf{b})|\mathbf{c}) \leq I(\mathbf{a};\mathbf{b}|\mathbf{c})$.
- **(P4) Iterated differential entropy :** Using Kolmogorov's formula (P2) and (2) we arrive at $h(\mathbf{a},\mathbf{b}) = h(\mathbf{a}|\mathbf{b}) + h(\mathbf{b})$.
- **(P5) Injective transformation in differential entropy :** Using a change of variables in the integrals of Definition 2.2, we have that if $\phi : \mathcal{B} \to \mathcal{A}$ is any given function then $h(\mathbf{a}|\mathbf{b}) = h(\mathbf{a} - \phi(\mathbf{b})|\mathbf{b})$.

[2]Throughout the paper, for simplicity, we refer to mutual information rate simply as information rate.

[3]Notice that 3.6.3 in [16] has a typographic mistake. The left hand side of the equality should be $I(\xi,\zeta)$.

III. Technical Framework

In this section, we introduce the performance measures used in our analysis, while discussing some of their properties and interpretations. In addition, we describe the basic feedback paradigm by means of a block diagram and three assumptions.

A. Performance Measures Under Asymptotic Stationarity: Definition of a Sensitivity-Like Function

In order to ascribe a frequency domain interpretation to our results, we adopt the following definition of asymptotic power spectral density.

Definition 3.1: **(Asymptotic Stationarity)** A given zero mean real stochastic process $\mathbf{a}$ is asymptotically stationary if the following limit exists for every $\gamma \in \mathbb{N}$:

$$\bar{R}_{\mathbf{a}}(\gamma) \stackrel{\text{def}}{=} \lim_{k\to\infty}\mathcal{E}[\mathbf{a}(k+\gamma)\mathbf{a}(k)]. \tag{4}$$

We also use (4) to define the following asymptotic power spectral density:

$$\hat{F}_{\mathbf{a}}(\omega) = \sum_{k=-\infty}^{\infty}\bar{R}_{\mathbf{a}}(k)e^{-j\omega k}. \tag{5}$$

Definition 3.2: **(A sensitivity-like function)** Consider the feedback loop shown in Fig. 1. If the stochastic processes $\mathbf{e}$ and $\mathbf{d}$ are asymptotically stationary then we define the following sensitivity-like function:

$$S_{\mathbf{d},\mathbf{e}}(\omega) \stackrel{\text{def}}{=} \sqrt{\frac{\hat{F}_{\mathbf{e}}(\omega)}{\hat{F}_{\mathbf{d}}(\omega)}}. \tag{6}$$

The following is a list of remarks regarding $S_{\mathbf{d},\mathbf{e}}(\omega)$:

- If the feedback system is linear, time-invariant and $\mathbf{c}$ is identically equal to zero then $S_{\mathbf{d},\mathbf{e}}(\omega)$ is the absolute value of the standard sensitivity function [4].
- In the general non-linear case, $S_{\mathbf{d},\mathbf{e}}$ will depend on the statistical properties of $\mathbf{d}$. It has been suggested in previous publications [25] that such a feature is intrinsic to feedback loops with general non-linear controllers. Limitations in terms of the ratio represented by $S_{\mathbf{d},\mathbf{e}}$ should be interpreted as follows: once we have a spectral model of the disturbance, say $\hat{F}_{\mathbf{d}}$, then limitations in $S_{\mathbf{d},\mathbf{e}}$ translate immediately to limitations in $\hat{F}_{\mathbf{e}}$. Clearly, for each spectral model of the disturbance, $S_{\mathbf{d},\mathbf{e}}$ gives as much information about $\hat{F}_{\mathbf{e}}$ as a classic sensitivity function would. As a consequence, our results, which express restrictions on $S_{\mathbf{d},\mathbf{e}}$, show, for any given disturbance spectrum $\hat{F}_{\mathbf{d}}$, that certain spectra $\hat{F}_{\mathbf{e}}$ are not attainable.
- If either $\mathbf{e}$ or $\mathbf{d}$ is not asymptotically stationary then $S_{\mathbf{d},\mathbf{e}}(\omega)$ is undefined. We have not tried to attribute a frequency domain interpretation based on non-stationary notions, such as wavelets or evolutionary power spectral

densities. In the absence of stationarity, we resort directly to entropy rates.

B. Performance Measures Using Entropy Rates

In the absence of asymptotic stationarity, we use entropy rates to gage performance, not only because it is technically convenient, but also because it is a fundamental quantity which can be related to other, more common, measures of performance. In standard texts, such as [24], [3], the entropy rate of a given stochastic process is interpreted as a measure of randomness, *energy* or disorder. The following inequalities relate the entropy rate with other performance measures[4]:

- for a general stochastic process $\mathbf{a}$, the following holds:

$$\limsup_{k\to\infty} \operatorname{Var}(\mathbf{a}(k)) \geq \frac{1}{2\pi e} 2^{2h_\infty(\mathbf{a})}$$

- if $\mathbf{a}$ is asymptotically stationary with an integrable power spectral density $\hat{F}_{\mathbf{a}}(\omega)$ then the following holds:

$$h_\infty(\mathbf{a}) \leq \frac{1}{4\pi}\int_{-\pi}^{\pi} \log\left(2\pi e \hat{F}_{\mathbf{a}}(\omega)\right).$$

C. Basic Feedback Paradigm and Assumptions

Throughout this article, we consider the general feedback scheme of Fig. 1. The following assumptions regarding the plant P, the feedback block K and the external excitation are made:

- Given n, the plant P is single input with state $\mathbf{x}(k)$ taking values in $\mathbb{R}^n$, and satisfying the following state-space equation:

$$\begin{aligned} \mathbf{x}(k+1) &= A\mathbf{x}(k) + B\mathbf{e}(k) \\ \mathbf{y}(k) &= C\mathbf{x}(k) \end{aligned} \tag{7}$$

  where the differential entropy of the initial state $\mathbf{x}(0)$ is assumed finite.
- The channel noise $\mathbf{c}$, the plant's initial condition $\mathbf{x}(0)$ and the disturbance $\mathbf{d}$ are mutually independent.
- The map K acts as $\mathbf{u}(k) = K(k, \mathbf{y}^k, \mathbf{c}^k)$, i.e., it deterministically maps past and present inputs to present outputs. As such, if K admits a state-space representation then the initial state must be deterministic.

As the reader will have the opportunity to infer, most of our results hold for multi-input plants. We have adopted the assumptions above just to make the paper more readable.

IV. Fundamental Limitations Created by Causality

Due to its importance, Bode's fundamental limitation has been extended to frameworks more general than the linear and time invariant one [29]. The multi-dimensional version was provided in [6], [7] while certain non-linear systems have been analyzed in [22], [23], [30]. Fundamental limits on tracking performance are addressed in [41], [42]. The authors of [46] have used properties of polynomials to obtain an extension of Bode's formula, under certain convergence conditions.

[4]The first fact follows from standard results in [3]. Since the last fact does not follow immediately from [3], a proof of this result can be found in ([12] Lemma 4.3)

Causality is the central assumption in obtaining a conservation law, which, under asymptotic stationarity assumptions, leads to an extension of Bode's result. The main result of this section is given in Theorem 4.2 and a frequency domain interpretation is provided in Lemma 4.3. We consider the scheme depicted in Fig. 1, where the feedback block K can be any causal function of $\mathbf{y}$ and $\mathbf{c}$, which includes maps that are non-linear, time-varying and operating on hybrid alphabets.

The following Lemma is the critical piece throughout this section:

Lemma 4.1: Consider the scheme of Fig. 1, where $\mathbf{d}^k, \mathbf{x}(0)$ and $\mathbf{c}^k$ are mutually independent, for all k. Causality of the feedback loop implies that the following holds:

$$I\left(\mathbf{d}(k);(\mathbf{u}^k,\mathbf{x}(0))\,|\mathbf{d}^{k-1}\right) = 0,\ k \in \mathbb{N}_+. \tag{8}$$

Proof: We start by realizing that causality of the feedback loop implies the following:

$$I(\mathbf{d}(k);(\mathbf{u}^k,\mathbf{x}(0),\mathbf{c}^k)|\mathbf{d}^{k-1}) \underset{\text{causality}}{=} I(\mathbf{d}(k);(\mathbf{x}(0),\mathbf{c}^k)|\mathbf{d}^{k-1}) \underset{\text{independence}}{=} 0 \tag{9}$$

where the first equality follows by noticing that $\mathbf{u}^k$ is a function of $\mathbf{x}(0), \mathbf{c}^k$ and $\mathbf{d}^{k-1}$. The fact that the aforementioned mutual information is equal to zero, follows from the mutual independence between $\mathbf{c}, \mathbf{x}(0)$ and $\mathbf{d}$. In order to arrive at (8) we apply the data processing inequality (P3)

$$0 \underset{(9)}{=} I(\mathbf{d}(k);(\mathbf{u}^k,\mathbf{x}(0),\mathbf{c}^k)|\mathbf{d}^{k-1}) \underset{(P3)}{\geq} I(\mathbf{d}(k);(\mathbf{u}^k,\mathbf{x}(0))|\mathbf{d}^{k-1}) \underset{(P1)}{\geq} 0 \tag{10}$$

□

The fact that (8) holds is enough to derive the following:

$$h(\mathbf{e}^k) = h(\mathbf{d}^k) + I(\mathbf{x}(0);\mathbf{e}^k) + \sum_{i=0}^{k} I(\mathbf{u}^i;\mathbf{e}(i)|\mathbf{e}^{i-1},\mathbf{x}(0)) \tag{11}$$

which, from the positivity property of mutual information (P1), implies the following:

$$h(\mathbf{e}^k) \geq h(\mathbf{d}^k). \tag{12}$$

The immediate implication of (12) is that causal feedback cannot reduce differential entropy. The proof of (11) goes as follows:

Proof of (11): Using (2) and Kolmogorov's formula (P2), we arrive at

$$I(\mathbf{d}(i);(\mathbf{u}^i,\mathbf{x}(0))|\mathbf{d}^{i-1}) = h(\mathbf{d}(i)|\mathbf{d}^{i-1}) - h(\mathbf{d}(i)|\mathbf{d}^{i-1},\mathbf{x}(0),\mathbf{u}^i). \tag{13}$$

Using Lemma 4.1 and (13), we find that

$$h(\mathbf{d}(i)|\mathbf{d}^{i-1}) = h(\mathbf{d}(i)|\mathbf{d}^{i-1},\mathbf{x}(0),\mathbf{u}^i). \tag{14}$$

Now notice that we can apply the change of variables $\mathbf{e} = \mathbf{d} + \mathbf{u}$ to (14) and use the injective transformation property of differential entropy (P5), to obtain

$$h(\mathbf{d}(i)|\mathbf{d}^{i-1}) = h(\mathbf{e}(i)|\mathbf{e}^{i-1}, \mathbf{x}(0), \mathbf{u}^i) \tag{15}$$

which summing over $i \in \{0, \ldots, k\}$ and using the iterated differential entropy property (P4), leads to

$$h(\mathbf{d}^k) = \sum_{i=0}^{k} h(\mathbf{e}(i)|\mathbf{e}^{i-1}, \mathbf{x}(0), \mathbf{u}^i). \tag{16}$$

By repeated application of (2) and Kolmogorov's formula (P2), we get

$$\begin{aligned} h(\mathbf{e}(i)&|\mathbf{e}^{i-1}, \mathbf{x}(0), \mathbf{u}^i) \\ &= h(\mathbf{e}(i)|\mathbf{e}^{i-1}) - I(\mathbf{x}(0); \mathbf{e}(i)|\mathbf{e}^{i-1}) \\ &\quad - I(\mathbf{u}^i; \mathbf{e}(i)|\mathbf{e}^{i-1}, \mathbf{x}(0)) \end{aligned} \tag{17}$$

which summing over $i \in \{0, \ldots, k\}$ and using (P4) and (P2), once more, leads to

$$\begin{aligned} \sum_{i=0}^{k} & h(\mathbf{e}(i)|\mathbf{e}^{i-1}, \mathbf{x}(0), \mathbf{u}^i) \\ &= h(\mathbf{e}^k) - I(\mathbf{x}(0); \mathbf{e}^k) \\ &\quad - \sum_{i=0}^{k} I(\mathbf{u}^i; \mathbf{e}(i)|\mathbf{e}^{i-1}, \mathbf{x}(0)). \end{aligned} \tag{18}$$

The final result follows from (18) and (16). □

The limiting case of (11) is stated in the following Theorem:

Theorem 4.2: (**Entropy inequality**) Consider the feedback system represented in Fig. 1. The following inequality holds:

$$h(\mathbf{e}^k) \geq h(\mathbf{d}^k) + I(\mathbf{x}(0); \mathbf{e}^k) \tag{19}$$

where equality holds if $\mathbf{u}$ is a function of $\mathbf{y}$ alone, or equivalently, if $\mathbf{c}$ is absent. Taking limits and dividing by k, we also have

$$h_\infty(\mathbf{e}) \geq h_\infty(\mathbf{d}) + \liminf_{k \to \infty} \frac{I(\mathbf{x}(0); \mathbf{e}^k)}{k}. \tag{20}$$

Proof: We obtain (19) by applying the positivity property of mutual information (P1) to (11). It remains to prove that equality in (19) is achieved if $\mathbf{u}$ is a function of $\mathbf{y}$ alone, or equivalently, if $\mathbf{c}$ is absent. We start by noticing that

$$\begin{aligned} I(\mathbf{u}^i&; \mathbf{e}(i)|\mathbf{e}^{i-1}, \mathbf{x}(0)) \\ &= h(\mathbf{e}(i)|\mathbf{e}^{i-1}, \mathbf{x}(0)) - h(\mathbf{e}(i)|\mathbf{e}^{i-1}, \mathbf{x}(0), \mathbf{u}^i). \end{aligned} \tag{21}$$

In addition, notice that if $\mathbf{c}$ is absent then $\mathbf{u}^k$ is a function of $\mathbf{e}^{k-1}$ (plant strictly proper) and of $\mathbf{x}(0)$, leading to

$$h(\mathbf{e}(i)|\mathbf{e}^{i-1}, \mathbf{x}(0), \mathbf{u}^i) = h(\mathbf{e}(i)|\mathbf{e}^{i-1}, \mathbf{x}(0)). \tag{22}$$

As such, we arrive at

$$\mathbf{c} \text{ is absent } \implies \forall i \in \mathbb{N}_+, I(\mathbf{u}^i; \mathbf{e}(i)|\mathbf{e}^{i-1}, \mathbf{x}(0)) = 0 \tag{23}$$

which concludes the proof. □

Notice that Theorem 4.2 did not assume stability or asymptotic stationarity. The subsequent Lemma specializes Theorem 4.2, under the assumption that the second moment of the state of the plant P is bounded. In addition, the Lemma provides bounds under various stationarity assumptions and its proof follows from (19) by means of techniques analogous to the ones used to prove Theorem 4.5 in [12].

Lemma 4.3: (**Extensions of Bode's Result**) Consider the feedback system depicted in Fig. 1, along with the following condition:

$$\sup_k \mathcal{E}[\mathbf{x}(k)^T \mathbf{x}(k)] < \infty \tag{24}$$

where $\mathbf{x}$ is the state of the plant P. The following holds:

- **Case 1** If (24) holds then $h_\infty(\mathbf{e}) \geq h_\infty(\mathbf{d}) + \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\}$
- **Case 2** If (24) holds and $\mathbf{e}$ is asymptotically stationary then the following is satisfied:

$$\frac{1}{4\pi} \int_{-\pi}^{\pi} \log(2\pi e \hat{F}_{\mathbf{e}}(\omega)) d\omega \geq h_\infty(\mathbf{d}) + \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\}. \tag{25}$$

- **Case 3** Under the conditions of Case 2 and the additional assumption that $\mathbf{d}$ is Gaussian auto-regressive and asymptotically stationary, the following is satisfied:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log(S_{\mathbf{d},\mathbf{e}}(\omega)) d\omega \geq \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\}. \tag{26}$$

Lemma 4.3, in Case 3, shows that if $\mathbf{e}$ and $\mathbf{d}$ are asymptotically stationary, with $\mathbf{d}$ Gaussian auto-regressive, then (26) holds. Notice that we require no assumptions on the probability density of $\mathbf{e}$ and that the asymptotic stationarity assumptions, on $\mathbf{d}$ and $\mathbf{e}$, are the weakest conditions under which power spectral densities are meaningful. We stress that if we want to establish a fundamental limitation based solely on the power spectral densities of $\mathbf{d}$ and $\mathbf{e}$, then we should *allow* $\mathbf{d}$ to have the *worst-case* probability distribution, which in this case is Gaussian auto-regressive. In addition, it is well known that a very large class of asymptotic power spectral densities can be generated by Gaussian auto-regressive processes. In particular, in [43] it is shown that the class of auto-regressive power spectral densities is sufficiently rich to approximate, with arbitrary accuracy, power spectral densities satisfying the Paley-Wiener condition. Therefore, (26) represents an extension of Bode's integral formula in a stochastic setting. The inequality in (26) indicates that not all of the frequency components can be attenuated, and that the *sensitivity* logarithmic integral is lower bounded by a constant, which depends on the degree of instability of the plant. This is a fundamental limit which cannot be breached by any control, including non-linear feedback over arbitrary alphabets. Another interesting aspect arises when the plant cannot be stabilized by a stable, linear and time-invariant controller [34]. Therefore, one could think that Lemma 4.3 may be conservative because the unstable poles of the controller should be present in the Bode constant given by $\sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\}$. However, the authors of [35] have shown that stable, linear and periodic controllers may be used in those cases. We should also mention that we are aware of an alternative Proof of Case 3 of Lemma 4.3, other than our information theoretic approach[5].

V. Information Rate Inequalities

Consider the feedback loop represented in Fig. 3, which is obtained by superimposing onto Fig. 1 an illustration of the fluxes of information, originating from the independent sources $\mathbf{x}(0)$ and $\mathbf{d}$. In this section, we prove that the flux of information *through* K must account for two independent sources: the plant's initial conditions $\mathbf{x}(0)$ and the external excitation $\mathbf{d}$. Section VI is an example of the applicability of such information rate inequalities, where we derive a new performance bound relating disturbance attenuation to feedback capacity.

We start with the following Lemma, where we show that the total information rate flowing in the loop is lower bounded by the additive contribution of the following information rates: (1) between the plant's initial condition $\mathbf{x}(0)$ and the error signal $\mathbf{e}$; (2) between $\mathbf{d}$ and $\mathbf{u}$.

Lemma 5.1: (**Fundamental Information Rate Inequality**) If $\mathbf{x}(k)$ is the state of the plant represented in Fig. 3, or equivalently Fig. 1, then the following holds:

$$\limsup_{k\to\infty} \frac{I((\mathbf{x}(0),\mathbf{d}^k);\mathbf{u}^k)}{k} \geq \liminf_{k\to\infty} \frac{I(\mathbf{x}(0);\mathbf{e}^k)}{k} + I_\infty(\mathbf{d};\mathbf{u}). \quad (27)$$

Intuitively, one can imagine a *cut* through K and that the left side of (27) represents the total information flow through the cut, which must be an upper bound to the individual fluxes quantified on the right hand side of (27).

Proof of Lemma 5.1: We start by using Kolomogorov's formula (P2) to write $I((\mathbf{x}(0),\mathbf{d}^k);\mathbf{u}^k) = I(\mathbf{x}(0);\mathbf{u}^k|\mathbf{d}^k) + I(\mathbf{u}^k;\mathbf{d}^k)$, which leads to the following inequality:

$$I((\mathbf{x}(0),\mathbf{d}^k);\mathbf{u}^k) \geq I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{d}^k) + I(\mathbf{u}^k;\mathbf{d}^k) \quad (28)$$

where we also used the transformation $\mathbf{e}(k) = \mathbf{d}(k) + \mathbf{u}(k)$ and the data processing property (P3) to establish that $I(\mathbf{x}(0);\mathbf{u}^k|\mathbf{d}^k) \geq I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{d}^k)$. On the other hand, using Kolmogorov's formula (P2) again, we get

$$I(\mathbf{x}(0);\mathbf{e}^k|\mathbf{d}^k) = I(\mathbf{x}(0);\mathbf{e}^k) - I(\mathbf{x}(0);\mathbf{d}^k) + I(\mathbf{x}(0);\mathbf{d}^k|\mathbf{e}^k). \quad (29)$$

Since $\mathbf{d}$ is independent from $\mathbf{x}(0)$, the second term, on the right-hand side of (29), vanishes and by substituting (29) into (28) we arrive at

$$I((\mathbf{x}(0),\mathbf{d}^k);\mathbf{u}^k) \geq I(\mathbf{x}(0);\mathbf{e}^k) + I(\mathbf{x}(0);\mathbf{d}^k|\mathbf{e}^k) + I(\mathbf{u}^k;\mathbf{d}^k). \quad (30)$$

The inequality (27) follows from (30) by the positivity of mutual information property (P1). □

[5]Prof. Alex Megretski (MIT) has recently suggested the only correct alternative proof (to the best of our knowledge) for Case 3 of Lemma 4.3, using one-step linear prediction. We should stress that our proof technique, unlike the one-step linear prediction method, can be adapted to prove the validity of Bode's integral in other configurations. In particular when the disturbance enters at the output of the plant, in which case the input to the controller should be viewed as the error signal.

Notice that Lemma 5.1 has no stability assumptions, but the following Theorem incorporates the effect of requiring that the plant P is stabilized.

Theorem 5.2: If the state of the plant, represented in Fig. 3, or equivalently Fig. 1, satisfies $\sup_{k\to\infty} \mathcal{E}[\mathbf{x}(k)\mathbf{x}(k)^T] < \infty$ then the following holds:

$$I_\infty(\mathbf{u};\mathbf{d}) \leq \limsup_{k\to\infty} \frac{I((\mathbf{x}(0),\mathbf{d}^k);\mathbf{u}^k)}{k} - \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\}. \quad (31)$$

Proof: For convenience, we provide the following abridged version of the statement in ([12], Lemma 4.1) [6].:

$$\sup_{k\to\infty} \mathcal{E}[\mathbf{x}(k)\mathbf{x}(k)^T] < \infty \Longrightarrow \liminf_{k\to\infty} \frac{I(\mathbf{x}(0);\mathbf{e}^k)}{k} \geq \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\}. \quad (32)$$

The proof follows immediately by substituting (32) into (27). □

A. Feedback in the Presence of a Communication Channel

Throughout this paper, we will adopt the following definition of channel:

Definition 5.1: (**Channel**) Let $\mathcal{V}$ be a given channel input alphabet, along with a stochastic process $\mathbf{c}$ which is assumed mutually independent of $\mathbf{x}(0)$ and $\mathbf{d}$. In addition, consider a map f acting as $\mathbf{z}(k) = f(k, \mathbf{v}^k, \mathbf{c}^k)$, where $\mathbf{z}$ and $\mathbf{v}$ represent the channel output and input processes, respectively. The pair $(f, \mathbf{c})$ defines a channel, which may be either memory-less or with memory [3].

Remark 5.1: (**Feedback Capacity**) Consider a communication channel $(f, \mathbf{c})$. Given a set of stochastic processes $\mathbb{V}$ specifying the allowable channel inputs, the channel has a well defined quantity denoted as C_f and named feedback capacity (for a definition see, for instance, ([8], page 1)). For any choice of an external stochastic process $\mathbf{w}$ independent of $\mathbf{c}$, the following inequality holds[7]

$$C_f \geq \sup_{k\geq 0} \sup_{g_e(k,\cdot,\cdot)\in \mathbb{G}_{\mathbb{V},\mathbf{w}}} \frac{I(\mathbf{w}^k;\mathbf{z}^k)}{k} \quad (33)$$

where $\mathbf{z}(k) = f(k, \mathbf{v}^k, \mathbf{c}^k)$ is the channel output, while the input to the channel is given by $\mathbf{v}(k) = g_e(k, \mathbf{w}^k, \mathbf{z}^{k-1})$ and $\mathbb{G}_{\mathbb{V},\mathbf{w}}$ is the set of maps for which the input of the channel is in the set $\mathbb{V}$. Notice that $g_e(k, \ldots, \ldots)$ should be viewed as an encoder which has access to the past outputs of the channel. Notice that the right side of (33) may not equal feedback capacity. The definition of feedback capacity presupposes that the encoder has non-causal access to the entire codeword ($\mathbf{w}$ in this case), as opposed to the causal restriction that we impose on g_e.

For the remainder of this section, we will adopt the feedback interconnection depicted in Fig. 3, where $\mathbf{z}$ is taken as the output

[6]Similar necessary conditions for stability can be found in the work by the authors of [19], [39], [47], [14]

[7]Notice that this inequality holds under information stability of the pair of sequences $(\mathbf{w}^k, \mathbf{z}^k)$. Indeed, from [16], we know that, under information stability conditions, the reliable information rate that is achievable between two stochastic processes is quantified by the information rate

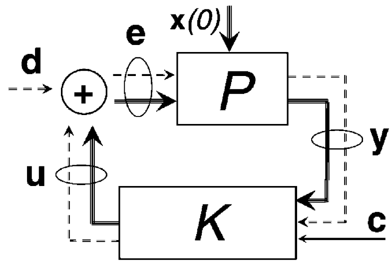

Fig. 3. Basic Feedback Paradigm, with an illustration of the fluxes of information generated by $\mathbf{x}(0)$ and $\mathbf{d}$.

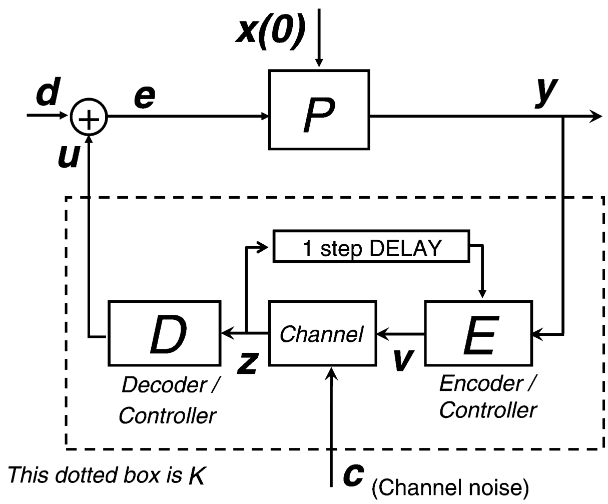

Fig. 4. Scheme representing a feedback system in the presence of a channel. The block K in Figs. 1 and 3 is represented above by the dotted box.

of a communication channel. It follows by inspection that the scheme in Fig. 4 is a particular case of Fig. 3, where the feedback block K consists of the dotted box. The encoder/controller E is a causal operator acting as $\mathbf{v}(k) = E(k, \mathbf{y}^k, \mathbf{z}^{k-1})$, while the decoder /controller operates as $\mathbf{u}(k) = D(k, \mathbf{z}^k)$. Also, notice that E has access to past channel outputs, and it may, or may not, choose to use such extra information. Still regarding the scheme of Fig. 4, it is important to notice that one can always construct a map g such that the following holds:

$$\mathbf{y}(k) = \begin{cases} g(k, \mathbf{z}^{k-1}, \mathbf{d}^k, \mathbf{x}(0)) & k \geq 1 \\ C\mathbf{x}(0) & k = 0 \end{cases}. \tag{34}$$

Remark 5.2: By substituting (34) into $\mathbf{v}(k) = E(k, \mathbf{y}^k, \mathbf{z}^{k-1})$, we can always construct a map g_{cl} leading to the following dynamic model for the channel input process $\mathbf{v}$ (see Fig. 4):

$$\mathbf{v}(k) = g_{\mathrm{cl}}(k, \mathbf{z}^{k-1}, \mathbf{d}^k, \mathbf{x}(0)). \tag{35}$$

Consider that C_f represents the feedback capacity of the channel in the scheme of Fig. 4. For the purposes of this paper, we only need to note that the following inequality holds:

$$\sup_{k \in \mathbb{N}_+} \frac{I((\mathbf{x}(0), \mathbf{d}^k); \mathbf{z}^k)}{k} \leq C_f \tag{36}$$

for all valid choices of $\mathbf{x}(0), \mathbf{d}, E$ and D, in the sense that any valid choice must guarantee that $\mathbf{v}$ belongs to the set of allowed channel input processes $\mathbb{V}$. In other words, feedback capacity is an upper-bound to the information rate between $(\mathbf{x}(0), \mathbf{d}^k)$ and $\mathbf{z}$, under the channel input constraint set $\mathbb{V}$. In order to see that (36) holds, we view the pair $(\mathbf{x}(0), \mathbf{d}^k)$ as the external process $\mathbf{w}$ (see Remark 5.1). The desired conclusion follows by noticing that g_{cl} in (35) belongs to the set $\mathbb{G}_{\mathbb{V},(\mathbf{x}(0),\mathbf{d})}$.

The computation of the feedback capacity is not easy in general. An exception is when the channel is memoryless, in which case C_f is identical to the standard Shannon capacity [3] that can be efficiently computed in most cases.

The following are examples of memory-less channels:

- **Additive white Gaussian channel**: $\mathcal{V} = \mathcal{C} = \mathbb{R}, \mathbf{c}$ is an i.i.d. white Gaussian sequence with variance σ_c^2 and $f(\mathbf{c}^k, \mathbf{v}^k) = \mathbf{c}^k + \mathbf{v}^k$. If the set $\mathbb{V}$ is specified by an input power constraint $\mathrm{Var}(\mathbf{v}(k)) \leq \sigma_v^2$ then the feedback capacity C_f is given by [3]:

$$C_f = \frac{1}{2}\log\left(1 + \frac{\sigma_v^2}{\sigma_c^2}\right)$$

- **Binary symmetric channel**: $\mathcal{V} = \mathcal{C} = \{0, 1\}, \mathbf{c}$ is an i.i.d sequence satisfying $\mathcal{P}(\mathbf{c}(k) = 1) = p_e$ and $f(\mathbf{c}^k, \mathbf{v}^k) = \mathbf{c}^k +_{\mathrm{mod}2} \mathbf{v}^k$. In this case the capacity is given by [3]:

$$C_f = -(p_e \log p_e + (1 - p_e)\log(1 - p_e))$$

- **Symmetric Quantizer** [3]: $\mathcal{V} = \mathbb{R}, \mathbf{c} = 0$ (not needed) and f is a symmetric quantizer with quantization interval δ and input amplitude constraint $\bar{v}$. This channel has capacity given by [3]

$$C_f = \log\left(2\frac{\bar{v}}{\delta}\right)$$

In view of Definition 5.1, the inequality in Theorem 5.2 can specialized according to the following Corollary:

Corollary 5.3: If the state of the plant, represented in Fig. 4, satisfies $\sup_k \mathcal{E}[\mathbf{x}(k)\mathbf{x}(k)^T] < \infty$ then the following holds:

$$I_\infty(\mathbf{u}; \mathbf{d}) \leq C_f - \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\} \tag{37}$$

where C_f is the feedback capacity of Definition 5.1.

Proof: Using the fact that $\mathbf{u}(k) = K(k, \mathbf{z}^k)$ and the data processing inequality (P3), we get

$$I((\mathbf{x}(0), \mathbf{d}^k); \mathbf{u}^k) \leq I((\mathbf{x}(0), \mathbf{d}^k); \mathbf{z}^k) \tag{38}$$

which leads to the desired result by direct substitution into (36) and (31). □

Equation (37) suggests that feedback capacity can be used to establish a universal upper-bound on $I_\infty(\mathbf{u}; \mathbf{d})$. As the feedback capacity approaches $\sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\}$ (the critical rate for the stabilization of P), the information rate $I_\infty(\mathbf{u}; \mathbf{d})$ decreases to zero. Furthermore, (37) also leads to a universal bound of performance for feedback systems, which is described in Section VI.

VI. Universal Performance Bound in the Frequency Domain

In Section V, we have proved an information rate separation principle which also leads to several upper-bounds on $I_\infty(\mathbf{d};\mathbf{u})$. Subsequently, we investigate how $I_\infty(\mathbf{d};\mathbf{u})$ may limit disturbance attenuation in the frequency domain. Throughout this section, we consider that $\mathbf{d}$ and $\mathbf{e}$ are asymptotically stationary. We also assume that $\mathbf{d}$ is Gaussian auto-regressive, but there are no assumptions on the probability density of $\mathbf{e}$.

From the upcoming Theorem 6.2, we conclude that if $I_\infty(\mathbf{d};\mathbf{u})$ is bounded then there is a universal limit on disturbance attenuation which cannot be predicted by existing results. In spite of having a similar form, the bound in Theorem 6.2 is unrelated to Bode's integral, in the sense that one does not imply the other. In Section IV, we have shown that causality is the main assumption behind our generalization of Bode's formula, while Theorem 6.2 relies on finite capacity feedback.

We start with the following Lemma, where we relate $I_\infty(\mathbf{d};\mathbf{u})$ with the ability to reject disturbances. In particular, we prove that as $I_\infty(\mathbf{d};\mathbf{u})$ approaches zero $\hat{F}_{\mathbf{e}}$ will converge to $\hat{F}_{\mathbf{d}} + \hat{F}_{\mathbf{u}}$, and that implies that disturbance attenuation is impossible over the whole frequency range. The definition below introduces a necessary concept for the statement of the aforementioned Lemma:

Definition 6.1: (**Joint Asymptotic Stationarity**) Let $\mathbf{a}$ and $\mathbf{b}$ be asymptotically stationary processes. We qualify $\mathbf{a}$ and $\mathbf{b}$ as jointly asymptotically stationary if the following limit exists for every $\gamma \in \mathbb{N}$:

$$\bar{R}_{\mathbf{b},\mathbf{a}}(-\gamma) = \bar{R}_{\mathbf{a},\mathbf{b}}(\gamma) \stackrel{\text{def}}{=} \lim_{k\to\infty} \mathcal{E}[\mathbf{a}(k+\gamma)\mathbf{b}(k)]. \tag{39}$$

We also use (39) to define the following asymptotic joint power spectral density:

$$\hat{F}_{\mathbf{a},\mathbf{b}}(\omega) = \sum_{k=-\infty}^{\infty} \bar{R}_{\mathbf{a},\mathbf{b}}(k) e^{-j\omega k}. \tag{40}$$

Lemma 6.1: Consider the scheme of Fig. 1, where $\mathbf{e}$ and $\mathbf{d}$ are assumed jointly asymptotically stationary, and $\mathbf{d}$ is also Gaussian auto-regressive. If $\log \hat{F}_{\mathbf{u}}(\omega)$ is bounded then (41), shown at the bottom of the page, holds.

Proof: In this proof, we will make use of the following non-causal, linear and time-invariant Wiener filter:

$$L^*(e^{j\omega}) = \frac{\hat{F}_{\mathbf{d},\mathbf{u}}(\omega)}{\hat{F}_{\mathbf{u}}(\omega)} \tag{42}$$

which provides the optimal estimate of $\mathbf{d}$ given $\mathbf{u}$, in the expected mean-square sense. In addition, given an integer l, we define the following auxiliary stochastic processes $\hat{\mathbf{d}}_{(W,l)}$ and $\mathbf{u}_{(W,l)}$ as:

$$\mathbf{u}_{(W,l)}(k) \stackrel{\text{def}}{=} \begin{cases} \mathbf{u}(k), & \text{if } k \le l \\ 0, & \text{otherwise} \end{cases} \tag{43}$$

$$\hat{\mathbf{d}}_{(W,l)}(k) \stackrel{\text{def}}{=} \begin{cases} (L^*\mathbf{u}_{(W,l)})(k), & \text{if } k \le l \\ 0, & \text{otherwise.} \end{cases} \tag{44}$$

We start by using the injective transformation of differential entropy property (P5) to establish the following:

$$h(\mathbf{d}^k|\mathbf{u}^k) = h\left(\tilde{\mathbf{d}}^k_{(W,k)}|\mathbf{u}^k\right) \tag{45}$$

where $\tilde{\mathbf{d}}^k_{(W,k)} \stackrel{\text{def}}{=} \mathbf{d}^k - \hat{\mathbf{d}}^k_{(W,k)}$.

Now, notice that from (2), we know that $h(\tilde{\mathbf{d}}^k_{(W,k)}) = I(\tilde{\mathbf{d}}^k_{(W,k)};\mathbf{u}^k) + h(\tilde{\mathbf{d}}^k_{(W,k)}|\mathbf{u}^k)$, which, from the positivity of mutual information property (P1), implies that $h(\tilde{\mathbf{d}}^k_{(W,k)}) \ge h(\tilde{\mathbf{d}}^k_{(W,k)}|\mathbf{u}^k)$. Therefore, (45) leads to

$$h(\mathbf{d}^k|\mathbf{u}^k) \le h\left(\tilde{\mathbf{d}}^k_{(W,k)}\right). \tag{46}$$

From (2), we know that $I(\mathbf{d}^k,\mathbf{u}^k) = h(\mathbf{d}^k) - h(\mathbf{d}^k|\mathbf{u}^k)$ so that (46) implies

$$I(\mathbf{d}^k;\mathbf{u}^k) = h(\mathbf{d}^k) - h(\mathbf{d}^k|\mathbf{u}^k) \ge h(\mathbf{d}^k) - h(\tilde{\mathbf{d}}^k_{(W,k)}). \tag{47}$$

Because L^* is a linear and time-invariant filter, even if non-causal, and since $|L^*(e^{j\omega})|$ is bounded, we conclude that L^* has a square integrable impulse response. The aforementioned properties, and the facts that $\mathbf{d}$ and $\mathbf{u}$ are assumed asymptotically stationary, allow the use of an argument similar to the one used to prove ([12], Lemma 4.3) to conclude that the following inequality holds:

$$\limsup_{k\to\infty} \frac{h(\tilde{\mathbf{d}}^k_{(W,k)})}{k} \le \frac{1}{4\pi}\int_{-\pi}^{\pi} \log\left(2\pi e \hat{F}_{\tilde{\mathbf{d}}}(\omega)\right) d\omega \tag{48}$$

where $\hat{F}_{\tilde{\mathbf{d}}} \stackrel{\text{def}}{=} \hat{F}_{\mathbf{d}}(\omega) - |L^*(e^{j\omega})|^2 \hat{F}_{\mathbf{u}}$ is the asymptotic power spectral density associated with the residual $\tilde{\mathbf{d}} = \mathbf{d} - L^*\mathbf{u}$ of the non-causal Wiener filter.

Accordingly, (47)–(48) and ([12], Lemma 4.3) can be used to infer the following:

$$I_\infty(\mathbf{d};\mathbf{u}) \ge \frac{1}{4\pi}\int_{-\pi}^{\pi} \log \times \left(\frac{\hat{F}_{\mathbf{d}}(\omega)}{\hat{F}_{\mathbf{d}}(\omega) - |L^*(e^{j\omega})|^2 \hat{F}_{\mathbf{u}}}\right) d\omega. \tag{49}$$

$$I_\infty(\mathbf{d};\mathbf{u}) \ge -\frac{1}{4\pi}\int_{-\pi}^{\pi} \log\left(1 - \frac{\left(\hat{F}_{\mathbf{e}}(\omega) - \left(\hat{F}_{\mathbf{d}}(\omega) + \hat{F}_{\mathbf{u}}(\omega)\right)\right)^2}{4\hat{F}_{\mathbf{d}}(\omega)\hat{F}_{\mathbf{u}}(\omega)}\right) d\omega \tag{41}$$

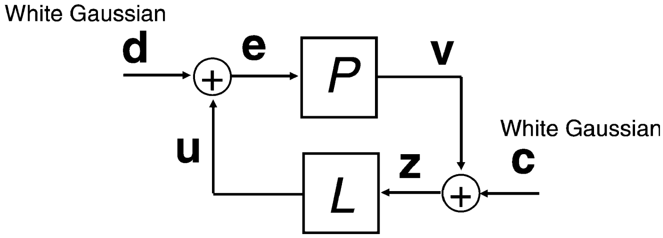

Fig. 5. Scheme representing a feedback system, in the presence of an AWGN channel and an asymptotically stationary Gaussian auto-regressive disturbance.

The proof follows by recognizing that 2 Re $\{\hat{F}_{\mathbf{d},\mathbf{u}}(\omega)\} = \hat{F}_{\mathbf{e}}(\omega) - \hat{F}_{\mathbf{d}}(\omega) - \hat{F}_{\mathbf{u}}(\omega)$, which can be used to infer the following:

$$|L^*(e^{j\omega})| \geq \frac{\hat{F}_{\mathbf{e}}(\omega) - \hat{F}_{\mathbf{d}}(\omega) - \hat{F}_{\mathbf{u}}(\omega)}{2\hat{F}_{\mathbf{u}}(\omega)}$$

□

The following Theorem provides a bound, which follows from (41) and does not depend on $\hat{F}_{\mathbf{u}}$.

Theorem 6.2: Consider the scheme of Fig. 1, where $\mathbf{e}$ and $\mathbf{d}$ are assumed jointly asymptotically stationary, and $\mathbf{d}$ is also Gaussian auto-regressive. The following holds:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \min\{0, \log S_{\mathbf{d},\mathbf{e}}(\omega)\} d\omega \geq -I_{\infty}(\mathbf{d};\mathbf{u}) \qquad (50)$$

where $S_{\mathbf{d},\mathbf{e}}(\omega)$ comes from Definition 3.2, i.e., $S_{\mathbf{d},\mathbf{e}}(\omega) = \sqrt{(\hat{F}_{\mathbf{e}}(\omega))/(\hat{F}_{\mathbf{d}}(\omega))}$.

Proof: The proof follows by infimizing the right-hand-side of (41), with respect to $\hat{F}_{\mathbf{u}}$ restricted to $\inf_{\omega \in (-\pi,\pi]} \hat{F}_{\mathbf{u}}(\omega) > 0$. An alternative proof of this Lemma for the non-asymptotically-stationary case may be found in [13]. □

A. Numerical Example

For the remainder of this subsection, consider the diagram of Fig. 5 with $\sigma_d = 1$, where P is a strictly proper plant and L is a linear and time invariant system (controller) with zero initial conditions. In particular, assume that P is a single-input single-output, strictly proper and minimum phase system of the form

$$P(z) = z^{-1}\phi(z)\frac{1}{\prod_{i=1}^{n_u}(1 - p_i z^{-1})} \qquad (51)$$

where $\phi(z)$ is an outer transfer function, n_u is the number of unstable poles of P and p_i represent such unstable poles satisfying $|p_i| > 1$. In addition, consider the following stabilizing dead-beat controller:

$$L(z) = z - \frac{1}{P(z)}. \qquad (52)$$

We don't know how conservative Theorem 6.2 is in general, but we will construct examples that illustrate that the bound can be arbitrarily tight. An example of application of such a Theorem is given in [38].

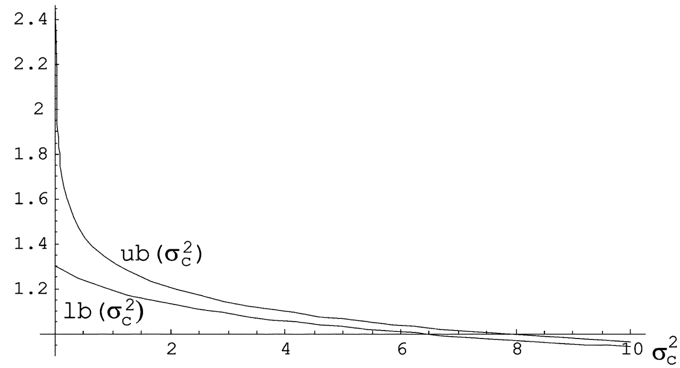

Fig. 6. Plot of the upper-bound and lower-bound, computed as a function of σ_c^2.

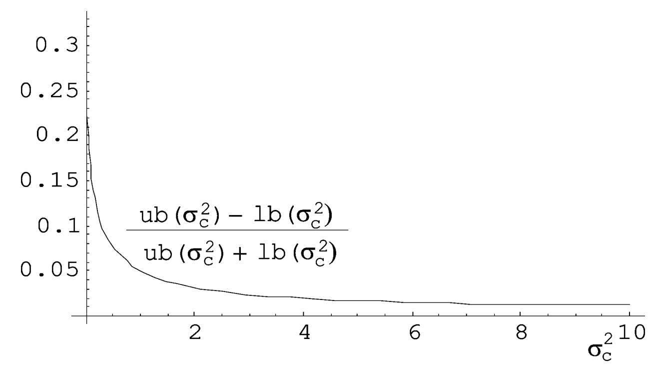

Fig. 7. Plot of the relative difference between the upper-bound and lower-bound, computed as a function of σ_c^2.

Under the assumption that the feedback loop is stable, we can use Lemma A 1, of the Appendix, to perform the following computation:

$$I_{\infty}(\mathbf{d};\mathbf{u}) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \log\left(1 + \frac{|P(e^{j\omega})|^2}{\sigma_c^2}\right) d\omega. \qquad (53)$$

From (53) we infer that inequality (50) will be tight if the following lower-bound and upper-bound are close:

$$\text{ub}\left(\sigma_c^2\right) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \log\left(1 + \frac{|P(e^{j\omega})|^2}{\sigma_c^2}\right) d\omega \qquad (54)$$

$$lb\left(\sigma_c^2\right) = -\frac{1}{4\pi}\int_{-\pi}^{\pi} \min\{0, \log S_{\mathbf{d},\mathbf{e}}(\omega)\} d\omega \qquad (55)$$

where $S_{\mathbf{d},\mathbf{e}}(\omega) = (\sqrt{1 + \sigma_c^2 |L(e^{j\omega})|^2})/(|1 - P(e^{j\omega})L(e^{j\omega})|)$.

In Fig. 6, we depict the numerical results for the following P:

$$P(z) = \frac{z^{-1}}{(1 - 1.5z^{-1})^{10}}. \qquad (56)$$

By inspection, one can argue that the bounds get more accurate for increasing values of σ_c^2 (see Fig. 7). Moreover, we have verified empirically that such relative accuracy can be made arbitrarily small by considering $P(z) = (z^{-1})/((1 - 1.5z^{-1})^n)$, with n arbitrarily large. We emphasize that the choice of the multiple pole of $P(z)$ was arbitrary. We have tried other values and the bounds behaved in a similar way (see Fig. 8).

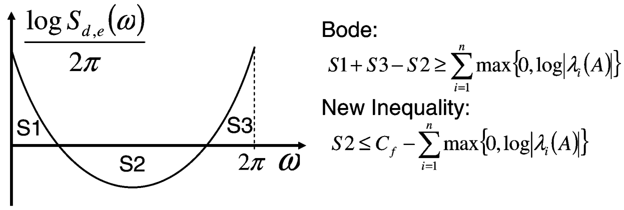

Fig. 8. Pictorial comparison between Bode's inequality and the new bound on disturbance attenuation, where $S1$ and $S3$ represent areas under the graph, while $S2$ is the area above the graph.

B. Finite Feedback Capacity: A Universal Bound on Disturbance Attenuation

Going back to the scheme of Fig. 4, the following Theorem characterizes a universal bound on disturbance attenuation in the presence of communication constraints.

Theorem 6.3: Consider the scheme of Fig. 4, where $\mathbf{e}$ and $\mathbf{d}$ are assumed jointly asymptotically stationary, with $\mathbf{d}$ Gaussian auto-regressive. If the state of the plant satisfies $\sup_k \mathcal{E}[\mathbf{x}(k)\mathbf{x}(k)^T] < \infty$ then the following holds:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \min\{0, \log S_{\mathbf{d},\mathbf{e}}(\omega)\} d\omega \geq \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\} - C_f \quad (57)$$

where C_f is the feedback capacity and A is the dynamic matrix of the plant.

Proof: The proof follows from Corollary 5.3 along with Theorem 6.2. □

We stress that the bound in (57) is valid for any channel and it depends only on the feedback capacity and on the unstable eigenvalues of A. This inequality could not be predicted from Bode's result nor from previous results.

We should also mention that there is a *good* intuitive reason why the positive part of the log-sensitivity integral is not present in (57). For a given $\mathbf{d}$, assume that $\mathbf{u} = \alpha\mathbf{u}_\perp + \mathbf{u}_d$, where $\mathbf{u}_\perp$ is some exogenous stochastic process, which is independent of $\mathbf{d}$. Clearly, increasing α does not increase $I_\infty(\mathbf{u},\mathbf{d})$ and, on the other hand, for each frequency ω, the power spectral density $F_{\mathbf{e}}(\omega)$ is an increasing function of α^2, with derivative given by $F_{\mathbf{u}_\perp}(\omega)d\omega$.

VII. Conclusion

By using notions, from Information Theory, such as mutual information and (differential) entropy, we have characterized conservation laws that hold under causality, which is a basic attribute of physical systems. In particular, we show that the differential entropy, induced by external excitation, cannot be reduced by causal feedback. This principle is related to the Bode integral formula, originally derived for linear and time-invariant feedback systems. The aforementioned analysis extends Bode's ideas to arbitrary feedback. In addition, we deduce information flow inequalities that can be used for establishing a universal bound of performance, in the frequency domain.

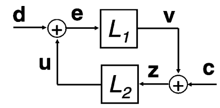

Fig. 9. Linear and time-invariant feedback scheme for the computation of information rates.

Appendix I
Auxiliary Results

Lemma A.1: Let $\mathbf{d}$ and $\mathbf{c}$ be white Gaussian sequences with positive variances σ_c^2 and σ_d^2. Consider the feedback loop of Fig. 9, where L_1 and L_2 represent linear and time-invariant systems, where L_2 is assumed stable. If the feedback loop is stable then the following holds:

$$I_\infty(\mathbf{d};\mathbf{u}) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \log\left(1 + \frac{|L_1(e^{j\omega})|^2\sigma_d^2}{\sigma_c^2}\right) d\omega. \quad (58)$$

Proof: We leave to the reader the detailed proof of the following equality:

$$I_\infty(\mathbf{u};\mathbf{d}) = \lim_{k\to\infty} \frac{I\left(\mathbf{d}_k^{k^2};\mathbf{u}_k^{k^2}\right)}{k^2} = \bar{I}(\mathbf{u};\mathbf{d}) \quad (59)$$

where $\bar{I}(\mathbf{u};\mathbf{d})$ is the stationary information rate, i.e., the information rate which we would get if the probability of the overall initial state, of the feedback loop in Fig. 9, was the stationary solution. The proof of (59) follows by using the fact that the elements of the covariance matrix of $(\mathbf{d}_k^{k^2}, \mathbf{u}_k^{k^2})$ converge uniformly and with exponential rate to the stationary solution. The convergence is exponential in k because the feedback loop is stable, linear and time-invariant. Indeed, (59) follows by noticing that, for the Gaussian case, the mutual information between $\mathbf{d}_k^{k^2}$ and $\mathbf{u}_k^{k^2}$ is given by $\log(\det(\Sigma_1(k)))/(det(\Sigma_2(k))det(\Sigma_3(k)))$, where $\Sigma_1(k), \Sigma_2(k)$ and $\Sigma_3(k)$ are the covariance matrices of $(\mathbf{d}_k^{k^2}, \mathbf{u}_k^{k^2}), \mathbf{d}_k^{k^2}$ and $\mathbf{u}_k^{k^2}$, respectively. The convergence of the determinants follows by Gershgorin's Circle Theorem.

On the other hand, from Theorem 10.2.1 [16], we know that:

$$\bar{I}(\mathbf{u};\mathbf{d}) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \log\left(1 + \frac{\left|\frac{L_1(e^{j\omega})L_2(e^{j\omega})}{1-L_1(e^{j\omega})L_2(e^{j\omega})}\right|^2 \sigma_d^2}{\left|\frac{L_2(e^{j\omega})}{1-L_1(e^{j\omega})L_2(e^{j\omega})}\right|^2 \sigma_c^2}\right) d\omega \quad (60)$$

Similar computations of the information rate for the Gaussian case can be found in [5]. □

Acknowledgment

The authors wish to thank. Dr. J. Doyle (Caltech) for his insightful suggestions, J. Freudenberg (UMich), R. Middleton (U. Newcastle), C. T. Abdallah (U. New Mexico), T. Başar (UIUC), P. Kokotovic (UCSB), and P. A. Iglesias (JHU), for their comments, B. Levine (UMD), O. Ayaso (MIT), and P. Narayan (UMD), for important feedback on this paper.

REFERENCES

[1] V. S. Borkar and S. K. Mitter, "LQG control with communication constraints," in *Communications, Computation, Control and Signal Processing: A tribute to Thomas Kailath*. Norwell, MA: Kluwer, 1997.

[2] H. W. Bode, *Network Analysis and Feedback Amplifier Design*. Princeton, NJ: Van Nostrand, 1945.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[4] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback Control Theory*. New York: Macmillan, 1992.

[5] N. Elia, "When bode meets shannon: Control-oriented feedback communication schemes," *IEEE Trans. Automat. Control*, vol. 49, no. 9, pp. 1477–, Sep. 2004.

[6] J. S. Freudenberg and D. P. Looze, *Frequency Domain Properties of Scalar and Multivariable Systems*. New York: Springer, 1988.

[7] U. Grenander and G. Szego, *Toeplitz Forms and Their Applications*. Berkeley, CA: Univ. of California Press, 1958.

[8] Y.-H. Kim, "Feedback capacity of the first-order moving average gaussian channel," *IEEE Trans. Inform. Theory*, vol. 52, no. 7, pp. 3063–3079, Jul. 2006.

[9] D. Liberzon, "On stabilization of non-linear systems with limited information feedback," in *Proc. IEEE Conf. Decision Control*, 2003, pp. 182–186.

[10] N. C. Martins, *Information Theoretic Aspects of the Control and Mode Estimation of Stochastic Systems*. Cambridge, MA: MIT, Aug. 2004.

[11] N. C. Martins, M. A. Dahleh, and N. Elia, "Feedback stabilization of uncertain systems in the presence of a direct link," *IEEE Trans. Automat. Control*, vol. 51, no. 3, pp. 438–447, Mar. 2006.

[12] N. C. Martins and M. A. Dahleh, "Fundamental limitations of disturbance attenuation in the presence of side information," *IEEE Trans. Automat. Control*, vol. 52, no. 1, pp. 56–66, Jan. 2007.

[13] N. C. Martins and M. A. Dahleh, "Fundamental limitations of performance in the presence of finite capacity feedback," in *Proc. Amer. Control Conf.*, 2005, vol. 1, pp. 79–86.

[14] G. N. Nair and R. J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM J. Control Optim.*, vol. 43, no. 2, pp. 413–436, 2004.

[15] G. N. Nair, R. J. Evans, I. M. Y. Mareels, and W. Moran, "Topological entropy and nonlinear stabilization," *IEEE Trans. Automat. Control*, vol. 49, no. 9, pp. 1585–1597, Sep. 2004.

[16] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*. San Francisco, CA: Holden Day, 1964.

[17] A. Sahai, "Anytime Information Theory," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, 2001.

[18] S. Tatikonda and S. K. Mitter, "Control under Communication Constraints," *IEEE Trans. Automat. Control*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.

[19] S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE Trans. Automat. Control*, vol. 49, no. 7, pp. 1196–1201, Jul. 2004.

[20] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints—I: State estimation problems," *IEEE Trans. Automat. Control*, vol. 42, no. 9, pp. 1294–1298, Sep. 1997.

[21] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints -II: Stabilization with limited information feedback," *IEEE Trans. Automat. Control*, vol. 44, no. 5, pp. 1049–1053, May 1999.

[22] G. Zang and P. A. Iglesias, "Nonlinear extension of bode's integral based on an information theoretic interpretation," *Syst. Control Lett.*, vol. 50, pp. 11–19, 2003.

[23] P. A. Iglesias, "An analogue of bode's integral for stable non-linear systems: Relations to entropy," in *Proc. IEEE CDC'01*, 2001, pp. 3419–3420.

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

[25] M. A. Dahleh and J. S. Shamma, "Rejection of persistent bounded disturbances: Nonlinear controllers," *Syst. Control Lett.*, vol. 18, pp. 245–252, 1992.

[26] G. N. Saridis, "Entropy formulation of optimal and adaptive control," *IEEE Trans. Automat. Control*, vol. 33, no. 8, pp. 713–721, 1988.

[27] S. Sarma, M. A. Dahleh, and S. Salapaka, "On time-varing bit-allocation maintaining stability: A convex parameterization approach," in *Proc. IEEE CDC'04*, 2004, pp. 1430–1435.

[28] A. S. Matveev and A. V. Savkin, "The problem of LQG optimal control via a limited capacity communication channel," *Syst. Control Lett.*, vol. 53, pp. 51–64, 2004.

[29] M. M. Seron, J. H. Braslavsky, and G. C. Goodwin, *Fundamental Limitations in Filtering and Control*. London, U.K.: Springer, 1997.

[30] M. M. Seron, J. H. Braslavsky, P. V. Kokotovic, and D. Q. Mayne, "Feedback limitations in nonlinear systems: From bode integrals to cheap control," *IEEE Trans. Automat. Control*, vol. 44, no. 4, pp. 829–833, Apr. 1999.

[31] J. P. Schalkwijk and M. T. Kailath, "A coding scheme for additive noise channels with feedbackI: No bandwidth constraint," *IEEE Trans. Inform. Theory*, vol. IT-12, no. 2, pp. 172–182, Apr. 1966.

[32] J. P. M. Schalkwijk, "A coding scheme for additive noise channels with feedbackPart II: Bandlimited signals," *IEEE Trans. Inform. Theory*, vol. IT-12, no. 2, pp. 183189–, Apr. 1966.

[33] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, Apr. 1997.

[34] D. C. Youla, J. J. Bongiorno, and C. N. Lu, "single-loop feedback stabilization of linear multivariable plants," *Automatica*, vol. 10, pp. 159–173, 1974.

[35] P. Khargonekar, K. Poola, and A. Tannenbaum, "Robust control of linear time-invariant plants using periodic compensation," *IEEE Trans. Automat. Control*, vol. AC-13, no. 11, pp. 1088–1096, Nov. 1985.

[36] S. Mitter and N. Newton, "Information and entropy flow in the Kalman–Bucy filter," *J. Stat. Phys.*, vol. 118, pp. 145–176, 2005.

[37] J. Braslavsky, R. Middleton, and J. Freudenberg, "Feedback stabilization over signal-to-noise ratio constrained channels," in *Proc. ACC*, 2004 [Online]. Available: ftp://warhol.newcastle.edu.au/pub/Reports/

[38] G. Vinnicombe, "Feedback networks," in *Proc. Control Uncertain Syst.: Modeling Approx., Design: Workshop Occasion Keith Glover's 60th Birthday*, Cambridge, UK, Apr. 2006, pp. 371–388.

[39] S. Yuksel and T. Basar, "Quantization and coding for decentralized LTI systems," in *Proc. IEEE CDC'03*, Maui, HI, Dec. 2003, pp. 2847–2852.

[40] K. Li and J. Baillieul, "Robust quantization for digital finite communication bandwidth (DFCB) control," *IEEE Trans. Automat. Control.*, vol. 49, no. 9, pp. 1573–1584, Sep. 2004.

[41] J. Chen, L. Qiu, and O. Toker, "Limitations on maximal tracking accuracy," *IEEE Trans. Automat. Control*, vol. 45, no. 2, pp. 326–331, Feb. 2000.

[42] R. M. Gray, *Entropy and Information Theory*. New York: Springer Verlag, 1991.

[43] M. B. Priestley, *Spectral Analysis and Time Series*. New York: Academic, Jan. 28, 1983, vol. 1–2.

[44] R. Ostrovsky, Y. Rabani, and L. J. Schulman, "Error-correcting codes for automatic control," in *Proc. 46th FOCS'05*, 2005, pp. 309–316.

[45] L. J. Schulman, "Coding for interactive communication," *IEEE Trans. Inform. Theory*, vol. 42, no. 6, pp. 1745–1756, Nov. 1996.

[46] T. M. Yi, J. Goncalves, B. Ingalls, H. Sauro, and J. C. Doyle, "A Fundamental limitation on the robustness of complex systems," to be published, 2008.

[47] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication linkpart I: Scalar systems," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3369–3395, Aug. 2006.

[48] E. A. Jonckheere, A. A. Hammad, and B. F. Wu, "Chaotic disturbance rejection a Kolmogorov-Sinai entropy approach," in *Proc. IEEE CDC'93*, 1993, vol. 4, pp. 3578–3583.

[49] H. K. Sung and S. Hara, "Properties of sensitivity and complementary sensitivity functions in single input and single output digital control systems," *Int. J. Control*, vol. 48, no. 6, pp. 2429–2439, 1988.

Nuno C. Martins (M'05) received the Licenciatura and the MS. degrees in electrical engineering from Instituto Superior Tecnico, Lisbon, Portugal, in 1994 and 1997, respectively, and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 2004.

Currently, he is an Assistant Professor with the Department of Electrical and Computer Engineering, University of Maryland, College Park. He is also affiliated with the Institute for Systems Research and He is on the Editorial Board of *Systems and Control Letters*. His research interests include fundamental limits of feedback, methods for the design of optimal networked control systems and the fusion between control theory and information theory.

Dr. Martins received the National Science Foundation CAREER Award in 2007 and the 2006 American Automatic Control Council O. Hugo Schuck Award (theory) and two fellowships, in 1999 and 2004, from the European Social Fund and the Portuguese Foundation for Science and Technology.

Munther A. Dahleh (S'84–M'97–SM'97–F'00) received the B.S. degree from Texas A & M University, College Station, TX, in 1983, and the Ph.D. degree from Rice University, Houston, TX, in 1987, all in electrical engineering.

Since then, he has been with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT), Cambridge, where he is now a Full Professor. He was a Visiting Professor at the Department of Electrical Engineering, California Institute of Technology, Pasadena, in Spring 1993. He has held consulting positions with several companies in the U.S. and abroad. He is interested in problems at the interface of robust control, filtering, information theory, and computation which include control problems with communication constraints and distributed mobile agents with local decision capabilities. He is the co-author of *Control of Uncertain Systems: A Linear Programming Approach* (Englewood Cliffs, NJ: Prentice-Hall) and co-author of *Computational Methods for Controller Design* (New York: Springer). He was an Associate Editor for the *Systems and Control Letters*. He is interested in model reduction problems for discrete-alpahbet hidden markov models and universal learning approaches for systems with both continuous and discrete alphabets. He is also interested in the interface between systems theory and neurobiology, and in particular, in providing an anatomically consistent model of the motor control systems.

Dr. Dahleh received the Ralph Budd Award in 1987 for the best thesis at Rice University, the George Axelby Outstanding Paper Award in 1987, the NSF Presidential Young Investigator Award in 1991, the Finmeccanica Career Development Chair, in 1992, the Donald P. Eckman Award from the American Control Council in 1993, the Graduate Students Council Teaching Award in 1995, the George Axelby Outstanding Paper Award, in 2004, and the Hugo Schuck Award for Theory. He was a Plenary Speaker at the 1994 American Control Conference, at the Mediterranean Conference on Control and Automation in 2003, and at the MTNS in 2006. He was an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.