Fundamental Limitations of Disturbance Attenuation in the Presence of Side Information

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

Abstract—In this paper, we study fundamental limitations of disturbance attenuation of feedback systems, under the assumption that the controller has a finite horizon preview of the disturbance. In contrast with prior work, we extend Bode's integral equation for the case where the preview is made available to the controller via a general, finite capacity, communication system. Under asymptotic stationarity assumptions, our results show that the new fundamental limitation differs from Bode's only by a constant, which quantifies the information rate through the communication system. In the absence of asymptotic stationarity, we derive a universal lower bound which uses Shannon's entropy rate as a measure of performance. By means of a case-study, we show that our main bounds may be achieved.

Index Terms—Fundamental limits, information constraints, preview control.

I. INTRODUCTION

SINCE it was first published in 1945 [1], Bode's integral equation is one of the most significant results in the theory of linear feedback. If $S(z)$ is the sensitivity transfer function [5], [37] of a single-input linear feedback loop, in discrete time, then Bode's integral equation can be written as

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

where $\mathcal{UP}$ are the unstable poles of the open-loop system [5], [37], which is assumed to be rational and strictly proper. By using feedback, one would expect that disturbance rejection can be improved. On the other hand, (1) quantifies a fundamental limitation which says that disturbance rejection can be, at most, *shaped* in frequency. Equivalently, $|S(e^{j\omega})|$ cannot be made small at all frequencies. Due to its importance, Bode's fundamental limitation has been extended to more general frameworks [29] than the linear and time invariant one. The multidimensional version was provided in [9], while the time-varying

case has been addressed in [13] and certain nonlinear systems have been analyzed in [40], [12], [30]. In recent publications, such as [20] and references therein, the study of fundamental limitations generalizes to controllers with preview.

Using an information theoretic formulation, Bode-like limitations of performance were derived for feedback systems where the controller belongs to a general class [18], [19], which might include systems operating on a discrete or finite alphabet. While causality is responsible for Bode's fundamental limitation, information constraints in the feedback loop give rise to a new limitation [18], [19]. Prior results on extending Bode's result to nonlinear systems, via an information theoretic approach, can be found in [12]. Articles [14], [15] establish a relationship between Shannon's entropy rate and Bode's integral [37] for linear and time-invariant feedback systems. The work in [6] explores the connection between Bode's integral formula for linear and time-invariant systems and the ability to transmit information over a Gaussian channel, by means of linear and time invariant systems acting as encoders and decoders. Bode's fundamental limitation is derived for a deterministic setting in [38], under certain convergence conditions.

A. Main Contributions

It is well known that the use of disturbance previews may improve controller performance [36]. In [34], one finds results in optimal preview control as well as a source of references to other related approaches. Recent results on fundamental limitations of tracking in the presence of reference preview are given in [2], [20].

In this publication, we consider the diagram of Fig. 1, where the controller has access to a remotely transmitted disturbance preview, represented as **r**. This scheme portrays a formulation where the disturbance results from a physical phenomenon, which must travel in space until it reaches the system. The travel time is represented as a delay of m units of time. At the same time, a remote preview signal **r** may be available to the controller, subject to information transmission/processing constraints at the remote preview system (RPS) block. We also adopt a Markovian model for the disturbance, where G is a stable auto-regressive shaping filter and **w** is the innovations process.

This work characterizes the fundamental limits of preview control in a general remote setting (see Fig. 1). Examples of remote preview systems can be found in animal life, such as the ones that use vision and hearing to perceive a future physical interaction. In these cases, the information/processing constraints arise from limited vision and hearing resolution as well as noise
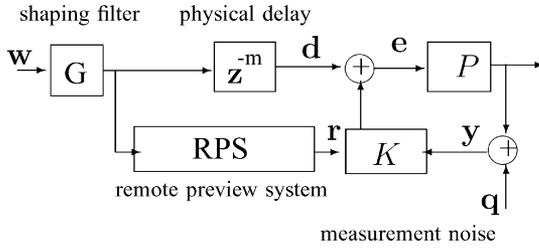
Fig. 1. Structure of a remote preview system.

and limited information processing in the brain [8]. Further examples are the information path of a heat-shock mechanism at the cellular level [7] as well as navigation engineering systems.

There are two extreme cases in this setup: the first is when the disturbance can be fully transmitted,[1] in that case the disturbance can be canceled by the controller, and the second is the absence of remote preview information, which is the classical framework. In this paper, we study the situation in between, i.e., we consider that $C > I_\infty(\mathbf{r}, \mathbf{d}) > 0$, where C is a finite positive constant, representing the Shannon capacity [3] of the RPS block, and $I_\infty(\mathbf{r}, \mathbf{d})$ is the mutual information rate[2] between the disturbance $\mathbf{d}$ and the remote preview signal $\mathbf{r}$, in bits per unit of time.

The following summarizes the contributions of this paper.

- We consider a new type of networked control system, where a preview of the disturbance is available to the controller via a general communication system. We derive an extension of Bode's integral formula for the aforementioned scheme.
- By making use of information theoretic principles, our bounds incorporate explicitly the information rate constraints at the remote preview system.
- Our derivations are valid for strictly proper linear and time-invariant plants, in the presence of arbitrary causal control, which includes time-varying, non-linear controllers operating on arbitrary alphabets.
- Our work brings a host of open problems, which are described at the end of this paper.

The paper is organized as follows. Sections I-B and C introduce the notation and the main definitions. The technical framework is given in Section II, where we also describe the measures of performance adopted in the paper. In Section III, we derive an entropy rate inequality which holds under general assumptions. The frequency domain interpretation of the aforementioned result is developed in Section IV, where we also establish a comparison with Bode's integral. A case-study is provided in Section V, characterizing a class of plants and remote preview systems for which the frequency domain lower-bounds of Section IV can be achieved. The conclusions and a discussion of research opportunities, is provided in Section VI.

B. Notation

The following notation is adopted.

[1] This would require a RPS block with infinite capacity.

[2] This quantity is precisely defined in Section I-C.

- Whenever it is clear from the context, we refer to a sequence $\{a(k)\}_0^\infty$ of elements in $\mathbb{R}^n$ as a. A finite segment of a sequence a is indicated as $a_{k_{\min}}^{k_{\max}} \stackrel{\text{def}}{=} \{a(k)\}_{k_{\min}}^{k_{\max}}$. If $k_{\max} < k_{\min}$ then $a_{k_{\min}}^{k_{\max}} = \emptyset$.
- Random variables are represented using boldface letters, such as $\mathbf{a}$.
- If $\mathbf{a}(k)$ is a stochastic process, then we use $a(k)$ to indicate a specific realization. Similar to the convention used for sequences, we may denote $\{\mathbf{a}(k)\}_0^\infty$ just as $\mathbf{a}$ and $\{a(k)\}_0^\infty$ as a. A finite segment of a stochastic process is indicated as $\mathbf{a}_{k_{\min}}^{k_{\max}}$.
- The probability density of a random variable $\mathbf{a}$, if it exists, is denoted as p_a. The conditional probability, given $\mathbf{b}$, is indicated as $p_{a\,|\,b}$.
- The expectation operator over $\mathbf{a}$ is written as $\mathcal{E}[\mathbf{a}]$.
- The variance of a real random variable $\mathbf{a}$ is given by $\mathrm{Var}(\mathbf{a}) \stackrel{\text{def}}{=} \mathcal{E}[(\mathbf{a} - \mathcal{E}[\mathbf{a}])^2]$.
- We write $\log_2(\,\cdot\,)$ simply as $\log(\,\cdot\,)$.
- We adopt the convention $0 \log 0 = 0$.

C. Basic Definitions of Information Theory

In this section, we summarize the main definitions of Information Theory used throughout the paper. We adopt [25], as a primary reference, because it considers general probabilistic spaces in a unified framework. The definitions and properties listed in this section hold under general assumptions, we refer to [25] for further details.

Definition 1.1 (from [25, p. 9]): The mutual information $I : (\mathbf{a}; \mathbf{b}) \rightarrow \mathbf{R}_+ \bigcup \{\infty\}$, between $\mathbf{a}$ and $\mathbf{b}$, is given by

$$I(\mathbf{a}; \mathbf{b}) = \sup \sum_{ij} \mathcal{P}_{\mathbf{a},\mathbf{b}}(E_i \times F_j) \log \frac{\mathcal{P}_{\mathbf{a},\mathbf{b}}(E_i \times F_j)}{\mathcal{P}_{\mathbf{a}}(E_i)\mathcal{P}_{\mathbf{b}}(F_j)}.$$

The supremum is taken over all partitions $\{E_i\}$ of $\mathcal{A}$ and $\{F_j\}$ of $\mathcal{B}$, where $\mathcal{A}$ and $\mathcal{B}$ are the alphabets of $\mathbf{a}$ and $\mathbf{b}$.

The definition of conditional mutual information can be found in [3] or in [25, p. 37].

Notice that, in Definition 1.1, $\mathcal{A}$ and $\mathcal{B}$ may be different. Without loss of generality, we consider probability spaces which are countable or $\mathbb{R}^q$, for some q. We also define the following quantities, denoted as differential entropy and conditional differential entropy, which are useful in the computation of the mutual information $I(\cdot, \cdot)$, for certain cases relevant in this paper.

Definition 1.2: If $\mathbf{a}$ is a random variable with alphabet $\mathcal{A} = \mathbb{R}^q$ 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 a random variable with alphabet $\mathcal{B} = \mathbb{R}^{q'}$ then we define the conditional differential entropy of $\mathbf{a}$ given $\mathbf{b}$ as:

$$h(\mathbf{a}\,|\,\mathbf{b}) = h(\mathbf{a}, \mathbf{b}) - h(\mathbf{b})$$
$$= -\int_{\mathbb{R}^{q'}} \left(\int_{\mathbb{R}^q} p_{a,b}(\gamma_a, \gamma_b) \log p_{a\,|\,b}(\gamma_a, \gamma_b) d\gamma_a \right) d\gamma_b. \quad (2)$$

If $\mathcal{B}$ is countable then $h(\mathbf{a} \mid \mathbf{b})$ is defined as

$$h(\mathbf{a} \mid \mathbf{b}) = - \sum_{\gamma_b \in \mathcal{B}} \int_{\mathbb{R}^q} p_{a,b}(\gamma_a, \gamma_b) \log p_{a \mid b}(\gamma_a, \gamma_b) d\gamma_a \quad (3)$$

Likewise, the quantity $h(\mathbf{a} \mid \mathbf{b}, \mathbf{c})$ is defined by incorporating another sum if the alphabet of $\mathbf{c}$ is discrete, or an integral if the alphabet is continuous.

Using [25, Th. 2.1.2], we know that the following holds:

$$I(\mathbf{a}; \mathbf{b}) = h(\mathbf{a}) - h(\mathbf{a} \mid \mathbf{b}) \quad (4)$$

where $h(\mathbf{a})$ and $h(\mathbf{a} \mid \mathbf{b})$ are assumed well defined.

In order to simplify our notation, we also define the following quantities.

Definition 1.3 (Information Rate): Let $\mathbf{a}$ and $\mathbf{b}$ be stochastic processes. The following is the definition of (mutual) information rate[3]:

$$I_\infty(\mathbf{a}; \mathbf{b}) = \limsup_{N \to \infty} \frac{I\left(\mathbf{a}_0^{N-1}; \mathbf{b}_0^{N-1}\right)}{N}.$$

The use of information rates 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 1.4 (Entropy Rate): For a given stochastic process, we also define entropy rate as:

$$h_\infty(\mathbf{a}) = \limsup_{N \to \infty} \frac{h(\mathbf{a}_m^{N-1})}{N - m} \quad (5)$$

where m is the time delay represented in Fig. 1.

In our formulation, $\mathbf{e}_0^{m-1}$ and $\mathbf{d}_0^{m-1}$ may be deterministic or may take values on a countable set. Since the differential entropy for these variables is undefined, we choose to define entropy rate as in (5). On the other hand, for any $\mathbf{e}_0^{m-1}$ and $\mathbf{d}_0^{m-1}$, there are no technical problems in using them with mutual information or conditional entropy, such as $I((\mathbf{e}_0^{m-1}, \mathbf{b}); \mathbf{c})$ and $h(\mathbf{b} \mid \mathbf{c}, \mathbf{e}_0^{m-1})$. In the real world, all signals in Fig. 1 should have some noise, with finite differential entropy, added to them. If that was the case then we could have defined entropy rate in the standard way. We chose not to add noise everywhere, because it would complicate the paper and it would lead to the same results and conclusions.

In this paper, we will refer to channels which are stochastic operators conforming to the following definition.

Definition 1.5 (Channel): Let $\mathcal{V}$ and $\mathcal{R}$ be given input and output alphabets, along with a stochastic process, denoted as $\mathbf{c}$, with alphabet $\mathcal{C}$. In addition, consider a causal map $f : \mathcal{C}^\infty \times \mathcal{V}^\infty \to \mathcal{R}^\infty$. The pair $(f, \mathbf{c})$ defines a channel. The following are examples of channels.

- **Additive white Gaussian channel**: $\mathcal{V} = \mathcal{R} = \mathcal{C} = \mathbb{R}$, $\mathbf{c}$ is an i.i.d. white Gaussian sequence with unit variance and $f(\mathbf{c}, \mathbf{v})(k) = \mathbf{c}(k) + \mathbf{v}(k)$.

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

- **Binary symmetric channel**: $\mathcal{V} = \mathcal{R} = \mathcal{C} = \{0, 1\}$, $\mathbf{c}$ is an i.i.d sequence satisfying $\mathcal{P}(\mathbf{c}(k) = 1) = p_e$ and $f(\mathbf{c}, \mathbf{v})(k) = \mathbf{c}(k) +_{\mathrm{mod2}} \mathbf{v}(k)$

For any given channel specified by $(f, \mathbf{c})$, the supremum of the rates, at which information can be reliably transmitted from the input to the output, is a fundamental quantity denoted as capacity [3]. The formal definition of capacity, denoted by C, can be found in [3], for which the following holds:

$$\sup_{p_{\mathbf{v}}} I_\infty(\mathbf{v}; f(\mathbf{v}, \mathbf{c})) \leq C. \quad (6)$$

D. Spectral Properties of Asymptotically Stationary Stochastic Processes

We adopt the following definition of asymptotic power spectral density.

Definition 1.6: A given zero mean real stochastic process $\mathbf{a}$ is (wide-sense) asymptotically stationary if the following limit exists for every $\gamma \in \mathbb{N}$:

$$\bar{R}_{\mathbf{a}}(\gamma) \stackrel{\text{def}}{=} \lim_{k \to \infty} \mathcal{E}[(\mathbf{a}(k + \gamma) - \mathcal{E}[\mathbf{a}(k + \gamma)])(\mathbf{a}(k) - \mathcal{E}[\mathbf{a}(k)])]. \quad (7)$$

We also use (7) to define the following asymptotic power spectral density:

$$\hat{F}_{\mathbf{a}}(\omega) = \sum_{k=-\infty}^{\infty} \bar{R}_{\mathbf{a}}(k) e^{-j\omega k}. \quad (8)$$

II. TECHNICAL FRAMEWORK AND ASSUMPTIONS

Regarding the general scheme of Fig. 1, the following assumptions are made.

- $\mathbf{w}$ is a scalar $(w(k) \in \mathbb{R})$, unit variance, identically and independently distributed stochastic process. For each k, $\mathbf{w}(k)$ is distributed according to a density $p_{\mathbf{w}}$, satisfying $|h_\infty(\mathbf{w}(k))| < \infty$.
- G is an all-pole stable filter of the form

$$G(z) = \frac{\alpha}{1 - \sum_{m=1}^{p} a_m z^{-m}} \quad (9)$$

where $p \geq 1$, a_i and $\alpha > 0$ are given. We chose this form of G, as a way to model the disturbance $\mathbf{d}$, not only because it is convenient that G^{-1} is well defined and causal, but also because there exists a very large class of power spectral densities that can be arbitrarily well approximated by $|G(e^{j\omega})|^2$ [26]. In addition, we assume that G has zero initial conditions.
- given n, P is a single input discrete-time plant with state $x(k) \in \mathbb{R}^n$, satisfying the following state–space equation:

$$\mathbf{x}(k+1) = \begin{bmatrix} \mathbf{x}_u(k+1) \\ \mathbf{x}_s(k+1) \end{bmatrix} = \begin{bmatrix} A_u & 0 \\ 0 & A_s \end{bmatrix} \mathbf{x}(k) + \begin{bmatrix} b_u \\ b_s \end{bmatrix} \mathbf{e}(k) \quad (10)$$

$$\mathbf{y}(k) = C\mathbf{x}(k), |\lambda_i(A_u)| \geq 1, |\lambda_i(A_s)| < 1 \text{ and } k \geq 0$$

where the state–space representation is assumed minimal. The state partitions $\mathbf{x}_u$ and $\mathbf{x}_s$ represent the unstable and stable open-loop dynamics, respectively. In addition, the initial state $\mathbf{x}(0)$ is a random variable satisfying $|h(\mathbf{x}_u(0))| < \infty$.

- The RPS system includes a communication channel. In addition, $\mathbf{q}, \mathbf{w}, \mathbf{x}(0)$, and $\mathbf{c}$ are mutually independent, where $\mathbf{c}$ represents the channel noise according to Definition 1.5.

- The measurement noise $\mathbf{q}$ is such that the following holds:

$$I\left(\mathbf{x}(0); \mathbf{u}_0^{m-1}\right) < \infty$$

meaning that the controller does not have access to an *exact description* of the initial state of the plant, for $k \in \{0, \ldots, m-1\}$. Since $\mathbf{u}_0^{m-1} = \mathbf{e}_0^{m-1}$, we could have equivalently required that $I(\mathbf{x}(0); \mathbf{e}_0^{m-1}) < \infty$.

- $\mathbf{e}$ is a scalar ($e(k) \in \mathbb{R}$) stochastic process, for which $\mathbf{e}_{k_{\min}}^{k_{\max}}$ has a probability density function for every finite $k_{\min}, k_{\max} \geq m$.

A. Performance Measures Using Entropy Rates and Asymptotic Power Spectra

In Section III, which comprises our most general results, we characterize limits of performance by means of a lower bound to the difference $h_\infty(\mathbf{e}) - h_\infty(\mathbf{d})$. In standard texts, such as [24] and [3], the entropy rate of a given stochastic process is interpreted as a measure of randomness or power. We use entropy rate to gage performance, not only because it is technically convenient, by allowing us to derive inequalities involving the information rate at the remote preview system (RPS) block, but also because it is a fundamental quantity which can be related to other, more common, measures of performance. The following is a list containing a few inequalities, which relate the entropy rate with other performance measures[4].

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

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

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

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

1) A Brief Note on the Comparison Between Kolmogorov–Sinai's and Shannon's Notions of Entropy: In [14] and [15], the authors study the conditions under which our definition of entropy rate, which is due to Claude Shannon [31], is comparable to Kolmogorov–Sinai's entropy rate [16], [32], [23]. It is concluded that, for stochastic as well as for chaotic continuously valued processes, Kolmogorov–Sinai's and Shannon's notions of entropy rate may not be compatible. Typically, the former is used in the classification of dynamical

systems according to their *degree of chaoticity*, while the latter is *usually* defined for stochastic processes. The comparative analysis in [14], [15] is interesting and fundamental. The Kolmogorov–Sinai entropy of a linear and time invariant system can be computed from the unstable eigenvalues of the dynamic matrix.

2) Definition of a Sensitivity-Like Function:

Definition 2.1: If the stochastic process $\mathbf{e}$, represented in Fig. 1, is asymptotically stationary[5] then we define the following sensitivity-like function:

$$S_{\mathbf{d},\mathbf{e}}(\omega) \overset{\text{def}}{=} \limsup_{\sigma_q \to 0} \sqrt{\frac{\hat{F}_{\mathbf{e}}(\omega)}{\hat{F}_{\mathbf{d}}(\omega)}}. \tag{11}$$

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

- If the RPS block is absent in Fig. 1 and K is linear and time-invariant then $S_{\mathbf{d},\mathbf{e}}(\omega)$ is the absolute value of the standard sensitivity function [5], [14].

- We define $S_{\mathbf{d},\mathbf{e}}(\omega)$ in the limit, as σ_q goes to zero, because we want $S_{\mathbf{d},\mathbf{e}}(\omega)$ to capture the effect of the RPS block in disturbance attenuation. In general, the existence of measurement noise, with $\sigma_q > 0$, would add an extra term on $\hat{F}_{\mathbf{e}}(\omega)$ that could mask the *beneficial* effect of the RPS block.

- in general, if K is non-linear, or if the RPS block is present, then $S_{\mathbf{d},\mathbf{e}}$ will depend on G (see [4]). 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 show, for any given disturbance spectrum, that there are limits to attenuation and that certain spectra $\hat{F}_{\mathbf{e}}$ are not attainable.

- in our formulation, $\mathbf{d}$ is asymptotically stationary, but $\mathbf{e}$ may not be. If $\mathbf{e}$ 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 other non-stationary notions, such as wavelets or evolutionary power spectral densities. In the absence of stationarity, we resort directly to entropy rates.

III. MAIN ENTROPY RATE INEQUALITY

The central result of this section is the main entropy rate inequality stated in Lemma 3.1. The frequency domain interpretations of Lemma 3.1 are developed in Section IV. For simplicity, we consider the scheme[6] of Fig. 2. In such a case, the remote preview system is constructed by means of an arbitrary channel and a general encoder.

In the rest of the paper, we adopt the following assumptions.

- A1) E and K are causal operators defined in the appropriate spaces, i.e., the output of E must belong to the

---

[4]The first fact follow from standard results in [3]. Since the second fact does not follow immediately from [3], we also include a proof in Lemma 4.3.

[5]When we refer to $\mathbf{e}$ as asymptotically stationary, we mean that $\mathbf{e}$ is asymptotically stationary for all $\sigma_q > 0$.

[6]Although our results are valid for the general scheme of Fig. 1, for simplicity, we consider the concrete scenario depicted in Fig. 2.
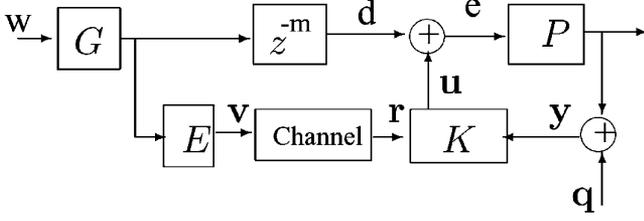
Fig. 2.   Structure of a remote preview system, using an encoder and a communication channel.

channel input alphabet, which might be discrete or continuous. Similarly, the output of the channel must be defined in the alphabet of $\mathbf{r}$, at the input of K (see Fig. 2).

- A2) (**Feedback well-posedness**) we assume that the feedback system is well-posed, i.e., that there exists a causal operator J such that the following is well defined:

$$\forall k \geq 0, \qquad \mathbf{u}(k) = J(\mathbf{x}(0), \mathbf{r}, \mathbf{d}, \mathbf{q})(k). \tag{12}$$

A. Basic Facts Involving Differential Entropy and Mutual Information

Here, we provide a list of seven properties used throughout the paper. The proofs can be found[7] in [3] or in [25], for arbitrary alphabets. For simplicity, we will refer to these properties by their number, i.e., P1)–P7).

- **P1**): $I(\mathbf{a}; \mathbf{b}) = I(\mathbf{b}; \mathbf{a}) \geq 0$ and $I(\mathbf{a}; \mathbf{b} \mid \mathbf{c}) \geq 0$.
- **P2**) Kolmogorov's formula[8] (3.6.6 in [25]):

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

- **P3**): (**data processing inequality**) [25, Th. 3.7.1]: If ϕ and θ are measurable functions in the appropriate probability spaces then $I(\phi(\mathbf{a}); \theta(\mathbf{b}) \mid \mathbf{c}) \leq I(\mathbf{a}; \mathbf{b} \mid \mathbf{c})$ and equality holds if ϕ and θ are injective[9].
- **P4**): From property P3), we conclude that $I(\mathbf{a}; (\mathbf{b}, \mathbf{c}) \mid \mathbf{d}) = I(\mathbf{a}; (\mathbf{b} - \mathbf{c}, \mathbf{c}) \mid \mathbf{d})$. Using (P2), such equality also leads to $I(\mathbf{a}; \mathbf{b} \mid (\mathbf{c}, \mathbf{d})) = I(\mathbf{a}; \mathbf{b} - \mathbf{c} \mid (\mathbf{c}, \mathbf{d}))$.
- **P5**): By means of P1) and recalling that $I(\mathbf{a}; \mathbf{b}) = h(\mathbf{a}) - h(\mathbf{a} \mid \mathbf{b})$, we infer that $h(\mathbf{a}) \geq h(\mathbf{a} \mid \mathbf{b})$, where equality holds if and only if $\mathbf{a}$ and $\mathbf{b}$ are independent. Likewise, we can use properties (P1)–(P2) to state that $I(\mathbf{a}; (\mathbf{b}, \mathbf{c})) \geq I(\mathbf{a}; \mathbf{b})$, which can be used with $I(\mathbf{a}; \mathbf{b}) = h(\mathbf{a}) - h(\mathbf{a} \mid \mathbf{b})$ to derive $h(\mathbf{a} \mid \mathbf{b}) \geq h(\mathbf{a} \mid (\mathbf{b}, \mathbf{c}))$.
- **P6**): Using a change of variables in the integrals of Definition 1.2, we reckon that if $\phi : \mathcal{B} \to \mathcal{A}$ is any given function then $h(\mathbf{a} \mid \mathbf{b}) = h(\mathbf{a} - \phi(\mathbf{b}) \mid \mathbf{b})$.
- **P7**) [3]: If $\mathbf{a}$ has a finite covariance matrix $\Sigma_a \in \mathbb{R}^{n \times n}$ then $h(\mathbf{a}) \leq (1/2) \log((2\pi e)^n \det(\Sigma_a))$.

---

[7]In [3] most of the results involving mutual information require that all the processes are either discrete or continuous.

[8]Notice that 3.6.3 in [25] has a typographic mistake. On the left hand side of the equality, the correct is $I(\xi, \zeta)$

[9]In [25] equality is guaranteed for everywhere dense ϕ and θ. Every time we say that a function is invertible or injective in this context we are implicitly assuming that it is everywhere dense.

B. Derivation of the Main Entropy Rate Inequality

The following lemma is the central result of this section. It shows that the entropy rate of the error signal is lower bounded by the entropy rate of the disturbance plus an extra term.

Lemma 3.1 (Main Entropy Rate Inequality): Consider the feedback interconnection represented in Fig. 2. For any encoder E and controller K, satisfying assumptions A1) and A2), the following holds:

$$h_\infty(\mathbf{e}) - h_\infty(\mathbf{d}) \geq \liminf_{N \to \infty} \frac{I\left(\mathbf{x}(0); \mathbf{e}_0^{N-1}\right)}{N} - I_\infty(\mathbf{r}; \mathbf{d}). \tag{13}$$

Remark 3.1: (Interpretation of (13)): The right-hand side of (13) consists of two terms: the first quantifies the information flow from the initial condition $\mathbf{x}(0)$ to the input of the plant $\mathbf{e}$, while the second is the information rate from the disturbance $\mathbf{d}$ to the controller preview input $\mathbf{r}$ (see Fig. 2). Notice that if the remote preview system is absent then the second term on the right hand side of (13) is zero while, by the non-negativity property of mutual information P1), the first term is always non-negative, implying that the entropy rate of $\mathbf{e}$ is no less than the entropy rate of $\mathbf{d}$. This fact holds for any controller K satisfying assumptions A1) and A2), and it indicates that entropy rate reduction can be achieved only in the presence of a remote preview system. Important frequency domain interpretations to Lemma 3.1 can be found in Section IV.

Proof of Lemma 3.1: We start by choosing arbitrary $k \geq m$ and using the fact that G has zero initial conditions to write:

$$h_\infty(\mathbf{d}) = h\left(\mathbf{d}(k) \mid \mathbf{d}_0^{k-1}\right) \tag{14}$$

By means of (P2), (14) and using $I(\mathbf{a}; \mathbf{b}) = h(\mathbf{a}) - h(\mathbf{a} \mid \mathbf{b})$, we obtain

$$h_\infty(\mathbf{d}) = h\left(\mathbf{d}(k) \mid \left(\mathbf{d}_0^{k-1}, \mathbf{x}(0), \mathbf{u}_0^k, \mathbf{q}_0^k\right)\right)$$
$$+ I\left(\left(\mathbf{x}(0), \mathbf{u}_0^k, \mathbf{q}_0^k\right); \mathbf{d}(k) \mid \mathbf{d}_0^{k-1}\right). \tag{15}$$

We proceed by re-writing each of the terms in the right-hand side of (15). By means of P6) and using the fact that $e(k) = d(k) + u(k)$, we find that

$$h\left(\mathbf{d}(k) \mid \left(\mathbf{d}_0^{k-1}, \mathbf{x}(0), \mathbf{u}_0^k, \mathbf{q}_0^k\right)\right)$$
$$= h\left(\mathbf{e}(k) \mid \left(\mathbf{d}_0^{k-1}, \mathbf{x}(0), \mathbf{e}_0^{k-1}, \mathbf{u}(k), \mathbf{q}_0^k\right)\right)$$
$$\underset{(\text{P5})}{\leq} h\left(\mathbf{e}(k) \mid \left(\mathbf{e}_0^{k-1}, \mathbf{x}(0)\right)\right)$$
$$\underset{(4) \text{ and (P2)}}{=} h\left(\mathbf{e}(k) \mid \mathbf{e}_0^{k-1}\right) - I\left(\mathbf{e}(k); \mathbf{x}(0) \mid \mathbf{e}_0^{k-1}\right). \tag{16}$$

Using the well-posedness assumption A2), together with P2) and P3), we can bound the second term on the right-hand side of (15) by means of the following inequality:

$$I\left(\left(\mathbf{x}(0), \mathbf{u}_0^k, \mathbf{q}_0^k\right); \mathbf{d}(k) \mid \mathbf{d}_0^{k-1}\right)$$
$$\leq I\left(\left(\mathbf{x}(0), \mathbf{r}_0^k, \mathbf{q}_0^k\right); \mathbf{d}(k) \mid \mathbf{d}_0^{k-1}\right) \tag{17}$$

which, since $\mathbf{x}(0), \mathbf{q}_0^k$ and $(\mathbf{r}_0^k, \mathbf{d}_0^k)$ are mutually independent, can be expressed as:

$$I\left(\left(\mathbf{x}(0), \mathbf{u}_0^k, \mathbf{q}_0^k\right); \mathbf{d}(k) \mid \mathbf{d}_0^{k-1}\right) \leq I\left(\mathbf{r}_0^k; \mathbf{d}(k) \mid \mathbf{d}_0^{k-1}\right). \tag{18}$$

By direct substitution of (18) and (16) into (15), we arrive at

$$h_\infty(\mathbf{d}) \le h\left(\mathbf{e}(k)\,\big|\,\mathbf{e}_0^{k-1}\right) - I\left(\mathbf{e}(k);\mathbf{x}(0)\,\big|\,\mathbf{e}_0^{k-1}\right)$$
$$+ I\left(\mathbf{r}_0^k;\mathbf{d}(k)\,\big|\,\mathbf{d}_0^{k-1}\right). \quad (19)$$

Now, choose arbitrary $N > m$ so that we have the following inequality based on (19):

$$h_\infty(\mathbf{d}) \le \frac{1}{N-m} \sum_{k=m}^{N-1} I\left(\mathbf{r}_0^k;\mathbf{d}(k)\,\big|\,\mathbf{d}_0^{k-1}\right)$$
$$+ \frac{1}{N-m}\left(\sum_{k=m}^{N-1} h\left(\mathbf{e}(k)\,\big|\,\mathbf{e}_0^{k-1}\right)\right.$$
$$\left. - \sum_{k=m}^{N-1} I\left(\mathbf{e}(k);\mathbf{x}(0)\,\big|\,\mathbf{e}_0^{k-1}\right)\right). \quad (20)$$

From P3) we know that for any $k \le N-1$ the following holds:

$$I\left(\mathbf{r}_0^k;\mathbf{d}(k)\,\big|\,\mathbf{d}_0^{k-1}\right) \le I\left(\mathbf{r}_0^{N-1};\mathbf{d}(k)\,\big|\,\mathbf{d}_0^{k-1}\right)$$

which, together with P2), P5), and (20), leads to

$$h_\infty(\mathbf{d}) \le \frac{1}{N-m}\left(h\left(\mathbf{e}_m^{N-1}\right) - I\left(\mathbf{e}_0^{N-1};\mathbf{x}(0)\right)\right)$$
$$+ \frac{1}{N-m}\left(I\left(\mathbf{r}_0^{N-1};\mathbf{d}_0^{N-1}\right) + I\left(\mathbf{e}_0^{m-1};\mathbf{x}(0)\right)\right). \quad (21)$$

Using the fact that $I(\mathbf{e}_0^{m-1};\mathbf{x}(0))$ is finite, we conclude the proof by considering the limit as $N \to \infty$ in (21). $\qquad\square$

C. Further Comments and Comparison With Prior Work

In [14] and [15], it is shown that for the standard configuration (no remote preview), with linear and time-invariant strictly proper loop gain, the entropy rate of the error signal is always lower bounded by the entropy rate of the disturbance. In the same publications, the authors relate such a property to Bode's integral [37], for linear and time-invariant systems. In the absence of a remote preview system, Lemma 3.1 is a natural extension of the result in [14], [15]. Indeed, Lemma 3.1 holds even if we consider arbitrary nonlinear *loop gains* with a delay of at least one time unit. In particular, inspection of the proof of Lemma 3.1 shows that the validity of the result does not depend on the linearity of the plant. In Section IV, we make use of the plant linearity, while requiring stability of the feedback loop and asymptotic stationarity of the error signal as a way to provide an extension of Bode's integral. However, even in Section IV, the controller is not required to be linear.

IV. Entropy Rate Inequalities Under Stability and Asymptotic Stationarity Assumptions

In this section, we specialize and interpret the main entropy rate inequality of Lemma 3.1 under stability and asymptotic stationarity assumptions. The main results of this section are Theorems 4.2 and 4.5.

A. Incorporating Stability

According to the following Lemma, stability suffices to derive a lower bound to the information rate from the initial state $\mathbf{x}(0)$ to the error signal $\mathbf{e}$. Such result follows from [35], [39], [22], [27] and it precisely characterizes the minimum information rate, between the initial condition and the error signal $\mathbf{e}$, which is necessary for stabilization. The aforementioned references address the problem of control under finite-rate constraints, but their proofs regarding the minimum stabilizing rate hold in general. An example of a control design method for such a class of problems is given in [28].

Lemma 4.1: Let $\mathbf{x}(k)$ be the solution of the state-space (10). If the plant is interconnected with a stabilizing feedback scheme, leading to mean-square stability, i.e., $\sup_k \mathcal{E}[(\mathbf{x}(k))^T\mathbf{x}(k)] < \infty$, then the following is satisfied:

$$\liminf_{N\to\infty} \frac{I\left(\mathbf{e}_0^{N-1};\mathbf{x}(0)\right)}{N} \ge \sum_{i=1}^n \max\{0, \log|\lambda_i(A)|\}. \quad (22)$$

Proof: If $A = A_s$ then we just use $I(\mathbf{e}_0^{N-1};\mathbf{x}(0)) \ge 0$. If $A \ne A_s$ then we consider the following homogeneous system:

$$\mathbf{x}_e(k+1) = A_u\mathbf{x}_e(k) + b_u\mathbf{e}(k) \quad x_e(0) = 0 \quad (23)$$

and define the estimate $\hat{\mathbf{x}}(k) = -A_u^{-k}\mathbf{x}_e(k)$. Since $\mathbf{x}_u(k) = A_u^k\mathbf{x}_u(0) + \mathbf{x}_e(k) = A_u^k(\mathbf{x}_u(0) - \hat{\mathbf{x}}(k))$, we know that

$$k\log\left|\det\left(A_u A_u^T\right)\right| + \log\det(R_{\mathbf{x}_{\text{error}}}(k))$$
$$= \log\det(R_{\mathbf{x}_u}(k)) < \beta < \infty \quad (24)$$

where $\mathbf{x}_{\text{error}}(k) = \hat{\mathbf{x}}(k) - \mathbf{x}_u(0)$, while $R_{\mathbf{x}_{\text{error}}}(k)$ and $R_{\mathbf{x}_u}(k)$ represent the covariance of $\mathbf{x}_{\text{error}}(k)$ and $\mathbf{x}_u(k)$, respectively. Since $\hat{\mathbf{x}}(k)$ is a function of $\mathbf{e}_0^k$, we use (P4) to arrive at:

$$I\left(\mathbf{x}(0);\mathbf{e}_0^{N-1}\right)$$
$$\ge I\left(\mathbf{x}_u(0);\mathbf{e}_0^{N-1}\right)$$
$$\ge h(\mathbf{x}_u(0)) - h(\hat{\mathbf{x}}(N-1) - \mathbf{x}_u(0)). \quad (25)$$

However, from P7) we know that

$$\limsup_{N\to\infty} \frac{h(\hat{\mathbf{x}}(N-1) - \mathbf{x}_u(0))}{N}$$
$$\le \limsup_{N\to\infty} \frac{\log\det(R_{\mathbf{x}_{\text{error}}}(N-1))}{2N}. \quad (26)$$

As a consequence, we can use (24) to get $\limsup_{N\to\infty}(h(\hat{\mathbf{x}}(N-1) - \mathbf{x}_u(0))/N) \le -\log|\det(A_u)|$. The proof follows by direct substitution. $\qquad\square$

B. Derivation of a General Bound Involving Entropy Rates

As we have discussed in Section II-A, we use $h_\infty(\mathbf{e}) - h_\infty(\mathbf{d})$ as a performance measure for the most general case, where we do not require $\mathbf{e}$ to be asymptotically stationary. The following Theorem provides a universal lower bound for $h_\infty(\mathbf{e}) - h_\infty(\mathbf{d})$ as a function of the unstable poles of P and the capacity of the remote preview channel.

Theorem 4.2: Consider the feedback interconnection represented in Fig. 2. In addition, assume that the state of P satisfies $\sup_k \mathcal{E}[(\mathbf{x}(k))^T \mathbf{x}(k)] < \infty$. For any encoder E and controller K, satisfying A1) and A2), the following is true:

$$h_\infty(\mathbf{e}) - h_\infty(\mathbf{d}) \geq \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\} - C \quad (27)$$

where C represents the capacity [3] of the remote preview channel.

Proof: By direct substitution of (22) into (13), we obtain:

$$h_\infty(\mathbf{e}) - h_\infty(\mathbf{d}) \geq \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\} - I_\infty(\mathbf{r}; \mathbf{d}) \quad (28)$$

The proof follows from the definition of channel capacity [3], i.e., $I_\infty(\mathbf{r}; \mathbf{v}) \leq C$ and by the data processing inequality (P3) we have $I_\infty(\mathbf{r}; \mathbf{d}) \leq I_\infty(\mathbf{r}; \mathbf{v}) \leq C$. □

C. Expressing Performance Limitations by Means of Asymptotic Power Spectral Densities: An Extension of Bode's Integral Formula

We start with the following Lemma, which establishes a connection between $h_\infty(\mathbf{e})$ and its asymptotic power spectral density $\hat{F}_\mathbf{e}$. At the end of this section, we state and prove Theorem 4.5, ascribing a frequency domain interpretation to Theorem 4.2.

Lemma 4.3: If $\mathbf{e}$ is an asymptotically stationary process[10] then the following holds:

$$\frac{1}{4\pi} \int_{-\pi}^{\pi} \log(2\pi e \hat{F}_\mathbf{e}(\omega)) d\omega \geq h_\infty(\mathbf{e}). \quad (29)$$

Remark 4.1: Notice that Lemma 4.1 holds under asymptotic stationarity assumptions, and that it remains valid regardless of the rate of convergence of the second moment statistics of the error signal $\mathbf{e}$. Such level of generality will carry on to Theorem 4.5. In general, feedback systems will exhibit stationary behavior only if the external excitation is stationary and the overall initial state (plant and controller) is distributed according to the stationary measure. Since the latter assumption is unrealistic, we gather that the nonstationary extension is relevant.

Proof of Lemma 4.3: Let $\tilde{\mathbf{e}}$ be a zero-mean Gaussian stochastic process such that $\mathcal{E}[\tilde{\mathbf{e}}(k+\gamma)\tilde{\mathbf{e}}(k)] = \mathcal{E}[\mathbf{e}(k+\gamma)\mathbf{e}(k)]$ holds. We can write

$$h_\infty(\mathbf{e}) \leq \limsup_{N\to\infty} \frac{h\left(\tilde{\mathbf{e}}_m^{N-1}\right)}{N-m} \quad (30)$$

where we used the fact that Gaussian distributions maximize differential entropy [3], under the same covariance matrix. In addition, we have:

$$\forall \gamma \in \mathbb{N}_+, \limsup_{N\to\infty} \frac{h\left(\tilde{\mathbf{e}}_m^{N-1}\right)}{N-m}$$
$$\leq \limsup_{N\to\infty} \frac{\sum_{k=m+\gamma}^{N-1} h\left(\tilde{\mathbf{e}}(k)\left|\tilde{\mathbf{e}}_{k-\gamma}^{k-1}\right.\right)}{N-m} \quad (31)$$

[10] In more rigorous terms, we should also require that $\hat{F}_\mathbf{e}$ is Lebesgue integrable. More details can be found in [10, pp. 64–65].

Now choose $\bar{\mathbf{e}}$ as a zero mean Gaussian stationary stochastic process with auto-correlation given by $\bar{R}_\mathbf{e}(\gamma)$, the asymptotic auto-correlation of $\mathbf{e}$. We can use such limit auto-correlation and substitute (31) in (30) to obtain:

$$\forall \gamma \in \mathbb{N}_+, h_\infty(\mathbf{e}) \leq h\left(\bar{\mathbf{e}}(\gamma)\left|\bar{\mathbf{e}}_0^{\gamma-1}\right.\right) \quad (32)$$

as well as the following limit:

$$h_\infty(\mathbf{e}) \leq \lim_{\gamma\to\infty} h\left(\bar{\mathbf{e}}(\gamma)\left|\bar{\mathbf{e}}_0^{\gamma-1}\right.\right) \quad (33)$$

On the other hand, we know that [3]:

$$\lim_{\gamma\to\infty} h\left(\bar{\mathbf{e}}(\gamma)\left|\bar{\mathbf{e}}_0^{\gamma-1}\right.\right) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log(2\pi e \hat{F}_\mathbf{e}(\omega)) d\omega. \quad (34)$$

□

By means of Lemma 4.3 and Theorem 4.2, we arrive at the following lemma.

Lemma 4.4: Consider the feedback interconnection represented in Fig. 2. In addition, assume that the state of P satisfies $\sup_k \mathcal{E}[\mathbf{x}^T(k)\mathbf{x}(k)] < \infty$. For any encoder E and controller K, satisfying A1) and A2) and such that $\mathbf{e}$ is asymptotically stationary, the following is true:

$$\frac{1}{4\pi} \int_{-\pi}^{\pi} \log(2\pi e \hat{F}_\mathbf{e}(\omega)) d\omega$$
$$\geq \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\} - C + h_\infty(\mathbf{d}) \quad (35)$$

where C represents the capacity of the RPS channel.

We can use Lemma 4.4 to state the following theorem, which provides a fundamental limitation in terms of $S_{\mathbf{d},\mathbf{e}}$.

Theorem 4.5: Consider the feedback interconnection represented in Fig. 2. In addition, assume that the state of P satisfies $\sup_k \mathcal{E}[\mathbf{x}^T(k)\mathbf{x}(k)] < \infty$. If the encoder E and the controller K are such that A1) and A2) are satisfied and $\mathbf{e}$ is asymptotically stationary then the following is true:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log(\sqrt{2\pi e} S_{\mathbf{d},\mathbf{e}}(\omega)) d\omega$$
$$\geq \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\} - C + h_\infty(\mathbf{w}) \quad (36)$$

where C represents the capacity [3] of the RPS channel. In addition, if $\mathbf{w}$ is Gaussian then (36) is given by

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log S_{\mathbf{d},\mathbf{e}}(\omega) d\omega \geq \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\} - C. \quad (37)$$

Remark 4.2: (Interpretation of (37)): By inspection, we infer that if the capacity of the RPS block is zero ($C = 0$) then (37) is comparable to Bode's classic result. If the capacity of the RPS block is nonzero then C quantifies a relaxation of Bode's original result. In Section V, we provide a case study where the lower bound in (37) is attained.

Proof of Theorem 4.5: Since $\mathbf{w}$ is i.i.d., we can write the following equality:

$$h_\infty(\mathbf{d}) = \lim_{k \to \infty} h\left(\mathbf{d}(k)\,\big|\,\mathbf{d}_m^{k-1}\right) = \log(\alpha) + h_\infty(\mathbf{w}) \quad (38)$$

where we used the facts that G has zero initial conditions and that, in the time-domain, G is represented as

$$\mathbf{d}(k) = \alpha\mathbf{w}(k) + \sum_{i=1}^{p} a_i \mathbf{d}(k-i)$$

Using the fact that $\int_{-\pi}^{\pi} \log|G(e^{j\omega})|d\omega = 2\pi\alpha$, we know that (38) can be rewritten as

$$h_\infty(\mathbf{d}) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \log|G(e^{j\omega})|^2 d\omega + h_\infty(\mathbf{w}). \quad (39)$$

Since $\mathbf{d}$ is asymptotically stationary and $\mathbf{w}(k)$ is unit variance, we also have that $\hat{F}_\mathbf{d}(\omega) = |G(e^{j\omega})|^2$ and (39) can be written as:

$$h_\infty(\mathbf{d}) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \log \hat{F}_\mathbf{d}(\omega)\, d\omega + h_\infty(\mathbf{w}) \quad (40)$$

From Lemma 4.4 and (40), we arrive at:

$$\forall \sigma_q > 0 \quad \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left(\sqrt{2\pi e \frac{\hat{F}_\mathbf{e}(\omega)}{\hat{F}_\mathbf{d}(\omega)}}\right) d\omega$$
$$\geq \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\} - C + h_\infty(\mathbf{w}) \quad (41)$$

After taking limits, and from the definition of $S_{\mathbf{d},\mathbf{e}}(\omega)$, we get

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log(\sqrt{2\pi e}\, S_{\mathbf{d},\mathbf{e}}(\omega))\, d\omega$$
$$\geq \limsup_{\sigma_q \to 0} \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left(\sqrt{2\pi e \frac{\hat{F}_\mathbf{e}(\omega)}{\hat{F}_\mathbf{d}(\omega)}}\right) d.\omega \quad (42)$$

The proof of (36) follows from (42) and (41). If $\mathbf{w}$ is Gaussian then $h_\infty(\mathbf{w}) = (1/2)\log(2\pi e)$ and (37) follows from (36) by direct substitution. $\square$

V. CASE STUDY: THE ADDITIVE WHITE GAUSSIAN CHANNEL CASE

In this section, we consider a particular case of the scheme of Fig. 2, where the RPS block is constructed by means of an encoder and an additive white Gaussian channel. We will show that if the plant is stabilizable by a stable, linear and time-invariant controller then the lower bound expressed in Theorem 4.5 is attainable.

A. Technical Framework for the Case-Study

For the purposes of this example, we consider the following.
- $\mathbf{w}(k)$ is unit variance, zero mean, white and Gaussian.
- P is stabilizable by means of a stable, linear and time invariant controller $H(z)$.
- $\mathbf{q}$ is i.i.d. zero mean and Gaussian with variance $\sigma_q^2 > 0$.

- The channel is specified by $\mathbf{r} = \mathbf{v} + \mathbf{c}$, where $\mathbf{c}(k)$ is white, zero mean Gaussian and the variance is $\sigma_c^2 > 0$. The channel has an input power constraint given by

$$\mathrm{Var}(\mathbf{v}(k)) \leq \bar{\sigma}_v^2, \qquad k \geq 0 \quad (43)$$

where $\bar{\sigma}_v^2$ is a pre-specified parameter. The capacity of such a Gaussian channel [3] is given by

$$C_G = \frac{1}{2}\log\left(1 + \frac{\bar{\sigma}_v^2}{\sigma_c^2}\right). \quad (44)$$

B. Achieving Equality of (37)

Consider the following selection[11] of E and K:

$$E = G^{-1}\bar{\sigma}_v \quad (45)$$
$$K(\mathbf{r}, \mathbf{y}) = -\left(z^{-m}\frac{\bar{\sigma}_v}{\bar{\sigma}_v^2 + \sigma_c^2}\right)G\mathbf{r} + H\mathbf{y} \quad (46)$$

where H is a stable linear and time-invariant stabilizing controller.

Referring to Fig. 2 and by adopting (45)–(46), standard computations lead to the following asymptotic power spectral density:

$$\hat{F}_\mathbf{e}(\omega) = \frac{\frac{\hat{F}_\mathbf{d}(\omega)}{\left(1 + \frac{\bar{\sigma}_v^2}{\sigma_c^2}\right)} + |H(e^{j\omega})|^2 \sigma_q^2}{|1 - P(e^{j\omega})H(e^{j\omega})|^2} \quad (47)$$

Using Definition 2.1 and (47), we arrive at the following:

$$S_{\mathbf{d},\mathbf{e}}(\omega) = \limsup_{\sigma_q \to 0}\sqrt{\frac{\hat{F}_\mathbf{e}(\omega)}{\hat{F}_d(\omega)}} = \frac{\left(\frac{1}{1 + \frac{\bar{\sigma}_v^2}{\sigma_c^2}}\right)^{\frac{1}{2}}}{|1 - P(e^{j\omega})H(e^{j\omega})|}. \quad (48)$$

By making use of (48), we can compute the following integral:

$$\int_{-\pi}^{\pi} \log S_{\mathbf{d},\mathbf{e}}(\omega)d\omega$$
$$= \int_{-\pi}^{\pi} \log\left(\frac{1}{|1 - P(e^{j\omega})H(e^{j\omega})|}\right)d\omega - 2\pi C_G. \quad (49)$$

By means of Bode's equality [5], [37] applied to the right hand side of (49), we can write:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log S_{\mathbf{d},\mathbf{e}}(\omega)d\omega = \sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\} - C_G. \quad (50)$$

We conclude that (50) attains the lower bound in Theorem 4.5.

C. Limitations in the Mean Square Sense

So far, we have derived fundamental limitations concerning the integral of the logarithm of $S_{\mathbf{d},\mathbf{e}}(\omega)$. In the subsequent anal-

---

[11]Notice that, from (9), G is causally invertible.

ysis, we use such results as a way to derive bounds involving mean square measures. Our inequalities relate the integral of $S_{\mathbf{d},\mathbf{e}}(\omega)^2$ with the unstable poles of the plant and $\mathcal{C}$, the capacity of the remote preview system. The interplay between logarithmic integrals in the frequency domain and various quadratic measures can be found in [21], for the standard configuration (no remote) in continuous time. The quest for universal mean square bounds in estimation over channels, as a function of the capacity, has seen recent developments in [11], where a list of relevant references can be found.

The following Proposition comprises the main tool used throughout this subsection.

Proposition 5.1: Consider the feedback interconnection represented in Fig. 2. In addition, assume that the state of P satisfies $\sup_k \mathcal{E}[\mathbf{x}^T(k)\mathbf{x}(k)] < \infty$. If the encoder E and the controller K are such that (A1)–(A2) are satisfied and $\mathbf{e}$ is asymptotically stationary then the following is true:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} 2\pi e\left(S_{\mathbf{d},\mathbf{e}}(\omega)\right)^2 d\omega$$
$$\geq 2^{2\left(\sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\} - C + h_\infty(\mathbf{w})\right)} \quad (51)$$

where C represents the capacity [3] of the remote preview channel.

Assuming that $\mathbf{w}$ is Gaussian, the following holds:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} (S_{\mathbf{d},\mathbf{e}}(\omega))^2 d\omega \geq 2^{2\left(\sum_{i=1}^{n} \max\{0, \log|\lambda_i(A)|\} - C\right)}$$
$$(52)$$

Proof: The proof follows from Jensen's inequality and Theorem 4.5. $\quad\square$

D. Example of the Achievability of the Mean Square Bounds

In this section, we construct an example where the lower-bound (52) is attained. The basic framework is described in Section V-A. In addition, consider a plant $P(z)$ given by

$$P(z) = z^{-1} \frac{\alpha(z)}{\prod_{i=1}^{l}(1 - p_i z^{-1})} \quad (53)$$

where α is a rational outer transfer function and p_i are the unstable poles of $zP(z)$.

Adopt the encoder and the controller given by (45)–(46), where the following H is selected:

$$H(z) = z\alpha(z)^{-1}\left(\prod_{i=1}^{l}(1 - p_i z^{-1})\right.$$
$$\left. - \prod_{i=1}^{l}(1 - (p_i)^{-1}z^{-1})\right). \quad (54)$$

From the Definition 2.1, we arrive at the following:

$$(S_{\mathbf{d},\mathbf{e}}(\omega))^2 = \lim_{\sigma_q \to 0} \frac{\hat{F}_{\mathbf{e}}(\omega)}{\hat{F}_d(\omega)} = \frac{\prod_{i=1}^{l}|p_i|^2}{\left(1 + \frac{\bar{\sigma}_v^2}{\sigma_c^2}\right)}. \quad (55)$$

Consequently, we can form the following integral:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} (S_{\mathbf{d},\mathbf{e}}(\omega))^2 d\omega = \frac{\prod_{i=1}^{l}|p_i|^2}{\left(1 + \frac{\bar{\sigma}_v^2}{\sigma_q^2}\right)}$$
$$= 2^{2\left(\sum_{i=1}^{l} \log|p_i| - C_G\right)} \quad (56)$$

Once we realize that p_i are the unstable poles of P we may also conclude that (56) implies achieving equality in (52).

VI. CONCLUSION, UNSOLVED PROBLEMS, AND OPPORTUNITIES FOR FUTURE RESEARCH

In this paper, we have derived disturbance attenuation bounds for a networked control scheme, where a finite horizon preview of the disturbance is remotely available. In our results, the capacity of the preview communication system is an essential quantity. Under asymptotic stationary assumptions, we have shown, for a given finite capacity, that Bode's integral formula can be extended. For the general case, we have derived an entropy rate inequality, which shows that the maximum *entropy reduction*, from the disturbance to the error signal, is bounded by the unstable poles of the open loop system minus the capacity of the preview communication system. By means of examples, we have illustrated that these bounds might be achieved.

Extension of the Framework: We have studied limitations of attenuation, from the disturbance to the error signal, and that leaves open opportunities in investigating extensions of our results to more general topologies. For instance, we expect that extensions of the bounds derived in Section IV might be possible and that, for more general schemes, they will also include the zeroes of the plant.

Incorporating the m-Delay: The m-delay depicted in Fig. 1 represents the travel time of the disturbance, counted from the instant it is generated until it reaches the plant. Our results provide bounds for performance, which are valid regardless of the m-delay, i.e., our results represent the worst case in terms of the m-delay. In order to reduce conservativeness and explicitly account for the m-delay, one needs to impose delay constrains at the remote preview scheme. Working with this constraint is very difficult and the extension of our results to such a setting is still an interesting topic of current research. The surprising structural results in [17], [33] may lead to efficient approaches to dealing with this problem.

Tightness of the Bound: In Section V, we provide an example where our bounds are tight for every m. This is a useful fact, because it implies that the remote preview system proposed in Section V achieves the lower bound, provided that the channel is additive, white and Gaussian and that P is stabilizable by means of a stable, linear and time-invariant H. If the remote preview system comprises an encoder and a non-Gaussian channel, then we expect that tight bounds, especially for small m-delays, will have to be obtained numerically.

also would like to express our gratitude to P. Narayan (Univ. Maryland) for carefully reading their first manuscript, to J. Chen (Univ. California, Riverside) for pointing us to a few references in Preview Control and to S. K. Mitter (M.I.T.) for interesting discussions. They are also grateful to the Associate Editor and to the reviewers for their feedback and suggestions, which have greatly improved the quality of this paper.

REFERENCES

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

[2] J. Chen, Z. Ren, S. Hara, and L. Qiu, "Optimal tracking performance: preview control and exponential signals," *IEEE Trans. Autom. Control*, vol. 46, no. 10, pp. 1647–1653, Oct. 2001.

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

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

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

[6] N. Elia, "When bode meets Shannon: Control-oriented feedback communication schemes," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1477–1488, Sep. 2004.

[7] H. El-Samad, H. Kurata, J. C. Doyle, C. A. Gross, and M. Khammash, "Surviving heat shock: Control strategies for robustness and performance," in *PNAS*, Feb. 22, 2005, vol. 102, no. 8, pp. 2736–2741.

[8] P. M. Fits, "The information capacity of the human motor system in controlling the amplitude of movement," *J. Experim. Psychol.*, vol. 47, pp. 381–391.

[9] J. S. Freudenberg and D. P. Looze, *Frequency Domain Properties of Scalar and Multivariable Systems*. Berlin, Germany: Springer-Verlag, 1988.

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

[11] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.

[12] P. A. Iglesias, "An analogue of Bode's integral for stable non-linear systems: Relations to entropy," *Proc. IEEE Conf. Decision Control*, pp. 3419–20, 2001.

[13] P. A. Iglesias, "Logarithmic integrals and system dynamics: An analogue of Bode's sensitivity integral for continuous-time, time-varying systems," *Linear Alg. Appl.*, vol. 343–344, pp. 451–71, 2002.

[14] E. A. Jonckheere and B. F. Wu, "Chaotic disturbance rejection and bode limitation," *Proc. Amer. Control Conf.*, pp. 2227–2231, 1992.

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

[16] A. N. Kolmogorov, "A new metric invariant of transitive dynamical systems and Lebesgue space automorphisms," *Dokl. Akad Nauk. SSR*, vol. 119, pp. 861–864, 1958.

[17] A. Mahajan and D. Teneketzis, On jointly optimal encoding, decoding and memory update for noisy real-time communication systems Dept. EECS, Univ. Michigan, Ann Arbor, MI, Control Group Report CGR-05-07, Oct. 2005.

[18] N. C. Martins and M. A. Dahleh, "Fundamental limitations in the presence of finite capacity feedback," *Proc. Amer. Control Conf.*, vol. 1, pp. 79–86, Jun. 8–10, 2005.

[19] N. C. Martins and M. A. Dahleh, Feedback control in the presence of noisy channels: Bode-like fundamental limitations of performance [Online]. Available: web.mit.edu/~nmartins/www MIT-LIDS Internal Publication #2647

[20] R. H. Middleton, J. Chen, and J. S. Freudenberg, "Tracking sensitivity and achievable H_∞ performance in preview control," *Automatica*, vol. 40, pp. 1297–1306, 2004.

[21] R. H. Middleton and J. H. Braslavsky, "On the relationship between logarithmic sensitivity integrals and limiting optimal control problems," in *Proc. IEEE Conf. Decision Control*, 2000, pp. 4990–4995.

[22] G. N. Nair and R. J. Evans, "Stabilization with data-rate-limited feedback: Tightest attainable bounds," *Syst. Control Lett.*, vol. 41, pp. 49–76, 2000.

[23] E. Ott, *Chaos in Dynamical Systems*. Cambridge, U.K.: Cambridge Univ. Press, 1993.

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

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

[26] M. B. Priestley, *Spectral Analysis and Time Series, Two-Volume Set: Volumes I and II (Probability and Mathematical Statistics)*. New York: Academic, Jan. 28, 1983.

[27] A. Sahai and S. K. Mitter, "The necessity and sufficiency of anytime capacity for control over a noisy communication link Part II: Vector systems," *IEEE Trans. Inform. Theory*, submitted for publication.

[28] S. Sarma, M. A. Dahleh, and S. Salapaka, "On time-varing bit-allocation maintaining stability: A convex parameterization approach," in *Proc. IEEE Conf. Decision Control*, Dec. 14–17, 2004, vol. 2, pp. 1430–1435.

[29] M. M. Seron, J. H. Braslavsky, and G. C. Goodwin, *Fundamental Limitations in Filtering and Control*. London, U.K.: Springer-Verlag, 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. Autom. Control*, vol. 44, no. 4, pp. 829–833, Apr. 1999.

[31] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, pp. 379–423, 1948.

[32] J. Sinai, "On the concept of entropy for a dynamical system," *Dokl. akad. Nauk. SSR*, vol. 124, pp. 768–771, 1959.

[33] D. Teneketzis, "On the structure of optimal real-time encoders and decoders in noisy communication," *IEEE Trans. Inform. Theory*, submitted for publication.

[34] G. Tadmor and L. Mirkin, "H_∞ control and estimation with preview part II: Fixed-size ARE solutions in discrete time," *IEEE Trans. Autom. Control*, vol. 50, no. 1, pp. 29–40, Jan. 2005.

[35] S. Tatikonda and S. K. Mitter, "Control under communication constraints," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.

[36] Z. Jarvis-Wloszek, D. Philbrick, M. A. Kaya, A. Packard, and G. Balas, "Control with disturbance preview and online optimization," *IEEE Trans. Autom. Control*, vol. 49, no. 2, pp. 266–270, Feb. 2004.

[37] B. F. Wu and E. A. Jonckheere, "A simplified approach to Bode's theorem for continuous-time and discrete-time systems," *IEEE Trans. Autom. Control*, vol. 37, no. 11, pp. 1797–1802, Nov. 1992.

[38] T. M. Yi, J. Goncalves, B. Ingalls, H. Sauro, and J. C. Doyle, A fundamental limitation on the robustness of complex systems in preparation.

[39] S. Yuksel and T. Basar, "Quantization and coding for decentralized LTI systems," in *Proc. IEEE Conf. Decision Control*, Dec. 9–12, 2003, vol. 3, pp. 2847–2852.

[40] 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.

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

Currently, he is an Assistant Professor with the Department of Electrical and Computer Engineering, the University of Maryland, College Park. He is also affiliated with the Institute for Systems Research and his research interests include fundamental limits of feedback and the fusion between control theory and information theory.

Dr. Martins received 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'01) received the B.S. degree from Texas A&M University, College Station, and the Ph.D. degree from Rice University, Houston, TX, both in electrical engineering, in 1983 and 1987, respectively.

Since 1987, he has been with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, where he is now a Full Professor.

Dr. Dahleh was the recipient of the Ralph Budd Award for the Best Thesis at Rice University in 1987, the George Axelby Outstanding Paper Award, a National Science Foundation 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, and the Graduate Students Council Teaching Award in 1995.

John C. Doyle (M'96) received the B.S. and M.S. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1977, and the Ph.D. degree in mathematics from the University of California, Berkeley, in 1984.

He is the John G. Braun Professor of Control and Dynamical Systems, Electrical Engineering, and Bioengineering at the California Institute of Technology, Pasadena.

Dr. Doyle received the 2004 IEEE Control Systems Award, the 1984 IEEE Centennial Outstanding Young Engineer Award, the 1984 Bernard Friedman Award, the 1983 American Automatic Control Council (AACC) Eckman Award, and the 1976 IEEE Hickernell Award. His best paper awards include the 1991 IEEEW. R. G. Baker Prize, the 1994 AACC O. Hugo Schuck Award, and the 1990 IEEE G. S. Axelby Award (twice).