The Effect of Delayed Side Information on Fundamental Limitations of Disturbance Attenuation

Yingbo Zhao, Vijay Gupta, and Jorge Cortés

***Abstract*— In this paper, we present a fundamental limitation of disturbance attenuation in discrete-time single-input single-output (SISO) feedback systems when the controller has delayed side information about the external disturbance. Specifically, we assume that the delayed information about the disturbance is transmitted to the controller across a finite Shannon-capacity communication channel. Our main result is a lower bound on the log sensitivity integral in terms of open-loop unstable poles of the plant and the characteristics of the channel, similar to the classical Bode integral formula. A comparison with prior work that considers the effect of preview side information of the disturbance at the controller indicates that delayed side information and preview side information play different roles in disturbance attenuation. In particular, we show that for open-loop stable systems, delayed side information cannot reduce the log integral of the sensitivity function whereas it can for open-loop unstable systems, even when the disturbance is a white stochastic process.**

I. Introduction

In a typical closed-loop control system as depicted in Fig. 1, the transfer function from the external disturbance $\mathbf{d}$ to the error signal $\mathbf{e}$ is a crucial measure of robustness and is known as the sensitivity function. Ideally, one prefers the magnitude of the sensitivity function to be small, which reflects good disturbance rejection performance.

However, Bode's integral equation [1] states that for an open-loop stable transfer function with relative degree greater than or equal to 2, the integral over all frequencies of the logarithm of the magnitude of the sensitivity function is 0. This result indicates that it is not possible to decrease the magnitude of sensitivity below 1 over all frequencies. Furthermore, by designing the controller, the integral of log sensitivity can only be shaped in frequency, which makes it a fundamental limitation on disturbance rejection. Freudenberg and Looze [2] generalized Bode's result to unstable open-loop systems and found that the integral of log sensitivity function is equal to the sum of the logarithm of the unstable open-loop poles.

A similar result is also known for discrete-time systems. For a discrete-time single-input single-output (SISO) open-loop system Σ_O (which includes both the plant and the controller) with a strictly proper transfer function $L(\omega)$, as shown in Fig. 1, the sensitivity function $S(\omega) = (1 + L(\omega))^{-1}$ satisfies, see [3], [4],

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log|S(\omega)|d\omega = \sum_{i:|p_i|>1} \log|p_i|, \qquad (1)$$

where the p_i's are the poles of $\omega \mapsto L(\omega)$ (i.e., the open-loop poles of Σ_O).

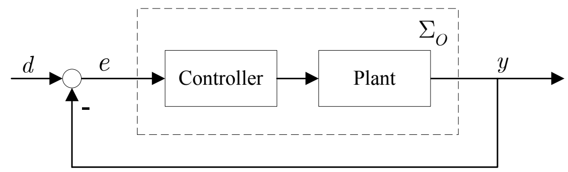

Fig. 1. Feedback system configuration with no side information. The random process d denotes the external disturbance, e denotes the error and y denotes the output of the system.

Due to its importance, the Bode integral formula has been extensively studied and further extended for multi-input multi-output (MIMO) systems [5]–[7], nonlinear systems [8], stochastic systems driven by Gaussian disturbance [9]–[12], switched systems [13], spatially invariant multi-agent systems [14], and time-varying systems [15]–[17].

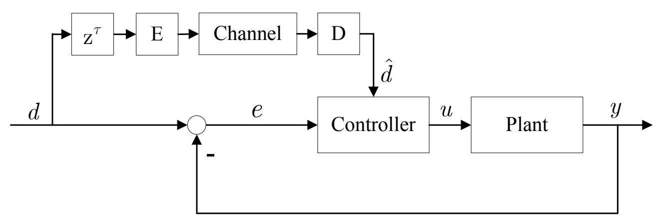

Fig. 2. Feedback system configuration when the controller has preview side information.

The work most closely related to the present one is [18], where the Bode integral equation (1) is extended to the case when the controller has a finite-horizon preview of the white disturbance d through a finite Shannon-capacity communication channel, as depicted in Fig. 2[1]. For this case, it holds that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log|S(\omega)|d\omega \geq \sum_{i:|p_i|>1} \log|p_i| - C, \qquad (2)$$

[1]We have slightly adapted the configuration in [18] to better suit our framework, without affecting its main result. First, the disturbance is now directly added to the output y instead of the control input u. The proofs of the main result in (2) for these two cases are similar, as shown in [12]. Second, instead of a physical delay block $z^{-\tau}$ between the disturbance and the plant, we introduce equivalently a preview block z^{τ} between the disturbance and the communication channel. Third, we assume that the disturbance is a white stochastic process.

Yingbo Zhao and Jorge Cortés are with the Department of Mechanical and Aerospace Engineering, University of California at San Diego, La Jolla, CA 92093. Vijay Gupta is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556. Emails: {yiz326,cortes}@ucsd.edu; vgupta2@nd.edu

where C represents the Shannon capacity of the preview channel. Inequality (2) shows that preview side information will, in general, improve the disturbance attenuation performance. In particular, if no preview information can be transmitted (C is zero), then (2) is comparable to Bode's classical result. If the disturbance can be fully transmitted (C is infinity), then the disturbance can be completely canceled by the controller and $S(\omega) = 0$ for all ω.

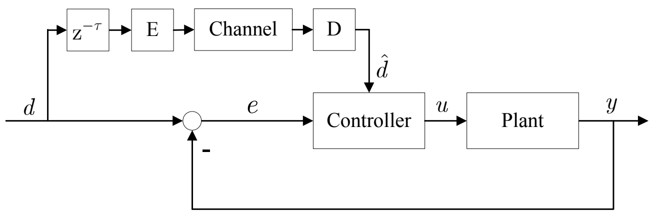

Fig. 3. Feedback system configuration when the controller has delayed side information.

In practice, it is often easier to obtain delayed side information (DSI) rather than preview side information (PSI). It seems reasonable to believe that, for a stochastic disturbance process that is white, DSI is not useful for disturbance attenuation because it provides no knowledge about the current or future disturbance. However, our main result of the paper shows the counterintuitive fact that even delayed side information of a white stochastic process d can improve the disturbance attenuation performance for unstable plants. For the problem setup depicted in Fig. 3, we obtain the following fundamental limitation

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log|S(\omega)|d\omega \geq \Big(\sum_{i:|p_i|>1} \log|p_i| - C\Big)^+, \quad (3)$$

where $(x)^+ \triangleq \max(x, 0)$ and C represents the Shannon capacity of the side communication channel. Note the similarity between (3) and Bode's integral equation without side information (1) and with preview side information (2).

Inequality (3) has various important implications. First, if the plant is stable, then no DSI can reduce the log integral of sensitivity. Second, if the plant is unstable, then DSI, even under a white stochastic disturbance process, can improve disturbance rejection. Third, unlike PSI, the contribution of DSI to the disturbance attenuation performance is upper bounded by the summation of the logarithm of the open-loop unstable poles. The intuition behind the latter fact is that DSI can only help stabilize the open-loop system but cannot reduce the controller's uncertainty about the disturbance.

The rest of the paper is organized as follows. Section II provides a brief review of basic information-theoretic concepts and Section III presents the problem statement. Section IV studies the effect of DSI on the differential entropy rate of the error signal. Section V presents the main result of the paper characterizing the effect of DSI on the log integral of sensitivity. Finally, we gather our conclusions and ideas for future work in Section VI.

II. Preliminaries

In this section we review basic notions and notation from stochastic processes and information theory following [19], [20]. Throughout the paper, we denote random variables and random processes using boldface letters. For any $j \leq k$, we let $x_j^k = [x(j), x(j+1), \ldots, x(k)]$ denote the row vector formed by a finite segment of a sequence $x = \{x(k)\}_{k=1}^\infty$. We omit the subscript j when it is equal to 0. A zero-mean stochastic process $\mathbf{x} = \{\mathbf{x}(k)\}_{k=1}^\infty$ is *asymptotically stationary* if

$$R_{\mathbf{x}}(n) \triangleq \lim_{k\to\infty} \mathbb{E}\left[\mathbf{x}(k)\mathbf{x}^T(k+n)\right] \quad (4)$$

exists for every $n \in \mathbb{Z}$. The power spectral density $\omega \mapsto \Phi_{\mathbf{x}}(\omega)$ of $\mathbf{x}$ is the discrete-time Fourier transform of $R_{\mathbf{x}}$. From [18], the *sensitivity function* $\omega \mapsto S_{\mathbf{x},\mathbf{y}}(\omega)$, between two asymptotically stationary stochastic processes $\mathbf{x}$ and $\mathbf{y}$ with power spectral densities $\Phi_{\mathbf{x}}$ and $\Phi_{\mathbf{y}}$, respectively, is

$$S_{\mathbf{x},\mathbf{y}}(\omega) \triangleq \sqrt{\Phi_{\mathbf{y}}(\omega)/\Phi_{\mathbf{x}}(\omega)}. \quad (5)$$

Given a continuous random vector $\mathbf{x}^k$ with probability density function $f(x^k)$ (in short $\mathbf{x}^k \sim f(x^k)$), the *differential entropy* $\mathbf{x}^k \mapsto h(\mathbf{x}^k)$ is

$$h(\mathbf{x}^k) \triangleq -\int f(x^k)\log f(x^k)\mathrm{d}x^k = -\mathbb{E}[\log(f(\mathbf{x}^k))].$$

For a continuously differentiable bijective function $\psi : \mathbb{R}^{k+1} \to \mathbb{R}^{k+1}$, one has

$$h\big(\psi(\mathbf{x}^k)\big) = h(\mathbf{x}^k) + \mathbb{E}[\log|\det(J)|], \quad (6)$$

where J is the Jacobian matrix of ψ with respect to x^k. For $\mathbf{x}^k \sim f(x^k)$ and $\mathbf{y}^k|\{\mathbf{x}^k = x^k\} \sim f(y^k|x^k)$, the *conditional differential entropy* of $\mathbf{y}^k$ given $\mathbf{x}^k$ is

$$h(\mathbf{y}^k|\mathbf{x}^k) \triangleq -\mathbb{E}_{\mathbf{x}^k,\mathbf{y}^k}[\log(f(\mathbf{y}^k|\mathbf{x}^k))].$$

Finally, the *mutual information* $I(\mathbf{x}^k;\mathbf{y}^k)$ between continuous random vectors $(\mathbf{x}^k,\mathbf{y}^k) \sim f(x^k,y^k)$ is defined as

$$I(\mathbf{x}^k;\mathbf{y}^k) = \int f(x^k,y^k)\log\frac{f(x^k,y^k)}{f(x^k)f(y^k)}\mathrm{d}x^k\mathrm{d}y^k.$$

The mutual information is symmetric (i.e., $I(\mathbf{x}^k;\mathbf{y}^k) = I(\mathbf{y}^k;\mathbf{x}^k)$) and nonnegative (i.e., $I(\mathbf{x}^k;\mathbf{y}^k) \geq 0$), which implies that conditioning reduces entropy (i.e., $h(\mathbf{y}^k) \geq h(\mathbf{y}^k|\mathbf{x}^k)$). Let $\mathbf{z}^k \sim f(z^k)$ and $(\mathbf{x}^k,\mathbf{y}^k)|\{\mathbf{z}^k = z^k\} \sim f(x^k,y^k|z^k)$. Denote the mutual information between $\mathbf{x}^k$ and $\mathbf{y}^k$ given $\{\mathbf{z}^k = z^k\}$ by $I(\mathbf{x}^k;\mathbf{y}^k|\mathbf{z}^k = z^k)$. Then, the conditional mutual information $I(\mathbf{x}^k;\mathbf{y}^k|\mathbf{z}^k)$ between $\mathbf{x}^k$ and $\mathbf{y}^k$ given $\mathbf{z}^k$ is defined as

$$I(\mathbf{x}^k;\mathbf{y}^k|\mathbf{z}^k) = \int I(\mathbf{x}^k;\mathbf{y}^k|\mathbf{z}^k = z^k)f(z^k)\mathrm{d}z^k.$$

The chain rule of differential entropy states that $h(\mathbf{x}^k) = \sum_{i=1}^k h(\mathbf{x}(i)|\mathbf{x}^{i-1})$ and implies the chain rule of mutual information $I(\mathbf{x}^k;\mathbf{y}^k) = \sum_{i=1}^k I(\mathbf{x}(i);\mathbf{y}^k|\mathbf{x}^{i-1})$.

The maximum entropy theorem states that if $\mathbf{x}^k \sim f(x^k)$ has covariance matrix $R_{\mathbf{x}^k} = \mathbb{E}\left[(\mathbf{x}^k)^T \mathbf{x}^k\right] > 0$, then

$$h(\mathbf{x}^k) \leq \frac{1}{2}\log\big((2\pi e)\det(R_{\mathbf{x}^k})\big) \tag{7}$$

with equality if and only if $\mathbf{x}^k$ is Gaussian distributed. If $\mathbf{x} \in \mathbb{R}$, we denote its covariance by $r_{\mathbf{x}}$ to indicate that it is a scalar. The above definitions can be extended to stationary stochastic processes. The differential entropy rate $\bar{h}(\mathbf{x})$ of a stationary continuous-valued stochastic process $\mathbf{x} = \{\mathbf{x}(k)\}_{k=1}^{\infty}$ is defined as

$$\bar{h}(\mathbf{x}) \triangleq \lim_{k\to\infty} \frac{1}{k} h(\mathbf{x}^k) = \lim_{k\to\infty} h(\mathbf{x}(k)|\mathbf{x}^{k-1}). \tag{8}$$

If $\mathbf{x}$ with $\mathbf{x}(k) \in \mathbb{R}^n$ has power spectral density $\Phi_{\mathbf{x}}$, then

$$\bar{h}(\mathbf{x}) \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{2}\log\big((2\pi e)^n \Phi_{\mathbf{x}}(\omega)\big)\mathrm{d}\omega \tag{9}$$

with equality if and only if $\mathbf{x}$ is a stationary Gaussian process.

III. Problem Formulation

Consider the problem setup depicted in Fig. 3 with the plant dynamics given by

$$\begin{aligned}\mathbf{x}(k+1) &= A\mathbf{x}(k) + B\mathbf{u}(k),\\ \mathbf{y}(k) &= H\mathbf{x}(k),\end{aligned} \tag{10}$$

where $\mathbf{x}(k) \in \mathbb{R}^n$ is the system state, $\mathbf{y}(k) \in \mathbb{R}$ is the system output and $\mathbf{u}(k) \in \mathbb{R}$ is the control input. At time $k \in \mathbb{Z}_{\geq 0}$, the controller has access to the error signal $\mathbf{e} = \mathbf{d} - \mathbf{y} \in \mathbb{R}$ as well as side information $\hat{\mathbf{d}}$ provided across a side communication channel with delay $\tau \in \mathbb{Z}_{>0}$. Formally, the side channel is defined by[2] an input set $\mathcal{M}$, an output set $\mathcal{N}$, and a transition probability mass function $p(n(k)|m(k))$. The latter characterizes the probability of the decoder receiving the output symbol $n(k) \in \mathcal{N}$ given that the symbol $m(k) \in \mathcal{M}$ is transmitted at time $k \in \mathbb{Z}_{\geq 0}$ by the encoder. We assume that the channel has Shannon capacity [19] equal to C (bits/sec).

At time $k \in \mathbb{Z}_{\geq 0}$, the encoder E maps $\mathbf{d}^{k-\tau}$ into a symbol $m(k) \in \mathcal{M}$, which is transmitted to the controller over the communication channel. The controller uses the error sequence $\mathbf{e}^k$ and the channel output sequence $\hat{\mathbf{d}}^k$ to generate a control sequence of the form

$$\mathbf{u}(k) = u_k(k, \hat{\mathbf{d}}^k, \mathbf{e}^k), \tag{11}$$

where $u_k : \mathbb{R}^{2k+3} \to \mathbb{R}$ is a time-varying, possibly nonlinear, function. We assume that the control laws in (11) are such that the random processes describing the closed-loop dynamics have well defined continuous joint probability density functions and are asymptotically stationary processes. In addition, we make the following assumptions.

Assumption 1: All the random processes have stable covariances, that is, the system is mean-square stable.

[2]The input/output set is finite for digital channels and infinite for analog channels.

Assumption 2: The disturbance process $\mathbf{d}$ is a zero-mean Gaussian process with independent and identically distributed (i.i.d.) random variables $\mathbf{d}(k)$. The plant's initial condition $\mathbf{x}(0)$ is a zero-mean random variable with covariance $R_{x(0)}$ and finite differential entropy, and is independent of the disturbance process $\mathbf{d}$.

In this framework, we follow an information–theoretic approach and compute a lower bound on

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log S_{\mathbf{d},\mathbf{e}}(\omega)d\omega \tag{12}$$

where $S_{\mathbf{d},\mathbf{e}}(\omega)$ is the sensitivity function defined in (5). Note that if the control law is linear time-invariant, then $S_{\mathbf{d},\mathbf{e}}(\omega)$ reduces to the magnitude of the transfer function between $\mathbf{d}$ and $\mathbf{e}$ at frequency ω.

IV. Effect of delayed side information on the differential entropy rate of the error

Here we study the impact that DSI has on the differential entropy rate of the error signal available to the controller. The results of this section will help us establish later in Section V our main result on the fundamental limitation of the log integral of the sensitivity.

Lemma 1: (A lower bound on the differential entropy rate of the error). For the feedback configuration in Fig. 3, the following holds

$$\bar{h}(\mathbf{e}) \geq \bar{h}(\mathbf{d}) + \big(\bar{I}(\mathbf{e}; \mathbf{x}(0)|\hat{\mathbf{d}}) - \bar{I}(\mathbf{d}; \hat{\mathbf{d}})\big)^{+}. \tag{13}$$

Proof: From the definition of mutual information, for any $k \in \mathbb{Z}_{\geq 0}$, one has

$$\begin{aligned}h(\mathbf{e}(k)|\mathbf{e}^{k-1}) &= h(\mathbf{e}(k)|\mathbf{e}^{k-1}, \hat{\mathbf{d}}^{k-1}, \mathbf{x}(0))\\ &\quad + I(\mathbf{e}(k); \hat{\mathbf{d}}^{k-1}, \mathbf{x}(0)|\mathbf{e}^{k-1}).\end{aligned} \tag{14}$$

For the first term on the right-hand side of (14) it holds that

$$\begin{aligned}h(\mathbf{e}(k)|\mathbf{e}^{k-1}, \hat{\mathbf{d}}^{k-1}, \mathbf{x}(0)) &\overset{(a)}{=} h(\mathbf{e}(k)|\mathbf{e}^{k-1}, \hat{\mathbf{d}}^{k-1}, \mathbf{x}(0), \mathbf{y}^k)\\ &\overset{(b)}{=} h(\mathbf{d}(k)|\mathbf{d}^{k-1}, \hat{\mathbf{d}}^{k-1}, \mathbf{x}(0))\\ &\overset{(c)}{=} h(\mathbf{d}(k)|\mathbf{d}^{k-1}),\end{aligned} \tag{15}$$

where (a) holds because $\mathbf{y}^k$ is a function of $(\mathbf{e}^{k-1}, \hat{\mathbf{d}}^{k-1}, \mathbf{x}(0))$, (b) follows from the relation that $\mathbf{e}(k) = \mathbf{d}(k) - \mathbf{y}(k)$, and (c) holds because $\mathbf{x}(0)$ and $\mathbf{d}$ are independent, $\mathbf{d}$ is white and the side information is *delayed*, i.e., $I(\hat{\mathbf{d}}^{k-1}; \mathbf{d}(k)|\mathbf{d}^{k-1}) = 0$.

For the second term on the right-hand side of (14), we take an arbitrary integer $N \geq k$ and obtain

$$\begin{aligned}&I(\mathbf{e}(k); \hat{\mathbf{d}}^{k-1}, \mathbf{x}(0)|\mathbf{e}^{k-1})\\ &\overset{(a)}{=} I(\mathbf{e}(k); \hat{\mathbf{d}}^{N}, \mathbf{x}(0)|\mathbf{e}^{k-1})\\ &\quad - I(\mathbf{e}(k); \hat{\mathbf{d}}_k^{N}|\hat{\mathbf{d}}^{k-1}, \mathbf{x}(0), \mathbf{e}^{k-1})\\ &\overset{(b)}{=} I(\mathbf{e}(k); \hat{\mathbf{d}}^{N}, \mathbf{x}(0)|\mathbf{e}^{k-1})\end{aligned}$$

$$
\begin{aligned}
&- I(\mathbf{d}(k); \hat{\mathbf{d}}_k^N | \hat{\mathbf{d}}^{k-1}, \mathbf{x}(0), \mathbf{d}^{k-1}) \\
&\overset{(c)}{\geq} I(\mathbf{e}(k); \mathbf{x}(0) | \mathbf{e}^{k-1}, \hat{\mathbf{d}}^N) \\
&- I(\mathbf{d}(k); \hat{\mathbf{d}}^N | \mathbf{d}^{k-1}, \mathbf{x}(0)), \qquad (16)
\end{aligned}
$$

where (a) follows from the chain rule of mutual information, (b) holds because $\mathbf{e}(k) = \mathbf{d}(k) - \mathbf{y}(k)$ and $\mathbf{y}^k$ is a function of $(\mathbf{e}^{k-1}, \hat{\mathbf{d}}^{k-1}, \mathbf{x}(0))$, and (c) follows from the fact that conditioning reduces entropy.

Substituting (15) and (16) into (14), and by summation from $k = 0$ to $k = N$, dividing by N, and taking the limit as $N \to \infty$, we obtain

$$\bar{h}(\mathbf{e}) \geq \bar{h}(\mathbf{d}) + \bar{I}(\mathbf{e}; \mathbf{x}(0) | \hat{\mathbf{d}}) - \bar{I}(\mathbf{d}; \hat{\mathbf{d}} | \mathbf{x}(0)). \qquad (17)$$

Moreover, from (14) and (15) and the nonnegativity of mutual information, we deduce that

$$\bar{h}(\mathbf{e}) \geq \bar{h}(\mathbf{d}). \qquad (18)$$

The result follows from the combination of (17) and (18) and the assumption that $\mathbf{x}(0)$ is independent with $(\mathbf{d}, \hat{\mathbf{d}})$. □

We make the following remarks on the result in Lemma 1.

Remark 1 (Case with no side information): The differential entropy rate is closely related to the power spectral density, as shown in (9), and itself can be used as a performance measure [18]. In the case when the side channel is absent, Lemma 1 reduces to

$$\bar{h}(\mathbf{e}) \geq \bar{h}(\mathbf{d}) + \bar{I}(\mathbf{e}; \mathbf{x}(0)), \qquad (19)$$

which is similar to [10, Theorem 4.2]. ■

Remark 2: (Delayed side information provides information about the initial condition). One can see from (13) that the differential entropy rate $\bar{h}(\mathbf{e})$ stems from two sources: the uncertainty in $\mathbf{d}(k)$ (characterized by the first term in the right-hand side of (13)) and information about the initial condition (characterized by the second term in the right-hand side of (13)) that is necessary for mean-square stabilization of the closed-loop system. Since the disturbance process $\mathbf{d}$ is assumed to be white and the side information $\hat{\mathbf{d}}$ is delayed, $\hat{\mathbf{d}}^k$ cannot provide any information about $\mathbf{d}(k)$. Nevertheless, even if $\hat{\mathbf{d}}$ is independent of $\mathbf{x}(0)$, it can still help to stabilize the system by providing conditional information about the initial condition, in other words, $I(\hat{\mathbf{d}}^k; \mathbf{x}(0) | \mathbf{e}^k)$ can be greater than zero. ■

Based on Lemma 1, we prove the following result.

Lemma 2: (DSI can reduce the differential entropy rate of the error). For the feedback configuration in Fig. 3, the following holds

$$\bar{h}(\mathbf{e}) \geq \bar{h}(\mathbf{d}) + \Big(\sum_{i:|\lambda_i(A)|>1} \log |\lambda_i(A)| - C \Big)^+, \qquad (20)$$

where $\lambda_i(A)$ are the eigenvalues of the system matrix A in (10).

Proof: The proof proceeds in two steps. First, by the chain rule of mutual information, one has

$$
\begin{aligned}
&I(\mathbf{e}^{N-1}; \mathbf{x}(0) | \hat{\mathbf{d}}^{N-1}) \\
&= I(\mathbf{e}^{N-1}, \hat{\mathbf{d}}^{N-1}; \mathbf{x}(0)) - I(\hat{\mathbf{d}}^{N-1}; \mathbf{x}(0)) \\
&\overset{(a)}{\geq} I(\mathbf{u}^{N-1}; \mathbf{x}(0)), \qquad (21)
\end{aligned}
$$

where (a) follows from the data processing inequality [19] and the assumption that $\mathbf{x}(0)$ is independent of $\mathbf{d}$ (and thus $\hat{\mathbf{d}}$). Therefore,

$$\bar{I}(\mathbf{e}; \mathbf{x}(0) | \hat{\mathbf{d}}) \geq \bar{I}(\mathbf{u}; \mathbf{x}(0)) \overset{(b)}{\geq} \sum_{i:|\lambda_i(A)|>1} \log |\lambda_i(A)|, \qquad (22)$$

where (b) follows from Assumption 1 and [18, Lemma 4.1].

Second, by the definition of channel capacity [19] it holds that $\bar{I}(\mathbf{d}; \hat{\mathbf{d}}) \leq C$, which combines with (22) and Lemma 1 yields the result. □

In the appendix we present an example where (b) in (22) is achieved.

V. Main result: effect of delayed side information on the log integral of sensitivity

In this section, we present the main result of this paper. The result shows that the log integral of sensitivity can only be reduced by delayed side information about the disturbance at the controller when the plant is unstable, and the reduction is no more than $\min(\sum_{i:|\lambda_i(A)|>1} \log |\lambda_i(A)|, C)$.

Theorem 1: (DSI can reduce the log integral of sensitivity). For the feedback configuration in Fig. 3 and under Assumptions 1 and 2,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log S_{\mathbf{d},\mathbf{e}}(\omega) d\omega \geq \Big(\sum_{i:|\lambda_i(A)|>1} \log |\lambda_i(A)| - C \Big)^+. \qquad (23)$$

Proof: Note that the log integral of the sensitivity function can be lower bounded by the difference of the entropy rates $\bar{h}(\mathbf{d})$ and $\bar{h}(\mathbf{e})$ of the disturbance and error processes as follows

$$
\begin{aligned}
&\frac{1}{2\pi} \int_{-\pi}^{\pi} \log S_{\mathbf{d},\mathbf{e}}(\omega) d\omega \\
&= \frac{1}{4\pi} \int_{-\pi}^{\pi} \big(\log \Phi_{\mathbf{e}}(\omega) - \log \Phi_{\mathbf{d}}(\omega) \big) d\omega \qquad (24a) \\
&= \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \Phi_{\mathbf{e}}(\omega) d\omega - \bar{h}(\mathbf{d}) \qquad (24b) \\
&\geq \bar{h}(\mathbf{e}) - \bar{h}(\mathbf{d}) \qquad (24c)
\end{aligned}
$$

Here (24a) follows from the definition of sensitivity function in (5), (24b) and (24c) follow from the maximum entropy theorem and the assumption that $\mathbf{d}$ is Gaussian (equality in (24c) holds iff $\mathbf{e}$ is a Gaussian process with power spectral density $\Phi_{\mathbf{e}}$). The result then follows from combining this inequality with Lemma 2. □

An important observation about the result in (23) is that, unlike PSI, the contribution of DSI to the disturbance attenuation performance is upper bounded by the summation of the logarithm of the open-loop unstable poles. The reason is that DSI can only help stabilize the open-loop system but cannot reduce the controller's uncertainty about the disturbance, cf. also Remark 2. In general, the lower bound in Theorem 1 is tight, as illustrated by the following example.

Example 1: (Lower bound on log integral of sensitivity is tight). Consider the configuration in Fig. 3 for the special case of a scalar plant

$$\begin{aligned}\mathbf{x}(k+1) &= a\mathbf{x}(k) + \mathbf{u}(k),\\ \mathbf{y}(k) &= \mathbf{x}(k),\end{aligned} \tag{25}$$

for some $|a| > 1$. Assume that $\mathbf{d}(0)$ is a zero-mean continuous random variable uniformly distributed over a compact support and $\mathbf{d}_1^\infty$ is a zero-mean stationary Gaussian process with i.i.d. random variables, independent with $\mathbf{d}(0)$. Let the side channel be a noiseless digital channel with capacity $C > \log|a|$ bits/sec.

Let the side channel transmit $\mathbf{d}(0)$ at every time step k, so that the controller has an increasingly better estimate $\hat{\mathbf{d}}_0(k)$ of $\mathbf{d}(0)$. In particular, we can find an encoder/decoder pair such that $\mathbb{E}(\|\mathbf{d}(0) - \hat{\mathbf{d}}_0(k)\|^2) \leq 2^{-2Ck}\,\mathbb{E}(\|\mathbf{d}(0)\|^2)$. A simple design is to divide the compact support of $\mathbf{d}(0)$ into 2^C boxes of equal size and transmit the index of the box where $\mathbf{d}(0)$ resides in and let $\hat{\mathbf{d}}(0) = \mathbb{E}(\mathbf{d}(0)) = 0$.

Under the control law

$$\mathbf{u}(k) = \begin{cases} a(\hat{\mathbf{d}}_0(0) - \mathbf{e}(0)), & k = 0,\\ a^{k+1}(\hat{\mathbf{d}}_0(k) - \hat{\mathbf{d}}_0(k-1)), & k \geq 1,\end{cases}$$

the corresponding closed-loop dynamics is given by

$$\mathbf{x}(k) = a^k(\hat{\mathbf{d}}_0(k-1) - \mathbf{d}(0)).$$

It then follows that

$$\begin{aligned}\mathbb{E}(\|\mathbf{x}(k)\|^2) &= a^{2k}\,\mathbb{E}(\|\hat{\mathbf{d}}_0(k-1) - \mathbf{d}(0)\|^2)\\ &\leq (2^{-C}a)^{2k}2^{2C}\,\mathbb{E}(\|\mathbf{d}(0)\|^2).\end{aligned}$$

Since $(2^{-C}a)^2 < 1$ by the assumption that $C > \log|a|$, $\mathbb{E}(\|\mathbf{x}(k)\|^2)$ converges to 0 exponentially and the closed-loop system is mean-square stable.

Notice that $\mathbb{E}[\mathbf{x}(k_1)\mathbf{d}^T(k_2)] = 0$, $\forall k_1, k_2 \in \mathbb{Z}_{>0}$. Moreover, $\lim_{k\to\infty} \mathbb{E}\left[\mathbf{x}(k)\mathbf{x}^T(k+n)\right] = 0$ because

$$\lim_{k\to\infty} \mathbb{E}\big(\|\mathbf{x}(k)\|^2 + \|\mathbf{x}(k+n)\|^2\big) = 0.$$

Then, from (4),

$$\begin{aligned}R_{\mathbf{e}}(n) &= \lim_{k\to\infty} \mathbb{E}\left[\mathbf{e}(k)\mathbf{e}^{\mathsf{T}}(k+n)\right]\\ &= \lim_{k\to\infty} \mathbb{E}\left[(\mathbf{x}(k) + \mathbf{d}(k))(\mathbf{x}(k+n) + \mathbf{d}(k+n))^T\right]\\ &= \lim_{k\to\infty} \mathbb{E}[\mathbf{d}(k)\mathbf{d}^T(k+n)] = R_{\mathbf{d}}(n),\end{aligned}$$

which means that $\Phi_{\mathbf{e}}(\omega) = \Phi_{\mathbf{d}}(\omega)$ (for stationary Gaussian processes, this also implies $\bar{h}_{\mathbf{e}} = \bar{h}_{\mathbf{d}}$) and thus $S_{\mathbf{d},\mathbf{e}}(\omega) = 1$, $\forall\omega \in [0, 2\pi)$. Therefore,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log|S_{\mathbf{d},\mathbf{e}}(\omega)|d\omega = 0 = \big(\log|a| - C, 0\big)^+,$$

which shows that we can achieve the lower bound for any $C > \log|a|$. ■

Remark 3: (Use of delayed side information versus the error signal for disturbance attenuation). Without any side information, information about the initial condition $\mathbf{x}(0)$ can only be carried by $\mathbf{e}$ to the controller for stabilization. In particular, a necessary condition for the closed-loop system to be mean-square stable is $\bar{I}(\mathbf{e};\mathbf{x}(0)) \geq \log|a|$, cf. [18, Lemma 4.1], which is a factor in the lower bound of $\bar{h}_{\mathbf{e}}$, cf. (19). Under the delayed side information $\hat{\mathbf{d}}$, the controller can stabilize the system without using the error signal $\mathbf{e}$ (except for $\mathbf{e}(0)$), as shown in Example 1. From this viewpoint, the delayed side information $\hat{\mathbf{d}}$ can improve the disturbance attenuation performance by taking over the stabilization task from the error signal $\mathbf{e}$. ■

VI. Conclusions

We have studied the effect of delayed side information on a Bode-like fundamental limitation for disturbance attenuation in discrete-time systems. Our result is valid for linear time-invariant plants with controllers that can be nonlinear and time-varying. We have shown the somewhat counterintuitive result that, for plants with unstable poles, delayed side information about the external stochastic disturbance can reduce the log integral of sensitivity even if the disturbance is white. Unlike the case of preview side information, we have also observed that delayed side information can only reduce the log integral of sensitivity by a limited amount. Future work will study the effect of delayed and preview side information on other fundamental limitations, such as the log integral of the complementary sensitivity function, and the extension of the results to multi-agent networked scenarios.

Acknowledgments

This research was partially supported by NSF Award CNS-1329619.

References

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

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

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

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

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

[6] J. Chen and C. N. Nett, "Sensitivity integrals for multivariable discrete-time systems," *Automatica*, vol. 31, no. 8, pp. 1113–1124, 1995.

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

[8] M. Seron, J. Braslavsky, P. Kokotovic, and D. Mayne, "Feedback limitations in nonlinear systems: From bode integrals to cheap control," *IEEE Transactions on Automatic Control*, vol. 44, no. 4, pp. 829–833, 1999.

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

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

[11] Y. Zhao, P. Minero, and V. Gupta, "Disturbance propagation analysis in vehicle formations: An information-theoretic approach," in *American Control Conference*, pp. 1338–1343, 2013.

[12] Y. Zhao, P. Minero, and V. Gupta, "On disturbance propagation in leader–follower systems with limited leader information," *Automatica*, vol. 50, no. 2, pp. 591–598, 2014.

[13] D. Li and N. Hovakimyan, "Bode-like integral for stochastic switched systems in the presence of limited information," *Automatica*, 2012.

[14] P. Padmasola and N. Elia, "Bode integral limitations of spatially invariant multi-agent systems," in *45th IEEE Conference on Decision and Control*, pp. 4327–4332, 2006.

[15] M. Seron, J. Braslavsky, and G. Goodwin, *Limitations in Filtering and Control*. London: Springer-Verlag, 1997.

[16] P. A. Iglesias, "Tradeoffs in linear time-varying systems: an analogue of bode's sensitivity integral," *Automatica*, vol. 37, no. 10, pp. 1541–1550, 2001.

[17] P. A. Iglesias, "Logarithmic integrals and system dynamics: an analogue of bode's sensitivity integral for continuous-time, time-varying systems," *Linear algebra and its applications*, vol. 343-344, pp. 451–471, 2002.

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

[19] T. Cover and J. Thomas, *Elements of Information Theory, 2nd Edition*. New York: Wiley, 2006.

[20] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

APPENDIX I
ACHIEVING THE EQUALITY IN (22)

We show through a specific scalar system example that the inequality in (b) in (22), see also [18, Lemma 4.1], can hold with equality. Consider the configuration in Fig. 1 with the scalar plant

$$\mathbf{x}(k+1) = a\mathbf{x}(k) + b\mathbf{u}(k), \quad |a| > 1. \tag{26}$$

Let Assumption 2 be satisfied and further assume that $\mathbf{x}(0)$ is Gaussian distributed. The state at any time k can be written as a function of the initial condition $\mathbf{x}(0)$ and the input sequence $\mathbf{u}^{k-1}$,

$$\mathbf{x}(k) = a^k\mathbf{x}(0) + a^k\sum_{j=0}^{k-1} a^{-(j+1)}b\mathbf{u}(j). \tag{27}$$

Define $\hat{\mathbf{x}}_0(k) \triangleq -\sum_{j=0}^{k-1} a^{-(j+1)}b\mathbf{u}(j)$ as an estimate of the initial condition $\mathbf{x}(0)$ at time $k \geq 1$ and let $\hat{\mathbf{x}}_0(0) = 0$. Denote $\mathbf{x}_e(k) \triangleq \mathbf{x}(0) - \hat{\mathbf{x}}_0(k)$ as the estimation error and rewrite (27) as

$$\mathbf{x}(k) = a^k\mathbf{x}_e(k). \tag{28}$$

Define $r_{\mathbf{x}_e}(k) \triangleq \mathbb{E}\left[\mathbf{x}_e(k)\mathbf{x}_e^{\mathsf{T}}(k)\right]$ and propose the following linear time-varying control law

$$\mathbf{u}(k) = K(k)\mathbf{e}(k) \triangleq \frac{-ab^{-1}r_{\mathbf{x}_e}(k)}{r_{\mathbf{x}_e}(k) + a^{-2k}r_d}\mathbf{e}(k). \tag{29}$$

Under this controller, the error covariance is obtained as

$$r_{\mathbf{x}_e}(k) = \frac{r_{x_0}}{(\sum_{i=0}^{k-1} a^{2i})r_d^{-1}r_{x_0} + 1}. \tag{30}$$

Note that the controller (29) makes $\hat{\mathbf{x}}_0(k)$ the minimum mean square error estimate of $\mathbf{x}(0)$. Thus, $\mathbf{x}_e(k)$ and $\mathbf{u}^{k-1}$ are independent, which means

$$I\left(\mathbf{x}_e(k); \mathbf{u}^{k-1}\right) = 0. \tag{31}$$

It can also be easily verified that the closed-loop system is mean-quare stable and

$$\lim_{k\to\infty} r_{\mathbf{x}}(k) = \lim_{k\to\infty} a^{2k}r_{\mathbf{x}_e}(k) = (a^2-1)r_d$$

is finite, which implies that $\lim_{k\to\infty} h(\mathbf{x}(k))/k = 0$ by (7). Moreover, the gain of the controller also converges to

$$\lim_{k\to\infty} K(k) = -b^{-1}(a - a^{-1}).$$

Finally, it holds that

$$\begin{aligned}
\lim_{k\to\infty}\frac{I\left(\mathbf{x}(0); \mathbf{u}^{k-1}\right)}{k} &= \lim_{k\to\infty}\frac{h(\mathbf{x}(0))}{k} \\
&\quad - \lim_{k\to\infty}\frac{h\left(\mathbf{x}(0)\,|\mathbf{u}^{k-1}\right)}{k} \\
&\overset{(a)}{=} -\lim_{k\to\infty}\frac{h\left(\mathbf{x}_e(k)\,|\mathbf{u}^{k-1}\right)}{k} \\
&\overset{(b)}{=} -\lim_{k\to\infty}\frac{h(\mathbf{x}_e(k))}{k} \\
&\overset{(c)}{=} \log|a| - \lim_{k\to\infty}\frac{h(\mathbf{x}(k))}{k} \\
&= \log|a|,
\end{aligned}$$

where (a) holds because $\lim_{k\to\infty} h(\mathbf{x}(0))/k = 0$, $\mathbf{x}(0) = \mathbf{x}_e(k) + \hat{\mathbf{x}}_0(k)$ and $\hat{\mathbf{x}}_0(k)$ is a function of $\mathbf{u}^{k-1}$, (b) follows from (31), (c) holds because of (28).