# Feedback Control Performance over a Noisy Communication Channel

J. S. Freudenberg
EECS Department
University of Michigan
Ann Arbor MI 48109
jfr@eecs.umich.edu

R. H. Middleton
Hamilton Institute
National University of Ireland Maynooth
Co Kildare, Ireland
richard.middleton@nuim.ie

*Abstract*— **We consider the problem of minimizing the variance in the output of a plant that is driven by a Gaussian disturbance using measurements of the plant output obtained from a Gaussian channel. For the special case in which only the variance at the terminal time is penalized, we derive an optimal linear time-varying communication and control strategy, and argue that nonlinear strategies cannot achieve better performance.**

## I. INTRODUCTION

Recent years have seen much interest in the limitations imposed on a feedback system by the presence of a noisy communication channel in the feedback path, as depicted in Figure 1. One problem is to determine the minimal channel capacity required to stabilize an open loop unstable plant. The solution to this problem is known for noise-free data rate limited channels [1]–[3] and additive Gaussian noise channels [4]. A more difficult problem is that of determining the optimal performance, in terms of disturbance attenuation, that is achievable given the presence of a noisy channel with fixed capacity.
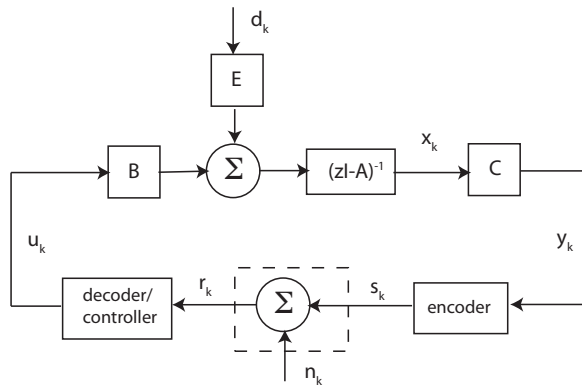


Fig. 1. Feedback control over a noisy communication channel.

In this paper the plant is assumed to be linear with initial state and white noise disturbance input that are Gaussian random variables, and the channel is assumed Gaussian. We characterize performance in terms of the variance of the plant output, and seek an encoder and decoder/controller to minimize the output variance at a specified terminal time. It is important to note that this problem has a nonclassical information pattern, in that the encoder and decoder/controller do not have access to the same information [5]. For such problems, the "separation" property of estimation and control need not hold [5], and the optimal communication and control strategies need not be linear [6]. Situations in which linear strategies may indeed be linear are studied in [7], [8]. We shall show that for our special measure of performance, i.e., output variance at a terminal time, the optimal control and communication strategies are in fact linear. We also show that, although there is no conflict between estimation and control for our problem, nevertheless the control and communication strategies must be carefully coordinated.

In Section II we define the problem of minimizing plant output variance at a fixed terminal time, and derive a formula for the optimal control to be applied at the last possible instant. This formula shows that the control problem reduces to one of estimating a single linear combination of states. In Section III we suppose that the encoder has access to additional information, and derive a linear time-varying strategy to estimate that combination of states used in the optimal control. We argue in Section IV that a nonlinear strategies cannot achieve better performance. In Section V we return to the original problem and show how to achieve the same performance achieved in Section III but without the additional information at the encoder. The strategies used are again linear and time-varyng. Conclusions are presented in Section VI.

Other researchers have studied feedback control performance over a communication channel; a partial review follows. The authors of [9] derive a lower bound on a measure of disturbance attenuation that is stated in terms of channel capacity; however, the realistic assumption of causality is not invoked in the proof. The authors of [10] study performance limitations imposed by a vector Gaussian channel, with one channel per state of the plant. The author of [11] relates the problem of feedback stabilization over a communication channel to that of communication over a channel with feedback. The authors of [12], [13] consider performance constraints imposed by noise free, data rate limited channels. The authors of [14] study the joint optimum design of communication and control strategies for feedback over noisy channels, and show that linear strategies are optimal only for first order linear systems with Gaussian noise and quadratic cost.

## II. PRELIMINARIES

The plant in Figure 1 is described by

$$x_{k+1} = Ax_k + Bu_k + Ed_k, \quad x_k \in \mathbb{R}^n, \ u_k, d_k \in \mathbb{R} \quad (1)$$

$$y_k = Cx_k, \quad y_k \in \mathbb{R}, \quad (2)$$

where $x_0$ is a zero mean Gaussian random variable with covariance $\Sigma_{0|-1}$ and $d_k$ is a stationary zero mean Gaussian white noise sequence with variance $\sigma_d^2$. The Gaussian channel has input power constraint $\mathcal{E}\{s_k^2\} \leq P$ and additive white noise with variance $\sigma_n^2$. The encoder $f_k()$ may be nonlinear and time-varying, and the channel input at time $k$ is allowed to depend on the sequence of current and previous plant outputs: $s_k = f_k(y^k)$, where $y^k = \{y_0, y_1, \ldots, y_k\}$. The decoder/controller $g_k()$ may also be nonlinear and time-varying, and the control input at time $k$ is allowed to depend on the sequence of current and previous channel outputs: $u_k = g_k(r^k)$, where $r^k = \{r_0, r_1, \ldots, r_k\}$.

The design problem is to choose the encoder $f_k()$ and decoder/controller and $g_k()$, $k = 0, \ldots, N$, to minimize the variance of the plant output at a terminal time $k = N + 1$. The optimal value of this cost function is given by

$$J^* = \inf_{f_k, g_k} \mathcal{E}\{y_{N+1}^2\}.$$

Denote the conditional expectation of $y_{N+1}$ given the channel output sequence $r^N$ by

$$\hat{y}_{N+1|N} = \mathcal{E}\{y_{N+1}|r^N\}$$

and the resulting estimation error by $\tilde{y}_{N+1|N} = \hat{y}_{N+1|N} - y_{N+1}$. It follows immediately from properties of the conditional expectation (cf. [15, p. 97]) that the variance of the conditional estimation error provides a lower bound on the optimal cost:

$$J^* \geq \mathcal{E}\{\tilde{y}_{N+1|N}^2\}. \quad (3)$$

Under mild hypotheses, the lower bound (3) can be achieved.

**Proposition 1** *Assume that $CB \neq 0$. Then the control*

$$u_N = -(CB)^{-1}CA\hat{x}_{N|N}, \quad (4)$$

*where $\hat{x}_{N|N} = \mathcal{E}\{x_N|r^N\}$, yields $\mathcal{E}\{y_{N+1}^2\} = \mathcal{E}\{\tilde{y}_{N+1|N}^2\}$. Moreover,*

$$\mathcal{E}\{y_{N+1}^2\} = \mathcal{E}\{(CA\tilde{x}_{N|N})^2\} + (CE)^2\sigma_d^2. \quad (5)$$

*Proof:* Substituting (4) into the state equations (1)-(2) shows that

$$y_{N+1} = CAx_N - CA\hat{x}_{N|N} + CEd_N,$$

and thus $\hat{y}_{N+1|N} = 0$ and $y_{N+1} = \tilde{y}_{N+1|N}$. The identity (5) follows because $\tilde{x}_{N|N}$ and $d_N$ are independent random variables. ∎

We have shown that the optimal control to apply at time $N$ sets the variance of the plant output to its theoretical minimum, given by the variance of the conditional estimation error (5). The term $(CE)^2\sigma_d^2$ in (5) is due to the disturbance at time $N$, which is uninfluenced by the control and communication

strategies because it has not yet propagated to the encoder. Hence the estimation problem reduces to that of minimizing $\mathcal{E}\{(CA\tilde{x}_{N|N})^2\}$, and the problem of minimizing the variance of the plant output reduces to that of minimizing the variance in the conditional estimate of the single linear combination of states $CAx_N$. Iterating the state equations (1)-(2) shows that

$$CAx_N = CA^{N+1}x_0 + \sum_{j=0}^{N-1} CA^{N-j}Ed_j + \sum_{j=0}^{N-1} CA^{N-j}Bu_j.$$

Since the control signal is known at the decoder, the only information that needs to be communicated over the channel is the "message"

$$m_N = CA^{N+1}x_0 + \sum_{j=0}^{N-1} CA^{N-j}Ed_j. \quad (6)$$

The problem of choosing the encoder $f_k()$, $k = 0, \ldots, N$ and decoder/controller $g_k()$, $k = 0, \ldots, N-1$ thus reduces to that of obtaining the best possible estimate of $m_N$ given $N + 1$ uses of the channel. If we denote the estimate and estimation error by $\hat{m}_{N|N} = \mathcal{E}\{m_N|r^N\}$ and $\tilde{m}_{N|N}$, then the variance in the plant output is given by

$$\mathcal{E}\{y_{N+1}^2\} = \mathcal{E}\{\tilde{m}_{N|N}^2\} + (CE)^2\sigma_d^2. \quad (7)$$

An obvious complication is that the complete message is not available until $k = N$, the last time step at which the channel may be used.

## III. AN ENCODER WITH MORE INFORMATION

In the feedback system of Figure 1, the encoder has access only to the sequence of plant outputs. Suppose instead, as shown in Figure 2, that the encoder has access to perfect measurements of the plant state, the plant input, and feedback from the channel output. We now propose a strategy for estimating the message $m_N$ under the assumption that the encoder has this additional information.
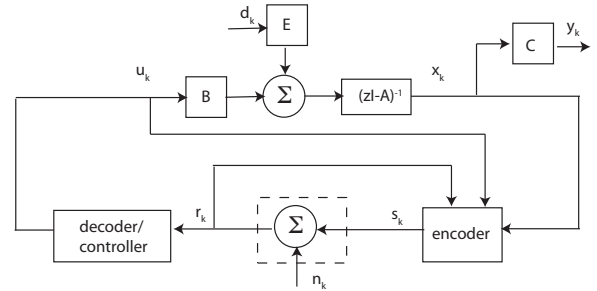
Fig. 2. Encoder with access to the plant state, plant input, and channel output.

We first show that access to the plant state and control input allows us to compute the primitive random variables $z_0 \triangleq CA^{N+1}x_0$ and $d_0, d_1, \ldots, d_{N-1}$. Clearly access to the state allows $z_0$ to be computed at $k = 0$. Access to the state and control input allows the encoder to determine the sequence of disturbance inputs. Choose an arbitrary row vector $F$ such that

233

$FE \neq 0$. Then the disturbance at time $k$ may be computed at time $k+1$:

$$d_k = (FE)^{-1}(Fx_{k+1} - FAx_k - FBu_k), \quad k = 0, \ldots, N-1.$$

The ability to compute the primitive random variables, in turn, allows the encoder to form an estimate of the message (6) at each time step:

$$m_0 = CA^{N+1}x_0 \tag{8}$$
$$m_1 = m_0 + CA^N E d_0 \tag{9}$$
$$\vdots$$
$$m_N = m_{N-1} + CAEd_{N-1}. \tag{10}$$

Note that $m_k$ is the best possible estimate of $m_N$ given the information available at time $k$: $m_k = \mathcal{E}\{m_N | z_0, d_0, \ldots, d_{k-1}\}$.

We next propose to transmit the estimates (8)-(10) over the communication channel, taking appropriate advantage of the noiseless feedback link to improve the quality of transmission. To do so, we adapt a technique for communicating over a channel with feedback that is described in [16, pp. 479-481] and [17, pp. 166-168], and depicted in Figure 3. The idea is
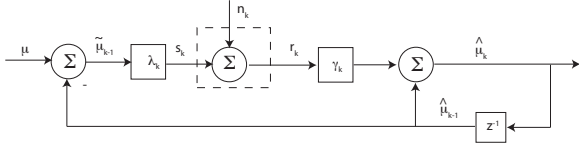


Fig. 3.   Communication over a channel with noiseless feedback.

to use a Gaussian channel $N+1$ times for the purpose of communicating a single message $\mu$, assumed to be a Gaussian random variable with zero mean and variance $\sigma_\mu^2$. Define the conditional estimate $\hat{\mu}_k \triangleq \mathcal{E}\{\mu | r^k\}, k = 0, \ldots, N$, the estimation error $\tilde{\mu}_k \triangleq \mu - \hat{\mu}_k$, and set $\hat{\mu}_{-1} = 0$. Then choosing $\lambda_k$ so that $\mathcal{E}\{s_k^2\} = P$, and $\gamma_k = (1/\lambda_k)(1 + \sigma_n^2/P)^{-1}$ results in an estimation error at time $N$ with variance $\mathcal{E}\{\tilde{\mu}_N^2\} = \sigma_\mu^2(1 + P/\sigma_n^2)^{-(N+1)}$, the minimum possible according to rate distortion theory [16], [17].

The scheme depicted in Figure 3 is not directly applicable to our situation, because the message $m_N$ is unavailable at the beginning of channel transmission, and we must instead transmit the estimates $m_k$. We thus modify the scheme in Figure 3 by noting that the sequence $m_k$ can be modeled as the response of a discrete integrator to initial condition $z_0$ and input $v_k = CA^{N-k}E d_k$, a white noise sequence with variance $\sigma_k^2 = (CA^{N-k}E)^2 \sigma_d^2$.

A Kalman filter [18] to estimate the state of the integrator $m_k$ given the sequence $r_k = \lambda_k m_k + n_k$ has the form

$$\hat{m}_{k|k} = \hat{m}_{k|k-1} + L_k(r_k - \lambda_k \hat{m}_{k|k-1}),$$
$$\hat{m}_{k+1|k} = \hat{m}_{k|k}$$

where

$$L_k = \frac{\lambda_k M_{k|k-1}}{\lambda_k^2 M_{k|k-1} + \sigma_n^2}, \tag{11}$$

and $M_{k|k-1} = \mathcal{E}\{\tilde{m}_{k|k-1}^2\}$ satisfies the Riccati difference equation

$$M_{k+1|k} = M_{k|k-1} - \frac{\lambda_k^2 M_{k|k-1}^2}{\lambda_k^2 M_{k|k-1} + \sigma_n^2} + \sigma_k^2, \tag{12}$$

with initial condition $M_{0|-1} = CA^{N+1}\Sigma_{0|-1}A^{(N+1)T}C^T$. Suppose we adjust $\lambda_k$ at each time step so that $\lambda_k^2 M_{k|k-1} = P$. Then (11) and (12) reduce to

$$L_k = \frac{1}{\lambda_k}\frac{P}{P + \sigma_n^2},$$
$$M_{k+1|k} = M_{k|k-1}\frac{\sigma_n^2}{P + \sigma_n^2} + \sigma_k^2.$$

The resulting estimation scheme may be implemented over a channel with noiseless feedback as shown in Figure 4.
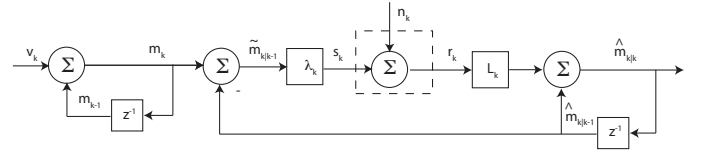


Fig. 4.   Communicating the output of a discrete integrator over a channel with feedback.

Note the similarity of the communication system in Figure 4 to that in Figure 3. In each figure $\lambda_k$ is chosen to satisfy the channel power constraint with equality. Once $\lambda_k$ is chosen, the formulas for $\gamma_k$ in Figure 3 and $L_k$ in Figure 4 are identical. The only difference is the presence of the disturbance sequence $v_k$ in Figure 4. With this disturbance present, the error in the estimate $\hat{m}_{N|N}$ may be obtained by iterating the Riccati equation (12):

$$\mathcal{E}\{\tilde{m}_{N|N}^2\} = M_{0|-1}\left(\frac{\sigma_n^2}{P + \sigma_n^2}\right)^{N+1} + \sum_{j=0}^{N-1}\sigma_j^2\left(\frac{\sigma_n^2}{P + \sigma_n^2}\right)^{N-j}. \tag{13}$$

Were this disturbance not present, the estimation error would equal that achieved with the procedure described in [16], [17].

We have shown how to use the additional information present at the encoder in Figure 4 to obtain an estimate of the "message" $m_N$ defined by (10), and whose estimation error has variance (13). As noted in the discussion preceding (10), the control signal is known at the decoder and may be used to obtain an estimate of $CAx_N$ from that of $m_N$; the estimation error will be the same for these two quantities. Hence substituting (13) into (5) and using the definitions of $M_{0|-1}$ and $\sigma_k^2$ shows that if we apply the optimal control (4) at time $k = N$, then the plant output will satisfy

$$\mathcal{E}\{\tilde{y}_{N+1|N}^2\} = CA^{N+1}\Sigma_{0|-1}A^{(N+1)T}C^T\left(\frac{\sigma_n^2}{P + \sigma_n^2}\right)^{N+1}$$
$$+ \sigma_d^2\sum_{j=0}^{N}(CA^{N-j}E)^2\left(\frac{\sigma_n^2}{P + \sigma_n^2}\right)^{N-j}. \tag{14}$$

Note that the control inputs at times $k = 0, \ldots, N-1$ are irrelevant and with no loss of generality may be set to zero.

234

## IV. OPTIMALITY OF LINEAR TIME-VARYING COMMUNICATION AND CONTROL

In Section III we increased the amount of information available to the encoder and derived linear time-varying communication and control strategies that resulted in the output at time $k = N+1$ having variance (14). It remains to determine whether nonlinear communication and control strategies may yield a smaller variance. Were the disturbance not present, the arguments from rate distortion theory used in [16], [17] would suffice to prove that there is no advantage to use of nonlinear strategies; however, our problem is more complicated due to the presence of the disturbance input. To address this question, we use arguments based on the concept of entropy power, introduced by Shannon in [19].

The entropy of a scalar zero mean Gaussian random variable $x$ with variance $\sigma^2$ is completely determined by its variance [20]: $h(x) = (1/2)\log 2\pi e\sigma^2$. For a non-Gaussian random variable, the relation between entropy and variance is only an inequality: $h(x) \leq (1/2)\log 2\pi e\sigma^2$. The latter fact complicates attempts to use entropy to study the propagation of variance through potentially nonlinear transformations. The entropy power of a random variable, defined as $N(x) = (1/2\pi e)e^{2h(x)}$, has the interpretation of being the variance of a Gaussian random variable with the same entropy as that of $x$ [19]. It follows that the entropy power of $x$ is a lower bound on the variance of $x$, with equality holding precisely when $x$ is Gaussian. The concept of entropy power and the associated entropy power inequality [20] were used in [2] to derive a lower bound on the variance of the (possibly vector valued) state of a plant under control over a noise-free data rate limited channel using a nonlinear time-varying encoder and decoder. The analysis of [2] was adapted in [21] to derive a similar lower bound for control over Gaussian channels. It is noted in [12], however, that bounds on variance obtained using entropy power arguments will be tight only in special cases, including that of a scalar system.

With the preceding for motivation, we have noted that with additional information at the encoder, the problem of using control to minimize variance at a terminal time reduces to that of sending estimates of $m_N$ over the channel. The encoding and decoding schemes we used to do so were linear and time-varying. By adapting the arguments of [21], however, it may be shown that nonlinear control and communication strategies offer no improvement. Indeed, denote the random conditional entropy [21] of the integrator state $m_k$ conditioned on past channel outputs by $h(m_k|r^{k-1})$, and the random conditional entropy power by

$$N(m_k|r^{k-1}) = (1/2\pi e)e^{2h(m_k|r^{k-1})}.$$

Let $n_k = \mathcal{E}\{N(m_k|r^{k-1})\}$. Then [20] $n_k \leq \mathcal{E}\{\tilde{m}_{k|k-1}^2\}$, and the results of [21] may be adapted to show that $n_k$ satisfies the difference inequality

$$n_{k+1} \geq \left(\sigma_n^2/(P+\sigma_n^2)\right)n_k + \sigma_k^2.$$

Since $m_0$ is Gaussian, $n_0 = M_{0|-1}$, and since $\tilde{m}_{k+1|k} = \tilde{m}_{k|k}$, it follows that

$$\mathcal{E}\{\tilde{m}_{N|N}^2\} \geq n_{N+1}$$

$$\geq M_{0|-1}\left(\frac{\sigma_n^2}{P+\sigma_n^2}\right)^{N+1} + \sum_{j=0}^{N-1}\sigma_j^2\left(\frac{\sigma_n^2}{P+\sigma_n^2}\right)^{N-j}. \tag{15}$$

Comparing (15) with (13), we see that use of nonlinear communication and control strategies is not effective at reducing the variance in the estimate of $m_N$ below the level achievable with linear strategies. By arguments similar to those used to derive (14), the same conclusion holds for the variance of the plant output $\mathcal{E}\{\tilde{y}_{N+1|N}^2\}$. Hence use of linear communication and control strategies is optimal for the system in Figure 2, whose encoder has access to additional information.

## V. THE ORIGINAL PROBLEM

We now return to the original problem depicted in Figure 1. We shall show that the performance achieved in Section III with *additional* information available at the encoder is in fact also achievable without such information. We saw in Section IV that nonlinear strategies offer no improvement when additional information is available. It follows that since we can achieve the same level of performance without the additional information, nonlinear strategies also offer no advantage for our original problem setup.

Our first result is to show that the encoder need not have access to the plant input or channel output to achieve the level of performance that was obtained using this information in Section III. The following result is proven in [22].

**Proposition 2** *Assume that the encoder has access to the states of the plant. Choose the channel input to satisfy*

$$s_k = \lambda_k H_k x_k, \tag{16}$$

*where*

$$H_k = CA^{N+1-k}, \qquad k = 0, \ldots, N, \tag{17}$$

*and, if $H_k\Sigma_{k|k-1}H_k^T \neq 0$,*

$$\lambda_k H_k\Sigma_{k|k-1}H_k^T = P, \qquad k = 0, \ldots, N. \tag{18}$$

*Choose the control law*

$$u_k = -(H_{k+1}B)^{-1}H_{k+1}A\hat{x}_{k|k}, \qquad k = 0, \ldots, N, \tag{19}$$

*with $\hat{x}_{k|k}$ given by*

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_k + AL_k(r_k - \lambda_k H_k\hat{x}_{k|k-1}),$$
$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k(r_k - \lambda_k H_k\hat{x}_{k|k-1}),$$

*where*

$$L_k = \lambda_k\Sigma_{k|k-1}H_k^T/(P+\sigma_n^2), \tag{20}$$

*and $\Sigma_{k|k-1}$ is the solution to the Riccati difference equation*

$$\Sigma_{k+1|k} = A\Sigma_{k|k-1}A^T$$
$$- \frac{A\Sigma_{k|k-1}H_k^T H_k\Sigma_{k|k-1}A^T}{H_k\Sigma_{k|k-1}H_k^T}\frac{P}{P+\sigma_n^2} + \sigma_d^2 EE^T.$$

235

*Then at time $k = N + 1$, the mean square value of the plant output satisfies $\mathcal{E}\{y_{N+1}^2\} = \mathcal{E}\{\tilde{y}_{N+1|N}^2\}$, where $\mathcal{E}\{\tilde{y}_{N+1|N}^2\}$ is given by* (14). ■

Our second result, also proven in [22], shows that under reasonable hypotheses access to the plant state is also unnecessary. Instead of measuring the states to form the channel input, we let the encoder have access to the state of an estimator whose input is a noise-free measurement of the plant output. In general estimators must have access to the plant input; as we shall see, this is not the case in the present situation. To distinguish these state estimates from those obtained by processing the channel output, we denote them by $\hat{x}_{k|k-1}^0 = \mathcal{E}\{x_k|u^{k-1}, y^{k-1}\}$ and $\hat{x}_{k|k}^0 = \mathcal{E}\{x_k|u^{k-1}, y^k\}$. The next result, taken from [22], shows that access to the plant input is unnecessary for the purpose of computing the estimate $\hat{x}_{k|k}^0$.

**Proposition 3** *Assume that $E = B$, that $CB \neq 0$, and that the transfer function $G(z) = C(zI - A)^{-1}B$ has no zeros outside the closed unit circle. Then the state estimate $\hat{x}_{k|k}^0$ may be obtained from the recursion*

$$\hat{x}_{k|k}^0 = A\hat{x}_{k-1|k-1}^0 + L_f(y_k - CA\hat{x}_{k-1|k-1}^0),$$

*where where $L_f = B(CB)^{-1}$. Furthermore, if $F_x(z)$ denotes the transfer function from the plant output $y_k$ to the state estimate $\hat{x}_{k|k}^0$, then $F_x(z)G(z) = (zI - A)^{-1}B$.* ■

It follows from Proposition 3 that *the response of the state estimate $\hat{x}_{k|k}^0$ to disturbance and control inputs is identical to the response of the system states to these signals*. It follows that if the estimator is initialized with the plant initial condition, $\hat{x}_{0|-1} = x_0$, then $\hat{x}_{k|k}^0 = x_k$, $\forall k$. If the initial plant state is unknown, then $\hat{x}_{k|k}^0 \to x_k$ as $k \to \infty$.

We have now shown that the level of performance achieved in Section III using an encoder that has access to the plant state, plant input, and channel output can be achieved without this additional information.

## VI. CONCLUSION

We have presented a solution to the problem of minimizing the variance of the plant output at a specified terminal time using measurements of the plant output that are obtained from a Gaussian channel in the feedback loop. Our solution involves linear time-varying communication and control strategies. There is no separation between the problems of communication and control, because the two strategies must be carefully synchronized, as shown in Proposition 2. Yet there is also no conflict, because the control input is chosen to aid in the estimation process up until the last time step, when estimation is no longer an issue and the control (4) may be used that sets the plant output equal to the estimation error. As examples presented in [22] show, the transient response associated with control and communication strategies designed to minimize variance at a terminal time may be very poor. This is unsurprising, and strategies that are suboptimal at the terminal time but possess better transient performance need

to be explored. Of course, penalizing transient as well as terminal variance will introduce conflict between control and estimation, and thus nonlinear strategies may prove superior.

### REFERENCES

[1] G. N. Nair and R. J. Evans, "Exponential stabilizabiity of finite-dimensional linear systems with limited data rates", *Automatica*, vol. 39, no. 2, pp. 585-593, April 2003.
[2] G. N. Nair and R. J. Evans, "Stabilizabiity of stochastic linear systems with finite feedback data rates", *SIAM Journal on Control and Optimization*, vol. 43, no. 2, pp. 413-436, July 2004.
[3] S. Tatikonda and S. K. Mitter, "Control under communication constraints", *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1056-1068, July 2004.
[4] J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg, "Feedback stabilization over signal-to-noise ratio constrained channels", *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1391-1403, August 2007.
[5] H. S. Witsenhausen, "Separation of Estimation and Control for Discrete Time Systems", *IEEE Transactions on Automatic Control*, vol. 59, no. 11, pp. 1557-1566, November 1971.
[6] H. S. Witsenhausen, "A counterexample in stochastic optimal control", *SIAM Journal of Control*, vol. 6, pp. 131-147, 1968.
[7] R. Bansal and T. Basar, "Stochastic Teams with Nonclassical Information Revisited: When is an Affine Law Optimal?", *IEEE Transactions on Automatic Control*, vol. 32, no. 6, pp. 554-559, June 1987.
[8] R. Bansal and T. Basar, "Solutions to a class of linear-quadratic-Gaussian (LQG) stochastic team problems with nonclassical information", *Systems and Control Letters*, vol. 9, pp. 125-130, 1987.
[9] N. C. Martins and M. A. Dahleh, "Feedback control in the presence of noisy channels: "Bode-like" fundamental limitations of performance", *IEEE Transactions on Automatic Control*, to appear, 2008.
[10] S. Tatikonda, A. Sahai, and S. K. Mitter, "Stochastic linear control over a communication channel" *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1549-1561, September 2004.
[11] N. Elia, "When Bode meets Shannon: Control-oriented feedback communication schemes" *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1477-1488, September 2004.
[12] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, "Feedback control under data rate constraints: An overview" *Proceedings of the IEEE*, vol. 95, no. 1, pp. 108-137, January 2007.
[13] A. S. Matveev and A. V. Savkin, "The problem of LQG optimal control via a limited capacity communication channel", *Systems and Control Letters*, vol. 53, no. 1, pp. 51-64, 2004.
[14] R. Bansal and T. Basar, "Simultaneous Design of Measurement and Control Strategies for Stochastic Systems with Feedback", *Automatica*, vol. 25, no. 5, pp. 679-694, 1989.
[15] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification, and Adaptive Control*, Prentice-Hall, New Jersey, 1986.
[16] R. G. Gallager, *Information Theory and Reliable Communication*, Wiley, New York, 1968.
[17] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*, Prentice-Hall, New Jersey, 1971.
[18] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall, New Jersey, 1979.
[19] C. E. Shannon, "Communication in the Presence of Noise", *Proceedings of the IRE*, vol. 37, no. 1, pp. 10-21, 1949.
[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd ed.* Wiley, New Jersey, 2006.
[21] J. S. Freudenberg, R. H. Middleton, and V. Solo, "The Minimal Signal-to-Noise Ratio Required to Stabilize over a Noisy Channel", *Proceedings of the 2006 American Control Conference*, Minneapolis MN, June 2006.
[22] J. S. Freudenberg, R. H. Middleton, and J. H. Braslavsky, "Minimum Variance Control over a Gaussian Communication Channel", *to appear in the Proceedings of the 2008 American Control Conference*, Seattle WA, June 2008.