Research Article

On Linear Coding Schemes for Stabilizing LTI Control with Multiple Sensors

Anna N. Kim

Department of Electronics and Telecommunication, Norwegian University of Science and Technology, 7491 Trondheim, Norway

Correspondence should be addressed to Anna N. Kim, annak@iet.ntnu.no

Received 11 August 2010; Accepted 12 November 2010

Copyright © 2010 Anna N. Kim. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We examine the problem of designing the encoding and control policies of a linear stochastic control system, where the communication channel between the plant state observer (sensor) and the controller is a lossy wireless channel that is constrained in terms of transmit power and bandwidth. For a first-order ARMA modeled plant with Gaussian statistics, when there are two sensors observing the plant, nonlinear encoding is shown to result in smaller cost at time instant $T = 1$ compared to the linear schemes, if transmissions are carried out over parallel Gaussian independent channels. In this paper, optimal linear coding schemes for the case of multiple sensors are examined. They are shown to minimize the control cost at the infinite time horizon, when the wireless channel is accessed using time division multiplexing. Our analysis is carried out for when separation between the state estimation and control is possible, and the optimal steady state control law is certainty equivalent. The distortion lower bound for estimating the plant state is derived, along with the necessary conditions on the transmit power that minimize the steady state control cost. We also propose a linear scheme that reaches the distortion bound asymptotically under relaxed conditions.

1. Introduction

Wireless sensors and communication have become an integral part of closed-loop control systems in the recent years. Consider for example, the following simple control application. We wish to regulate the temperature of a room. Multiple sensor nodes are placed at different locations in the room for taking temperature measurements. The measurements are then sent to the controller, which utilizes the sensor observations when issuing new control commands to the heating or cooling elements. Naturally, benefits such as flexible placement, reduced costs on installation, and easier maintenance can be expected if sensor nodes communicate to the controller via wireless channels. This, however, also means that the sensor nodes are now sharing the communication medium, which is lossy and in general also restricted in terms of bandwidth and power. Such information constraints make the networked control systems differ from the classic wired setting. As pointed out in [1], designing the optimal communication and control policies in this case is no simple task, since the changes in one directly influence the outcome of the other. The controller may have to tolerate large errors and/or delays in sensor observations, while the sensors may have to encode and transmit their measurements subject to strict controller requirements, in addition to those imposed by the wireless channel. The situation may be further complicated by the network topology if we take into account possible multihop and relaying between sensor nodes and sensor/controller. Although many challenges in optimization of communication and control in these scenarios remain open [1], if we restrict our attention to linear and stochastic plants, and if the control objective is to minimize the linear quadratic Gaussian (LQG) control cost, there are a few nice, established results.

In particular, suppose a memoryless additive Gaussian noise channel is between the plant state observation (performed by a single sensor), and the controller, and the plant is described by the first-order ARMA (autoregressive moving average) model. And further, suppose that the encoder at the sensor has access to the current plant state and all past channel outputs, while the controller only has access to all past and current channel outputs. Simultaneous optimization of the encoding and control strategy that minimizes the LQG cost with finite and infinite time horizon leads to linear schemes for both the encoder at the sensor node and the controller [2].
In practice, the sensor observation may experience noise. In the case that the observation noise is additive memoryless Gaussian, optimality of linear encoding and control still holds when the noisy observations are transmitted over parallel Gaussian channels [3]. If we extend this setting to one that includes two sensor nodes, we arrive at a rather different result.

When both sensors are used to measure the plant state with certain observation noise, their measurements are correlated. If these measurements (intercorrelated Gaussian) are transmitted over two parallel Gaussian channels, it was shown in [4] that for time horizon of \( T = 1 \), a nonlinear encoding scheme can result in a lower LQG cost, compared to the optimal linear strategy.

The findings of [4] is predominantly of a communication nature. The control applied to the plant is limited to time horizon \( T = 0 \) and is also linear. In this paper, we further extend the network topology to more than two sensors. The angle of our approach to the multisensor control problem is also rooted in communications. Namely, we ask ourselves the question “exactly how well can the linear coding schemes perform in such multisensor control system?” Naturally, the answer is trivial if we restrict the control horizon to \( T = 0 \), or the problem can be unmistakably complex if our aim is to simultaneously optimize communication and control for an infinite time horizon. We can, however, simplify our problem at hand by using the divide and conquer strategy. That is, we look at the case when communication (i.e., estimation of state) and control can be separated and then focus solely on the communication aspect of the problem. Note also that the term “communication” in this paper is confined to the application and physical layer of the communication protocol stack. Unquestionably, other layers can/will be involved when a more complex network is considered.

Our contribution in this paper is the following. We first identify the condition under which communication and control can be separated, when our goal is to minimize the control cost at the infinite time horizon. We then derive the distortion lower bound for multiple sensors observing the plant state and transmitting over the AWGN channel. We show when this distortion lower bound is achievable using linear coding schemes. Using these results, we present the necessary and sufficient conditions for stabilizing the plant. We then propose another linear coding scheme which achieves the minimum distortion asymptotically under relaxed conditions.

The basis of much of our findings is not new within the communications domain. In particular, linear coding and transmission with and without feedback have existed since the 1960s. However, further insights can be gained by taking these approaches and the new variations within the context of control. For instance, with a defined communication cost required to stabilize the control system, tradeoffs can be determined when designing implementable schemes. Note also that in this paper we restrict our attention to a simple first-order plant model. The main motivation is that certain issues regarding multisensor communication in this setting are not resolved until now. Our contribution is one step towards studying communication and control that involves a more complex plant. Keep in mind also that the more general linear systems constitute a significant portion of control systems, where the plant dynamics are more easily tractable. In addition, as we will show in the paper, linear control policies play a significant role in communication and control separation.

The rest of the paper is organized as follows. In Section 2, we present the proper formulation of our problem and show the conditions under which separation of communication and control can take place. The main results of the paper are summarized in Section 3. Numeric examples are given in Section 4 to provide further insights on how the different linear schemes perform under different sensor network conditions. We conclude the paper in Section 5 along with possible future work.

2. Problem Formulation and the Preliminaries

Figure 1 depicts the general structure of the closed-loop control system. The plant state \( s(t) \) at time instant \( t \) is observed by \( M \) sensors, which are placed physically apart over a defined area. Each sensor experiences observation noise \( w_m(t) \). The sensor measurements are encoded with the help of additional side information \( \bar{a}(t-1) \) to produce channel input symbols \( x_m(t) \). The side information \( \bar{a}(t-1) \) constitutes what is available at the controller prior to receiving the current channel output \( y_m(t) \). It is made available at the encoders via a noiseless feedback channel. The exact content of \( \bar{a}(t) \) will be discussed later. On the receiving side of the lossy wireless communication channel, we consider a joint decoder that processes all channel outputs. The decoder is set to be colocated with the controller, so it has access to the controller information. This is indicated by the dotted arrow to the decoder. The decoder output \( z(t) \) is used by the controller to produce the control signal \( r(t) \), which is subsequently fed back to the plant.

We consider a discrete time, LTI (linear-time-invariant) scalar plant that is defined by the following state equation:

\[
s(t + 1) = \gamma s(t) + r(t) + d(t), \quad \forall t \geq 0,
\]

where \( \gamma \) is a known constant, \( d(t) \) is the plant disturbance and \( s(t), d(t), r(t) \in \mathbb{R} \). Both \( s(0) \) and \( d(t) \) are independently and identically distributed (i.i.d.) zero mean memoryless Gaussian random variables (r.v.’s) of variance \( \sigma_S^2(0) \) and \( \sigma_D^2 \), respectively. Note that we do not restrict \( \gamma \), so the plant may be unstable without control.
Figure 2 illustrates the communication part of the closed-loop control system in detail. Here we have

\[ u_m(t) = s(t) + w_m(t) \quad m = 1, 2, \ldots, M, \]

where \( u_m(t) \) is the observation of \( s(t) \) at sensor \( m \). The observation noise \( w_m(t) \) is zero mean, memoryless Gaussian with variance \( \sigma_m^2 \) and is independent of \( s(t) \). In addition, the observations \( u_m(t), m = 1, 2, \ldots, M \) are conditionally independent of each other, given \( s(t) \). Each sensor node is equipped with its own encoder, denoted by function \( F_m(\cdot) \). So \( x_m(t) = F_m(u_m(t), \tilde{a}(t - 1)) \). We consider an information pattern where \( \tilde{a}(t - 1) \) consists of past controls \( r_{0}^{t-1} \) and decoder output \( z_{0}^{t-1} \), where the notation \( r_{0}^{t-1} \) means \( [r(0), \ldots, r(t-1)] \). In other words, the encoder has full knowledge. Naturally, depending on the application, the amount of feedback information available at the encoder may vary, or the feedback quality itself may be limited. It is, however, also important to examine the system under more idealized scenario, so that we can determine the respective performance bounds. A brief discussion on other information pattern and feedback limitations is presented at the end of the paper.

The channel input \( x_m(t) \) from node \( m \) is subjected to an average transmit power constraint \( P_m \), which is defined as

\[ \frac{1}{T} \sum_{t=0}^{T} E[x_m(t)^2] \leq P_m. \]

The sensor nodes access the wireless channel using orthogonal transmission, for example, TDMA or FDMA (time or frequency division multiple access). This is illustrated in the figure as \( M \) parallel independent subchannels. Furthermore, we assume that if the plant produces the state signal at a rate of \( 2B \) symbols per second, that is, the state signal occupies a bandwidth of \( B \) symbols per second, then the wireless channel is said to have a total bandwidth of \( M \cdot B \), where each subchannel has bandwidth \( B \). The subchannels are corrupted by additive white Gaussian noise \( n_m(t) \) of zero mean and variance \( \sigma_n^2 \). The channel noises are also mutually independent. At the receiver side, the decoder output \( z(t) \) is produced using the decoding function \( \tilde{g} \), and past control signals. So \( z(t) = \tilde{g}(y_{m=1\sim M}(t), r_{0}^{t-1}) \).

We define the steady state Linear Quadratic Gaussian (LQG) cost of the above control system as

\[ J = \lim_{T \to \infty} \sup_{T} \frac{1}{T} E \left[ \sum_{t=0}^{T-1} s^2(t) + \kappa r^2(t) \right] \]

for some \( \kappa > 0 \). Our goal is to design the encoder-controller pair, such that the cost function is minimized for the plant defined by (1), under the above specified communication constraints.

Before proceeding further, we will first simplify our problem by identifying the optimal control policy.

**Proposition 1.** The optimal control for the system is the certainty equivalence control law, given that the encoding functions \( F_m \) are linear.

![Figure 2: A sensor network with M sensors and orthogonal transmission.](image_url)

**Proof.** The proof follows the same argument as shown in [5]. We present here a brief summary of the main points.

For the plant in (1), under full state observation, that is, when the controller has direct access to \( s(t) \), the optimal steady state control law is the certainty equivalent (CE) control law [6]. That is, the control signal is the state signal with a linear gain: \( r(t) = l \cdot s(t) \), where

\[ l = -\frac{hy}{h + \kappa}, \]

and \( h \) satisfies the scalar Riccati equation:

\[ h = 1 + \frac{y^2h\kappa}{h + \kappa}. \]

The communication channel from the sensors to controller turns the fully observed plant into a partially observed one. In order for the CE control law also to be optimal for system (1), that is, \( r(t) = l \cdot z(t) \), where \( z(t) \) is the decoder’s estimate of the plant state, the control must have no dual effect (as stated in [5, Proposition 3.1]). This means that the estimation error variance of the state signal is independent of the control applied for all \( t \). Memoryless Gaussian channels with any linear encoders in fact satisfy the necessary conditions for insuring no dual effect of the control signal, as shown in [5, Lemma 3.1]. Subsequently, for system (1), the optimal control remains to be certainty equivalent as given by (5).

The partially observed control problem is then reduced to a fully observed one, by modeling the estimated state signal as the new state:

\[ z(t+1) = yz(t) + r(t) + \tilde{a}(t), \]

where \( \tilde{a}(t) \) is the new information being transmitted over the channel.

And the optimal control is

\[ r(t) = -\frac{hy}{h + \kappa} z(t) = -\frac{hy}{h + \kappa} \tilde{s}(t), \]

where \( \tilde{s}(t) = E[s(t) | y_{m=1\sim M}(t), r_{0}^{t-1}] \).
Define $\Delta(t, M) = \mathbb{E}[|s(t) - \hat{s}(t)|^2]$. The cost function can be rewritten as [5]

$$\tilde{J}(\mathcal{F}_{m=1-M}, \tilde{g}) = \frac{h^2 \gamma^2}{h + \kappa} \Delta_{t-\infty}(t, M) + h\sigma_d^2.$$  \hfill (9)

Since the second term is a constant that depends on the control system parameters $\gamma, \kappa$ and plant disturbance $\sigma_d^2$. Our objective of minimizing the cost function as given in (9) is met, if there exists linear encoding schemes that in fact minimizes $\Delta(t, M)$, for all $t$, under the given communication constraints. We will show next that this is indeed possible.

3. Summary of Results

In this section, we present a summary of our findings in the form of lemmas and theorems, along with the necessary proofs.

We begin by analyzing the source of observation in the sensor network, $s(t)$. Given that the side information $\tilde{a}(t-1) = [z_0^{t-1}, z_{t-1}^{l-1}]$ is available at the sensor nodes, each encoder can compute the following:

$$s(t) + w_m(t) - \mathbb{E}[s(t) | z_0^{t-1}] = \gamma(s(t-1) - \bar{s}(t-1))$$
$$+ d(t-1) + w_m(t)$$
$$= \gamma\lambda(t-1) + d(t-1) + w_m(t),$$  \hfill (10)

where $\lambda(t-1) = s(t-1) - \bar{s}(t-1)$ is the estimation error of the state signal at instant $t-1$.

Define the innovation process $i(t)$ as

$$i(t) = \gamma\lambda(t-1) + d(t-1).$$  \hfill (11)

The communication scenario we consider is thus equivalent to having the sensors observing the innovation $i(t)$, without any side information. As the coding operations we consider here are linear, we can assume therefore, without loss of generality, that the innovation is memoryless and Gaussian distributed with zero mean and variance $\sigma_i^2(t)$. It is then sufficient to evaluate $\Delta(t, M)$ in terms of the innovation's variance.

Lemma 1. Given a sensor network as depicted in Figure 2, at time $t$, the distortion $\Delta(t, M)$ is lower bounded by $\Delta_i(t, M)$:

$$\Delta_i(t, M) = \frac{\sigma_i^2(t)}{1 + \sigma_i^2(t) \sum_m \frac{1}{\sigma_{w_m}^2(t)}}$$
$$\times \left(1 + \frac{\sigma_i^2(t) \sum_m \frac{1}{\sigma_{w_m}^2(t)}}{\prod_m \left(1 + P_m/\sigma_d^2\right)}\right),$$  \hfill (12)

where $\sigma_i^2(t)$ is the variance of innovation.

This statement remains true even when arbitrary collaborations between the sensor nodes are allowed or when feedbacks are present at some or all sensor nodes.

Proof. We prove the above distortion lower bound by considering an idealized communication scenario. This approach has been applied in other context (see [7, 8]).

Let the $M$ sensor nodes be connected by wires or ideal links (ideal links in the communication sense refer to lossless channels with infinite capacity,), so that full collaboration is possible. This is equivalent to replacing the sensors with a single “fusion node”. We have then a point-to-point (p2p) channel between the fusion node and the decoder. The resulting structure constitutes the so-called “remote sensing” problem [9]. The optimal performance of a p2p communication system is characterized by equality between the rate-distortion function of the source and the channel capacity $R(D_r) = C$. The $M$ parallel memoryless Gaussian subchannels, now fused into one, has a total capacity of

$$C = \sum_{m=1}^{M} \log_2 \left(1 + \frac{P_m}{\sigma_d^2}\right).$$  \hfill (13)

The distortion-rate function of the remote sensing problem is [9]

$$D_r(t, R) = \frac{\sigma_i^2(t)}{1 + \Psi \left(1 + 2^{-2R\Psi}\right)},$$

where $\Psi = \sigma_i^2(t) \sum_m \frac{1}{\sigma_{w_m}^2(t)}.$  \hfill (14)

Replacing $R$ with $C$ in the above equation, we arrive at the claimed formula.

Remark 1. If the sensor observations are perfect, that is, $w_m(t) = 0$, and if all the sensors have the same transmit power constraint, that is, $P_m = P$, then the distortion lower bound becomes

$$\Delta_i^p(t, M) = \frac{\sigma_i^2(t)}{(1 + P/\sigma_d^2)^\Psi}.$$  \hfill (15)

This is a known result which describes the optimal performance theoretically attainable (OPTA) for transmitting a memoryless Gaussian source over a memoryless Gaussian channel that has $M$ times higher bandwidth [10]. When $M = 1$, linear encoding, that is, simple scaling of the source symbols to meet the transmit power constraint, is optimal. For $M > 1$, nonlinear coding schemes can easily outperform their linear counterparts. The optimal linear scheme, which is also referred to as BPAM (block pulse amplitude modulation) offers only 3 dB gain in source signal-to-distortion ratio (SDR), with each doubling of $M$ [11]. Nonlinear scheme such as that proposed in [12], which uses direct analogue mappings, offers significant gain in SDR compared to BPAM, especially when the channel signal-to-noise ratio (CSNR) is high. However, it is still 6 dB away from OPTA. In addition, it was proven in [13] that nonlinear coding schemes will have the quadratic distortion bounded away from $\Delta_i^p(t, M)$, as $M$ increases, when the source dimensionality is large but finite.

On the other hand, if the wireless channel is accessed via TDMA, causal feedback from the decoder can be utilized.
together with linear schemes. We will show next that this approach can be extended to sensors with noisy observations.

**Lemma 2.** In the case of TDMA, the distortion lower bound of (12) can be achieved using a linear scheme with noiseless causal feedback and full sensor collaboration.

**Proof.** Consider again that the sensor nodes are connected by wires to allow full collaboration. An MMSE (minimum-mean-squared-error) estimation of \( \hat{t}(i) \) can be made using all available noisy observations. Denote this estimation as \( \tilde{t}(i) \). And

\[
\tilde{t}(i) = \frac{E \left[ \hat{t}(i) \sum_{m=1}^{M} u_m(t) \right]}{E \left[ \left( \sum_{m=1}^{M} u_m(t) \right)^2 \right]} \sum_{m=1}^{M} u_m(t)
\]

which has variance:

\[
\sigma_{\tilde{t}}^2(t) = \frac{\sigma_i^2(t) \sum_{m=1}^{M} (1/\sigma_{wm}^2(t))}{\sigma_i^2(t) \sum_{m=1}^{M} (1/\sigma_{wm}^2 + 1)}.
\]

The variance of the estimation error is

\[
D_e(t, M) = \frac{\sigma_{\tilde{t}}^2(t)}{1 + \Delta^2 \gamma^2(t) \sum_{m=1}^{M} (1/\sigma_{wm}^2(t)).
\]

Note that this is also the term outside the parenthesis of (12).

Let one of the M sensors, call it sensor 1, transmit the estimation \( \tilde{t}(i) \) with power constraint \( P_1 \), using a linear scaling factor \( F_1(t) \). The decoder \( G_1(t) \) makes MMSE estimation of \( \tilde{t}(t) \) and feeds back the innovation \( \tilde{t}(t) - \hat{t}(t) \). At the decoder, the newly received and decoded innovation is added to the previous reconstruction to produce the signal for feedback, so

\[
\hat{\tilde{t}}(t) = \hat{t}(t) + G_2(t) \left( \tilde{t}(t) - \hat{t}(t) \right) + n_2(t).
\]

Repeat this process with the corresponding coding functions \( F_m(t) \) and \( G_m(t) \):

\[
F_m(t) = \sqrt{\frac{P_m}{D_m(t)}}, \quad G_m(t) = \frac{F_m(t) D_{m-1}}{P_m + \sigma_n^2},
\]

\[
D_m(t) = \frac{D_{m-1} \sigma_n^2}{P_m + \sigma_n^2}, \quad D_0(t) = \sigma_i^2(t).
\]

After \( M \) channel uses, we have

\[
D_M(t, M) = \frac{\sigma_{\tilde{t}}^2(t)}{\prod_{l=1}^{M} \left( 1 + P_m/\sigma_n^2 \right)}.
\]

Since the two distortions due to estimation and distortion from transmission are uncorrelated, we can simply sum \( D_e(t, M) \) and \( D_M(t, M) \) to get the end-to-end distortion, which is the same as \( \Delta(t, M) \) as desired.

**Remark 2.** The basic principle behind our proof can be traced back to [14–16], and it is also known as the Schalkwijk-Kailath scheme. By transmitting only the innovation at each channel use, optimal distortion can be reached for memoryless Gaussian source channels with integer channel to source bandwidth ratio. In other words, if \( w_m(t) = 0 \), the above described linear coding scheme will lead to \( \Delta^*(t, M) \) of (15). With each feedback, the contribution of the channel noise is reduced. It has, however, no effect on the observation noise. In fact \( D_e(t, M) \) is the absolute distortion lower bound of the \( M \) sensor network since this is same as having the fusion node wired directly to the decoder.

We are now ready to compute the steady state cost function \( J \) and determine the necessary conditions the cost is minimized. The result is the following theorem.

**Theorem 1.** The necessary condition to stabilize the scalar LTI control system of (1) is that the transmit power \( P_m \) of each sensor \( m \) satisfies

\[
\prod_{l=1}^{M} \left( 1 + \frac{P_m}{\sigma_n^2} \right) > \gamma^2.
\]

This condition is sufficient under TDMA with a linear sensing strategy, with full sensor collaboration and using noiseless causal feedback.

**Proof.** We need to determine the steady state distortion \( \Delta(t, M) \) and identify the conditions when it is nonnegative and bounded.

From (11), we see that the variance of innovation at time \( t \) is

\[
\sigma_{\tilde{t}}^2(t) = \gamma^2 \Delta(t, M) + \sigma_n^2.
\]

Substitute \( \sigma_{\tilde{t}}^2(t) \) to (12) of Lemma 1. At the steady state, \( \Delta(t, M) \) and \( \Delta(t-1, M) \) become \( \Delta(t, \infty) \). Equation (12) can then be rewritten as the following second-order polynomial:

\[
C_1 \Delta(t, \infty)^2 + C_2 \Delta(t, \infty) + C_3 = 0,
\]

where

\[
C_1 = \gamma^2 \sum_{l=1}^{M} \frac{1}{\sigma_{wm}^2(l)} \left( 1 - \frac{\gamma^2}{\Phi} \right),
\]

\[
C_2 = 1 - \gamma^2 + \sigma_n^2 \left( 1 - \frac{2\gamma^2}{\Phi} \right) \sum_{l=1}^{M} \frac{1}{\sigma_{wm}^2},
\]

\[
C_3 = -\sigma_n^2 \frac{\gamma^4 \sum_{l=1}^{M} (1/\sigma_{wm}^2)}{\Phi}.
\]

To ensure that there exists a nonnegative root, we need \( C_1 > 0 \) and thus \( \Phi > \gamma^2 \), which is (22) as expected.

The sufficient condition follows directly from Lemma 2.

**Remark 3.** This result in fact coincides with the necessary condition for stabilizing system (1), when encoders observe
\( s(t) \) directly without noise and transmit over a memoryless Gaussian channel. In other words, the observation noise at the sensor nodes does not contribute to the overall information exchanged within the control system. Therefore, ideally, no communication resources should be allocated to them.

Remark 4. To achieve \( \Delta_i(t, M) \) using the linear scheme described in Lemma 2, it is required that (1) the signalling rate of the wireless channel is \( M \) times higher than the sampling rate of \( s(t) \), and (2) full cooperation is possible between the sensor nodes. The first requirement is inherent of time division multiple access, which is in general not a concern. The second requirement can be realized by either connecting the sensor nodes by wire or allocating additional transmit power and bandwidth for inter-sensor communications. However, if either wiring is unfeasible or when long sensor network lifetime is desired, other means of coding and transmission have to be considered. Next, we present a simple alternative which is also linear.

**Theorem 2.** Under TDMA, there exists a linear coding scheme which does not require collaboration between the sensor nodes and is able to reach \( \Delta_i(t, M) \) asymptotically, when either \( M \to \infty \) or when \( P_m/\sigma_n^2 \gg \sigma_i^2(t)/\sigma_{w_n}^2(t) \).

**Proof.** The basic scheme is similar to that given in the proof for Lemma 2, in the sense that noiseless causal feedback is required. The main difference is that here the feedback is made available to each sensor prior to its channel use, instead of the ideal fusion node. This means that we can effectively utilize the same feedback channel that made side information \( \tilde{a}(t) \) available at the sensor nodes.

The \( M \) sensors take turns in transmitting the difference between its observation and the previous reconstruction from the feedback channel, subject to its transmit power constraint. At the decoder, the newly received and decoded innovation is summed together with the previous reconstruction to produce the current estimation of the source. The set of linear coder/decoders are

\[
\tilde{P}_m(t) = \frac{P_m}{\sigma_i^2(t) + \sigma_{w_n}^2(t)}, \quad \tilde{G}_m(t) = \frac{P_m}{\sigma_i^2(t) + \sigma_{w_n}^2(t)},
\]

\[
\tilde{D}_m(t) = \frac{\tilde{D}_{m-1}(t)\sigma_{w_n}^2(t)}{\tilde{D}_{m-1}(t) + \sigma_{w_n}^2(t)} \left( 1 + \frac{\tilde{D}_{m-1}(t)}{\sigma_{w_n}^2(t)(1 + P_m/\sigma_n^2)} \right)
\]

and \( \tilde{D}_0(t) = \sigma_i^2(t) \).

When \( P_m/\sigma_n^2 \gg \sigma_i^2(t)/\sigma_{w_n}^2(t) \), we can approximate \( \tilde{D}_m(t) \) as

\[
\tilde{D}_m(t) \approx \frac{\tilde{D}_{m-1}(t)\sigma_{w_n}^2(t)}{\tilde{D}_{m-1}(t) + \sigma_{w_n}^2(t)}.
\]

After \( M \) transmissions, we have

\[
\tilde{D}_M(t) = \frac{\sigma_i^2(t)}{1 + \sigma_i^2(t) \sum_{t=1}^{M} \left( 1/\sigma_{w_n}^2(t) \right)},
\]

which equals \( \Delta_i(t, M) \) under the same conditions.

When \( M \to \infty \), from (12) it is clear that the distortion lower bound approaches 0. Rewrite (27) as \( \tilde{D}_m(t) = f(\tilde{D}_{m-1}(t)) \). The fixed point \( p \) of the iterative function \( f(\cdot) \) is when \( p = f(p) \). A direct solution of \( p \) shows that \( \tilde{D}_m(t) \) goes to 0 as \( M \to \infty \). If we apply the convergence test by evaluating \( |f'(p)| \), where \( f'(\cdot) \) is the first-order derivative of the function, we can see that \( \tilde{D}_m(t) \) converges linearly to 0 as \( M \to \infty \), instead of the exponential convergence in \( \Delta_i(t, M) \).

4. Numerical Examples

We evaluate the different linear coding schemes from the previous section using numerical examples and compare them with the distortion bound \( \Delta_i(t, M) \). In addition, we will examine how observation noise and size of the sensor network \( M \) affect the overall performance. For sake of simplicity and easier graphical interpretation, we set \( P_m = P \). And performance is assessed in terms of SDR (i.e., the inverse of the distortion normalized by the variance of \( i(t) \)) over a range of CSNR.

We first look at the simple case of two sensor nodes. To see the effect of observation noise \( w_1(t), w_2(t) \), we define the correlation coefficient between the sensor observations as

\[
\rho(t) = \frac{E[i(t)^2]}{E[(i(t) + w_1(t))(i(t) + w_2(t))]},
\]

so \( 0 < \rho \leq 1 \). The best linear scheme without causal feedback has an encoder that scales the source to satisfy the transmit power constraint, while at the decoder an MMSE estimation of \( i(t) \) is made using both received channel symbols. We have then the following end-to-end distortion:

\[
\delta(t, 2) = \sigma_i^2(t) - \frac{E[i(t)g_1(t)]^2}{E[y_1^2(t)]} - \frac{E[i(t)g_2(t)]^2}{E[y_2^2(t)]},
\]

If the observation noise has the same variance \( \sigma_n^2(t) \) then we can express \( \delta(t, 2) \) as

\[
\delta(t, 2) = \sigma_i^2(t) \frac{(1 - \rho)P + \sigma_n^2}{(1 + \rho)P} + \frac{\sigma_n^2}{\rho^2}.
\]

Note when \( \rho = 1 \), that is, perfect observation, we have

\[
\delta^*(t, 2) = \frac{\sigma_i^2(t)}{2P/\sigma_n^2 + 1}.
\]

So the best linear scheme is the same as transmitting the source over one channel with double the power.

We will see how “poorly” the best linear scheme without feedback performs relative to \( \Delta_i(t, M) \) for various \( \rho \). The result is shown in Figure 3. Interestingly enough, the largest gap to OPTA (that from \( \Delta_i(t, 2) \), (12)) occurs when correlation is very high, or in other words, when the observation noise variance is very small relative to the variance of innovation. For smaller \( \rho \) (below 0.7) or noisier observations, one may well choose the simple linear scheme without much sacrifice in performance.
For a network with more than 2 sensors, we compare the performance of the linear schemes proposed in Theorem 2 using causal feedback, along with the best linear scheme without feedback, relative to OPTA for different $M$. Let $\sigma_i^2(t) = 2$ and the observation noise variance $\sigma_{w_i}^2(t)$ be uniformly distributed between 0 and 1. We present the SDR versus CSNR results in Figure 4. As expected, the linear scheme with feedback closes the gap to OPTA significantly compared to the one without feedback in very low CSNR. When CSNR is at 15 dB, the gap to OPTA is already negligible for $M = 5$. This overlapping point moves further in the direction of lower CSNR as $M$ increases.

5. Conclusions and Future Work

In this paper, we surveyed a few established coding and transmission techniques in communication and re-examined their functions within the framework of closed-loop control. We studied linear coding schemes for an LQG control with $M$ sensors for plant observation and an AWGN communication channel between the sensors and controller of constrained bandwidth and transmit power. We showed that, contrary to the counter example given in [4], when the orthogonal channel access scheme is TDMA, causal feedback can be utilized so that optimal and asymptotically optimal linear coding schemes exist for the encoders at the sensor nodes. Since plant estimation and control can be separated, the associating control policy in this case is the certainty equivalent law. We also showed that when the encoders have full information, that is, both the past channel outputs and control signals, the necessary powers that minimize the steady state LQG cost correspond to that when there is no observation noise at the sensor nodes.

Naturally if the channel access scheme is for example FDMA, causal feedback will not be feasible, and the above mentioned linear schemes will be reduced to that in (31). However, as we pointed out in the numerical example, gains of nonlinear encoding schemes are only meaningful when observations are highly correlated. Another issue with the use of causal feedback is the quality of the feedback channel. If the feedback channel is bandwidth limited and has constrained power, losses are expected, as noted in [17], for the case of memoryless Gaussian source with no observation noise. In the remote sensing scenario as we consider here, the effect of noise from the feedback channel should be analyzed while taking the observation noise and the number of sensors into account. As we can see from (12), it is the term outside the parenthesis that dominates the end-to-end distortion.

We would also like to note that, if the wireless channel is a Gaussian multiple access channel (MAC), then linear coding scheme is again optimal when the sensor network is symmetric, that is, equal transmit power and equal observation noise variance for all sensor nodes [18]. The sensor nodes only need to scale their observations to satisfy the power constraint, while at the decoder side, MMSE estimations of the source is made from the sum of the received channel symbols.

The information pattern we consider in this work is one that results in full information at the encoders. If, for example, that the prior channel outputs are not available at the encoders, as shown in [5], that linear scheme that calculates the innovation without prior channel outputs will not be able to stabilize an unstable plant due to error propagation of unsynchronized encoder decoder. For an already stable plant, to minimize the LQG cost in this case requires coding schemes that operate across correlated sensor observations and along the correlation in time of the plant state. Keep in mind that, like FDMA, causal feedback is not applicable here. The best possible linear coding scheme is that discussed in [19]. How to incorporate in the control setting is part of our ongoing work.
Acknowledgment

The author is grateful to Professor Tor Arne Johansen of Department of Engineering Cybernetics, NTNU for valuable discussions.

References

[1] 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, 2007.
[2] R. Bansal and T. Başar, “Simultaneous design of measurement and control strategies for stochastic systems with feedback,” Automatica, vol. 25, no. 5, pp. 679–694, 1989.
[3] R. Bansal and T. Başar, “Solutions to a class of linear-quadratic-Gaussian (LQG) stochastic team problems with nonclassical information,” Systems and Control Letters, vol. 9, no. 2, pp. 125–130, 1987.
[4] S. Yüksel and S. Takikonda, “A counter example in distributed optimal sensing and control,” IEEE Transactions on Automatic Control, vol. 54, no. 4, pp. 841–844, 2009.
[5] S. Tatikonda, A. Sahai, and S. Mitter, “Stochastic linear control over a communication channel,” IEEE Transactions on Automatic Control, vol. 49, no. 9, pp. 1549–1561, 2004.
[6] D. P. Bertsekas, Dynamic Programming and Optimal Control, Athena Scientific, Belmont, Mass, USA, 3rd edition, 2007.
[7] M. Gastpar and M. Vetterli, Source-Channel Communication in Sensor Networks, Lecture notes in Computer Science, Springer, Berlin, Germany, 2003.
[8] H. Behroozi, F. Alajaji, and T. Linder, “On the optimal power-distortion region for asymmetric Gaussian sensor networks with fading,” in Proceedings of IEEE International Symposium on Information Theory (ISIT ’08), pp. 1538–1542, July 2008.
[9] R. Dobrushin and B. S. Tsybakov, “Information transmission with additional noise,” IRE Transactions on Information Theory, vol. 8, no. 5, pp. S293–S304, 1962.
[10] T. Berger, Rate Distortion Theory: Mathematical Basis for Data Compression, Prentice Hall Series in Information System Sciences, Prentice-Hall, Upper Saddle River, NJ, USA, 1971.
[11] K. H. Lee and D. P. Petersen, “Optimal linear coding for vector channels,” IEEE Transactions on Communications, vol. 24, no. 12, pp. 1283–1290, 1976.
[12] F. Hekland, P. A. Floor, and T. A. Ramstad, “Shannon-Kotel’nikov mappings in joint source-channel coding,” IEEE Transactions on Communications, vol. 57, no. 1, pp. 94–105, 2009.
[13] A. Ingber, I. Leibowitz, R. Zamir, and M. Feder, “Distortion lower bounds for finite dimensional joint source-channel coding,” in Proceedings of the IEEE International Symposium on Information Theory (ISIT ’08), pp. 1183–1187, July 2008.
[14] T. J. Cruise, “Achievement of rate-distortion bound over a additive white noise channel using a noisless feedback link,” Proceedings of the IEEE, vol. 55, pp. 583–584, 1967.
[15] J. P. M. Schalkwijk and L. I. Bluestein, “Transmission of analog waveforms through channels with feedback,” IEEE Transactions on Information Theory, vol. 13, no. 4, pp. 617–619, 1967.
[16] T. Kailath, “An application of Shannon’s rate-distortion theory to analog communication over feedback channels,” Proceedings of the IEEE, vol. 55, no. 6, pp. 1102–1103, 1967.
[17] A. N. Kim and T. A. Ramstad, “On bandwidth expansion with noisy feedback,” IEEE Communications Letters. In press.
