Signal Estimation over Channels with SNR Constraints and Feedback

Erik Johannesson

Automatic Control LTH, Lund University, Sweden
erik@control.lth.se

Abstract: We consider the problem of designing optimal linear encoders and decoders for estimation and transmission of a signal over an analog communication channel with additive noise and a signal-to-noise ratio (SNR) constraint. The objective is to minimize the variance of the estimation error at the receiving end, subject to a real-time constraint. It has been previously shown that, for channels without feedback, a solution can be obtained by solving a mixed norm minimization problem and performing a spectral factorization, see Johannesson et al. (2010b). In this paper, these results are extended to the case when the encoder has access to feedback from the channel output. This results in improved performance, which is demonstrated by a numerical example.

Keywords: Control under communication constraints, Control over networks, Real-time coding, Encoder-decoder design

1. INTRODUCTION

The problem studied in this paper lies in the intersection of estimation, communication and control theory. It can be seen as a modification of the Wiener-Kolmogorov filtering problem, as a real-time coding problem, or as a feed-forward compensator design problem. Depending on which aspect one wishes to focus on, the motivation of this paper can be drawn from each of these three fields.

The objective of the Wiener-Kolmogorov filtering problem is to estimate a signal that is measured with additive noise, under a mean square error criterion, see Hayes (1996). The design of the optimal estimation filter can be formulated in the frequency domain as the minimization problem:

\[ \| (z^{-k} - K)F \|^2_F + \| KG \|^2_F \]  

(1)

where \( k \) is the allowed time delay, \( F \) and \( G \) represent the frequency characteristics of the interesting signal and the measurement noise respectively, and \( K \) is the design variable.

This problem is here modified to a setting where the measurement and the estimation are performed at two different locations, see Fig. 1. The communication constraint between the two locations is modeled by an additive white noise (AWN) channel with feedback and a constraint on the signal-to-noise ratio (SNR). It is shown that the inclusion of a channel between the two parts of the filter induces an additional term in the cost (1).

As the present goal is to communicate a signal over a channel, under a real-time constraint, with minimal distortion, the problem under study may also be regarded as a communication problem. Information-theoretic tools seem to be of little use when there are time constraints, since classical communication theory does not worry about time delays, see Shannon (1948). It is also interesting, in this context, to note the well-known fact that the capacity of a discrete memoryless channel does not increase if feedback becomes available, see for example Massey (1990).

However, real-time coding problems have lately been studied with increasing interest. See for example Mahajan and Teneketzis (2009) for an overview. Here, the real-time communication problem is approached using more control-oriented methods.

It is worth noting that there are cases when the real-time constraint is without importance. For example, under certain conditions, it turns out that optimality can be achieved without coding. For example, this is the case when a white, Gaussian source is to be sent over an additive white Gaussian noise (AWGN) channel with a mean square error criterion, see Gastpar (2002). When the source is generated by a linear filter it may be enough
to send scaled innovations over an AWGN channel, see Martins et al. (2007).

These examples are somewhat counter-intuitive since a large allowed time delay usually makes the communication problem much easier in practice. Here, both the introduction of additive measurement noise at the coder and the access to feedback from the channel output makes the real-time constraint important. The measurement noise gives an incentive to filter the signal at the same time that it is coded, and the benefit of the feedback also grows with time.

In a control perspective, the problem can be interpreted as that of designing a feed-forward compensator in a networked control setting. A remote sensor/encoder has access to noisy measurements of the disturbance that is to be counteracted. The measurements are filtered and information about the disturbance is transmitted over the communication channel to the decoder/controller, which in turn can compensate.

1.1 Problem Description and Main Result

The block diagram in Figure 1 gives a schematic representation of the problem investigated in this paper. A signal is measured, together with some additive noise, at one location. An encoder is able to filter and encode information about the measurements and send it over an AWGN channel with a constraint on the SNR, to a decoder that forms an estimate of the signal. The encoder also has access to the past output of the channel. This may represent an actual feature of the communication channel, or constitute a model of the quantization error in a digital channel, which is known exactly by the encoder. See Silva et al. (2010b) for an example of the latter.

The task is to design the encoder and the decoder such that the estimation error is minimized. The estimation has to occur in real-time, as determined by the transfer function \( P \). Besides containing a fixed time delay, \( P \) may include general dynamics that shapes the signal that is to be estimated.

The main result of this paper is that the joint design of an optimal linear encoder-decoder pair can be formulated as a quasicone convex optimization problem followed by a spectral factorization. It is demonstrated that the access to feedback from the channel output improves performance.

1.2 Relations to Earlier Work

A lot of research efforts in the control community have recently been aimed at problems related to communication limitations. An overview of the research on networked control systems and control with data rate constraints, as well as a thorough list of references, can be found in Goodwin et al. (2008) and Nair et al. (2007) respectively. This paper is based on a modified version of the problem studied by Johannesson et al. (2010b). In that paper, the setup is the same as here except the encoder does not have access to feedback of the channel output. It was shown that the design problem can be formulated as a mixed \( \mathcal{H}_2 \) and \( \mathcal{H}_1 \) norm minimization problem, in which the relative weight of the two norms is determined by the channel capacity. These results have been generalized to the MIMO case by Johannesson et al. (2010a). The solution method in this paper is also similar to the one used by Johannesson et al. (2011) to design optimal controllers with two degrees of freedom for control over an AWGN channel with an SNR constraint.

The problem setup, both here and in Johannesson et al. (2010b), is inspired by Martins et al. (2007), where information-theoretic concepts were used to find a lower bound on the reduction of entropy rate made possible by side information communicated through a channel with given capacity. Under stationarity assumptions, this was used to derive a lower bound, which is a generalization of Bode’s integral equation, on a sensitivity-like function. Even though the problem architectures are similar, there are some important differences: The main difference is that Martins et al. (2007) gives performance bounds for a general communication channel while our papers treat synthesis for specific channel model, namely AWGN channels with SNR constraints, with or without feedback. Furthermore, there are differences in the employed performance metrics: Here, the variance of the error is minimized. In Martins et al. (2007), a lower bound is achieved on the integral of the logarithm of a sensitivity-like function. Also, in Martins et al. (2007), a feedback controller is placed at the receiving end. On the other hand, the setup in Johannesson et al. (2010b) and this paper is generalized by the inclusion of measurement noise at the sensor as well as the possibility of general dynamics in \( P \).

1.3 Notation

Denote the unit circle by \( T \). We define on \( \mathbb{T} \) the Smirnov class of functions \( \mathcal{N}^+ \) and, for \( 1 \leq p \leq \infty \), the Lebesgue \( \mathcal{L}_p \) and Hardy spaces \( \mathcal{H}_p \), in the usual manner. For more details, consult standard textbooks such as Rudin (1986) and Garnett (1981).

For \( 1 \leq p < \infty \) and scalar transfer functions \( X \) and \( Y \), define the \( p \)-norms

\[
\|X\|_p = \left( \frac{1}{2\pi} \int_0^{2\pi} |X(e^{i\omega})|^p \, d\omega \right)^{1/p}
\]

and the quantity

\[
(X,Y) = \frac{1}{2\pi} \int_0^{2\pi} X^*(e^{i\omega})Y(e^{i\omega}) \, d\omega.
\]

A transfer function \( X \in \mathcal{H}_p \) is said to be outer if the set \( \{Xq \mid q \text{ is a polynomial in } z^{-1}\} \) is dense in \( \mathcal{H}_p \). If \( X \) is a rational transfer function, then this is equivalent to \( X(z) \neq 0 \) for \( |z| > 1 \).

Equalities and inequalities involving functions evaluated on \( \mathbb{T} \) are to be interpreted as holding almost everywhere on \( \mathbb{T} \). That is, the subset of \( \mathbb{T} \) in which the (in)equality does not hold is of measure zero. The arguments to transfer functions will sometimes be omitted when these are clear from the context.

2. PROBLEM FORMULATION

Consider the system structure in Fig. 2. All the blocks represent LTI systems and the input signals are mutually independent white noise sequences with zero mean and
identity variance. The goal is to estimate the signal \( f \) after it has passed through \( P \), which in the nominal case is a fixed time delay but may be a general LTI system. The filters \( F \) and \( G \) are shaping filters for the signal \( f \) and the measurement noise, respectively.

![Diagram of system](image)

Fig. 2. Structure of the system. With \( F, G \) and \( P \) given, the objective is to design \( B, C \) and \( D \) such that the estimation error \( E(e^2) \) is minimized. The dashed box represents the encoder.

The communication channel is assumed to be an AWN channel with a constraint on the transmission power. Specifically,

\[
r(k) = t(k) + n(k), \quad E(t(k)^2) \leq \sigma^2
\]

where \( k \) is the time index, \( t \) is the transmitted variable, \( r \) is the received variable, \( n \) is the channel noise, and \( \sigma > 0 \) determines the maximum instantaneous transmission power. Since the transmission power constraint is equivalent to an SNR constraint, see Silva et al. (2010a), we shall refer to it as the SNR constraint.

The decoder \( D \) takes the channel output \( r \) as input and produces the estimate of \( Pf \). The output of the encoder, which consists of \( C \) and \( B \), is the transmitted signal \( t \). The encoder has two inputs: The measurements \( f + Gv \), where \( Gv \) is additive measurement noise, is the input to \( C \). The other input is delayed by one time step. Since the encoder remembers its past output, it can subtract \( t \) from \( r \) to obtain the channel noise \( n \), delayed by one time step, which is the input to \( B \). By linearity, we can assume that the encoder has this structure without loss of generality.

We assume that \( F, G, P \in \mathcal{H}_\infty \) and that

\[
\exists \epsilon > 0 \text{ such that } |F|^2 + |G|^2 \geq \epsilon
\]

that is, \( F \) and \( G \) have no common zeros on the unit circle. It is also assumed that the whole system is in stationarity.

For given \( F, G \) and \( P \), the objective is to design stable and causal \( B, C \) and \( D \) so that the stationary variance of the estimation error \( e \) is minimized. We expect to find the optimal linear \( B, C, D \), but make no claim that linear solutions are optimal per se.

From Fig. 2, we see that

\[
t = CFu + CGv + Bz^{-1}n
\]

\[
e = (P - DC)F u - DCGv - D(1 + Bz^{-1})n
\]

To conclude, the problem of minimizing \( E(e^2) \) subject to \( E(t^2) \leq \sigma^2 \), is formulated as:

**Problem 1:**

\[
\text{minimize}_{B, C, D} \left\| (P - DC)F \right\|_2^2 + \left\| DCG \right\|_2^2 + \left\| D(1 + Bz^{-1}) \right\|_2^2
\]

subject to

\[
\|CF\|_2^2 + \|CG\|_2^2 + \|B\| \leq \sigma^2
\]

3. **SOLUTION**

The objective (4) is not convex in \( B, C, D \). However, it will be shown in this section how to convert Problem 1 into a quasiconvex optimization problem.

This section is divided into three subsections. In the first, it is shown how to find the optimal \( C \) and \( D \) if their product \( DC \) and \( B \) are given. In the second subsection, this result is used to show equivalence between problem 1 and the minimization of another functional. Finally, in the third subsection, it is demonstrated how to implement a numerical solution.

### 3.1 Optimal Factorization

Note that the two first terms in (4) only depend on the product \( DC \). Hence, the idea is to consider \( K = DC \) and \( B \) as fixed and find a factorization of \( K \) that minimizes the remaining term, \( \|D(1 + Bz^{-1})\|_2^2 \), subject to the channel constraint (5). Given the value of the third term in (4) in terms of \( K \) and \( B \), it will be possible to optimize over \( K \) and \( B \). A factorization of the optimal \( K \) will then give the optimal \( C \) and \( D \).

In order to simplify the proceeding exposition, we will introduce the transfer function \( H \), as follows. By (3) and Theorem 4 (in appendix), there exists an outer function \( H \in \mathcal{H}_2 \) such that

\[
|H|^2 = |F|^2 + |G|^2
\]

and thus

\[
\|CH\|_2^2 = \|CF\|_2^2 + \|CG\|_2^2.
\]

Since \( F, G \in \mathcal{H}_\infty \) it follows that \( H \in \mathcal{H}_\infty \). Moreover, since \( H \) is outer it follows from (3) that \( H^{-1} \in \mathcal{H}_\infty \) as well.

For given \( B \in \mathcal{H}_2 \) satisfying \( \sigma^2 - \|B\|_2^2 = \alpha > 0 \) and a given product \( K = DC \in \mathcal{L}_1 \), we will now solve the problem of finding the optimal factorization. The solution is given by the following lemma. Note that it is not possible to obtain the minimum if it is required that \( D \in \mathcal{H}_2 \). Instead we must allow \( D \in \mathcal{N}^+ \). See Section 5 for a discussion of the implications.

**Lemma 1.** Suppose \( \alpha > 0 \), \( K \in \mathcal{H}_1, B \in \mathcal{H}_2, \) and \( H \in \mathcal{H}_\infty \) with \( H^{-1} \in \mathcal{H}_\infty \). Then the minimum

\[
\min_{C \in \mathcal{H}_2, D \in \mathcal{N}^+} \left\| D(1 + Bz^{-1}) \right\|_2^2
\]

subject to the constraints

\[
K = DC, \quad \|CH\|_2^2 \leq \alpha
\]

is attained. The minimum value is

\[
\frac{1}{\alpha} \left\| KH(1 + Bz^{-1}) \right\|_1^2.
\]

Moreover, if \( K = 0 \), then the minimum is achieved by \( D = 0 \) and any \( C \in \mathcal{H}_2 \) that satisfies (8). If \( K \) is not identically zero, then \( C \) and \( D \) are optimal if and only if \( C \in \mathcal{H}_2, D = KC^{-1} \in \mathcal{N}^+ \) and

\[
|C|^2 = \frac{\alpha}{\|KH(1 + Bz^{-1})\|_1} \quad \text{on } T.
\]
Proof. If $K = 0$, then the proof is trivial, so assume that $K$ is not identically zero. Then $C$ is not identically zero and $D = KC^{-1}$. Cauchy-Schwarz’s inequality gives
\[
\|D(1 + Bz^{-1})\|_2^2 \|CH\|_2^2 \geq |\langle KC^{-1}(1 + Bz^{-1})\rangle|, |CH|_2^2 = \|KH(1 + Bz^{-1})\|_1^2.
\]
This shows that (9) is a lower bound on (7). Equality holds if and only if $|\langle KC^{-1}(1 + Bz^{-1})\rangle| = \lambda |CH|$ for some $\lambda \in \mathbb{R}$ and $\|CH\|_2^2 = \alpha$. It is easily verified that this is equivalent to the optimality condition (10).

It only remains to verify the existence of $C \in \mathcal{H}_2$ and $D = KC^{-1} \in N^+$ such that (10) holds. Since $B$ is causal, $1 + Bz^{-1} \in \mathcal{H}_2$ is not identically zero and it follows from Lemma 3 (in appendix) that
\[
\log |KH^{-1}(1 + Bz^{-1})| = \log |KH^{-1}| + \log |1 + Bz^{-1}| \in \mathcal{L}_1
\]
Hence, it follows from Theorem 4 that there is an outer function $C \in \mathcal{H}_2$ such that (10) holds. Finally, since $K \in \mathcal{H}_1$ and $C \in \mathcal{H}_2$ is outer, it follows that $D = KC^{-1} \in N^+$.\]
which is equivalent to
\[
\sum_{n=1}^{N} \begin{bmatrix} \sigma^2 + 1 + 2 \Re(\hat{b}_n) & 0 \\ 0 & \gamma + (2 \Re(p_n^* k_n) - |p_n|^2)|f_n|^2 \end{bmatrix} - \begin{bmatrix} 1 & 1 & \ldots & 1 \\ 1 & 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} \hat{b}_n^T \nabla \gamma + \frac{1}{2} \Re(p_n^* k_n) - |p_n|^2 |f_n|^2 \end{bmatrix} \geq 0.
\]

Using Schur complement again, this is equivalent to
\[
L = \begin{bmatrix} 1 & \ldots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} \hat{b}_1 + 1 & \hat{b}_2 + 1 & \ldots & \hat{b}_N + 1 \\ k_1 h_1 & k_2 h_2 & \ldots & k_N h_N \end{bmatrix} \begin{bmatrix} L_{N+1} \\ 0 \\ \ldots \end{bmatrix} \geq 0,
\]
where
\[
L_{N+1} = N(\sigma^2 + 1) + 2 \sum_{n=1}^{N} \Re(\hat{b}_n),
\]
\[
L_{N+2} = N \gamma + \sum_{n=1}^{N} (2 \Re(p_n^* k_n) - |p_n|^2)|f_n|^2.
\]

The problem of minimizing \( \gamma \) subject to \( L \geq 0 \) and (14) can be cast as a semidefinite program with second-order cone constraints. The problem is quasiconvex for every \( N \).

By definition of the integral,
\[
\lim_{N \to \infty} \psi_N(B, K) = \psi(B, K)
\]
\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} |\hat{b}_n|^2 = \|B\|^2_2,
\]
so the value of the approximated problem tends arbitrarily close to (11) as \( N \) becomes sufficiently large. Obviously, when implementing the minimization program, the transfer functions \( K \) and \( B \) have to be parametrized using finite basis representations. The accuracy of the approximated problem clearly depends on this representation as well.

We now outline an algorithm for obtaining the solution to Problem 1:

1. Perform a spectral factorization in order to determine \( H \in \mathcal{H}_\infty \) with \( H^{-1} \in \mathcal{H}_\infty \) such that (6) holds.
2. Parameterize \( K \) and \( B \) by finite basis representations, for example as FIR filters.
3. Choose \( N \) large, determine the grid points and calculate \( p_n, f_n, h_n, k_n \) and \( \hat{b}_n \) for \( k = 1 \ldots N \).
4. Solve the quasiconvex optimization problem of minimizing \( \gamma \) subject to (14) and (15).
5. Determine \( C \) as a stable and outer spectral factor of \( |C|^2 \) according to (12). Finally, let \( D = KC^{-1} \).

4. NUMERICAL EXAMPLE

Suppose \( P = z^{-2} + 0.5z^{-7}, \ F = \frac{1}{z^{-0.5}}, \ G = 1 \) and \( \sigma = 1 \). \( K \) and \( B \) are parameterized as FIR filters with 20 and 19 coefficients, respectively. The minimization is implemented in Matlab using Yalmip, see Löfberg (2004), and ScDuMi, see Sturm (1999), with a grid distance of 0.0025. The impulse responses of the resulting transfer functions can be seen in Figures 3–6.

The minimum value for this problem is 0.90, to compare with the minimum value of 1.00 for the corresponding problem with no access to feedback from the channel, see Johannesson et al. (2010b)

![Fig. 3. Impulse response of K, the product of C and D, in the numerical example. Note that the timing of the peaks coincide with the exponents of P.](image1)

![Fig. 4. Impulse response of the encoder C in the numerical example.](image2)

![Fig. 5. Impulse response of the decoder D in the numerical example.](image3)

![Fig. 6. Impulse response of B, the feedback part of the encoder, in the numerical example. Note that the direct term is strongly negative, resulting in a transmission that is negatively correlated with the channel noise in the previous time step.](image4)
5. CONCLUSION

This paper treats the joint design of optimal linear encoders and decoders for filtering and transmission of a signal over an AWN channel with an SNR constraint and access to feedback of the channel output. The estimation is subject to a real-time constraint. The problem can be motivated as a distributed estimation problem, as a real-time communication problem or as a feed-forward compensator design problem in networked control.

The paper provides an extension of Johannesson et al. (2010b) to the feedback channel case and utilizes similar solution techniques, although the present problem is slightly more complicated.

It has been demonstrated that channel feedback can improve the performance. This is in some contrast to the classical information-theoretic result, see for example Massey (1990). Obviously, the difference is due to the real-time nature of the problem, illustrating that channel feedback can be of practical use when time-delays are important.

The fact that the optimal $D$ is of class $\mathcal{N}^+$ rather than $\mathcal{H}_2$, deserves some comment. Note first that all transfer functions from the input signals to the other signals in the system are in fact square-integrable. That $D$ necessarily is not, is unimportant as long as the model is correct. There is, however, no robustness guarantees towards disturbances at the input of $D$, or in the feedback to the encoder. To obtain such robustness properties, one would have to include additional noise signals, which would severely complicate the problem. However, this may not be a practical issue since the actual solution obtained from the optimization may very well turn out to have a $D \in \mathcal{H}_2$.

Possible future research questions include:

- If $P$, $F$ and $G$ are rational, will the optimal $C$ and $D$ also be rational (that is, implementable with finite memory)?
- Are linear solutions optimal? Under what conditions?
- Is the minimum in Theorem 2 attained?

APPENDIX

The following lemma consists of one of the results stated in Theorem 17.17 in Rudin (1986).

Lemma 3. Suppose $0 < p \leq \infty$, $X \in \mathcal{H}_p$, and $X$ is not identically zero. Define

$$\tilde{X}(e^{i \omega}) = \lim_{r \to 1^+} X(re^{i \omega}).$$

Then $\log |X| \in \mathbf{L}_1$.

The following theorem by Szegő (1975) is a generalization of the Fejér-Riesz Theorem.

Theorem 4. (Szegő). Suppose that $f(\omega)$ is a non-negative function on $\omega \in [-\pi, \pi]$, that is Lebesgue integrable and that

$$\int_{-\pi}^{\pi} \log f(\omega) \, d\omega > -\infty.$$ 

Then there exists $X \in \mathcal{H}_2$ such that $X(z) \neq 0$ for $|z| > 1$ and for almost all $\omega \in [-\pi, \pi]$ it holds that

- $X(e^{i \omega}) = \lim_{r \to 1^+} X(re^{i \omega})$ exists
- $f(\omega) = |X(e^{i \omega})|^2$.

REFERENCES


