MIMO Encoder and Decoder Design for Signal Estimation

Johannesson, Erik; Ghulchak, Andrey; Rantzer, Anders; Bernhardsson, Bo

2010

Citation for published version (APA):
MIMO Encoder and Decoder Design for Signal Estimation

Erik Johannesson, Andrey Ghulchak, Anders Rantzer and Bo Bernhardsson

Abstract—We study the joint design of optimal linear MIMO encoders and decoders for filtering and transmission of a vector-valued signal over parallel Gaussian channels subject to a real-time constraint. The objective is to minimize the sum of the estimation error variances at the receiving end. The design problem is nonconvex, but it is shown that a global optimum can be found by solving a related two-stage problem. The first stage consists of a mixed norm minimization problem, where the 2-norm corresponds to the error variance in a corresponding Wiener-Kolmogorov filtering problem and the 1-norm is induced by the channel noise. The second stage consists of a matrix spectral factorization.

I. INTRODUCTION

The problem studied in this paper lies in the intersection of estimation, communication and control. It is related to Wiener-Kolmogorov filtering, real-time coding and feed-forward compensator design. The problem may be motivated from each of these three perspectives, depending on which aspect one wishes to focus on.

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 [6]. The design of the optimal estimation filter can be formulated in the frequency domain as the minimization problem:

\[ \| (z^{-k} - B)F\|_2^2 + \| BG \|_2^2 \] (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 \( B \) is the design variable.

In this paper we generalize this problem to a setting where the measurement and the estimation are performed in two different locations. The additive white Gaussian noise (AWGN) channel is used to model the communication constraint between the two locations. It is shown that the inclusion of a channel between the two parts of the filter induces an additional term, a weighted 1-norm of \( B \), in the cost (1).

The problem under study may also be regarded as a communication problem since we wish to communicate a signal over a channel, with minimal distortion, subject to a real-time constraint. Classical communication theory does not worry about time delays [15] so information-theoretic tools seem to be of little use. However, real-time coding problems have lately been studied with increasing interest. See for example [8] for an overview.

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 AWGN channel with a mean square error criterion [3]. When the source is generated by a linear filter it may be enough to send scaled innovations over an AWGN channel [10].

These examples are somewhat counter-intuitive since a large allowed time delay usually makes the communication problem much easier in practice. Here, the introduction of additive measurement noise at the coder makes the real-time constraint important. The reason is that the noise gives an incentive to filter the signal at the same time that it is coded.

In a control perspective, the problem can be interpreted as that of designing a feed-forward compensator with access to remote and noisy measurements of the disturbance that is to be counteracted. In this context, an encoder filters the measurements and transmits information about the disturbance to the decoder/controller, which in turn can compensate.

A. Main Result

The main result of this paper is that the joint design of an optimal linear MIMO encoder-decoder pair for parallel Gaussian channels can be formulated as a convex optimization problem followed by a matrix spectral factorization. Specifically, it takes the form of a mixed norm minimization problem, where the relative weight of the two norms is determined by the maximum transmission power.

B. General Problem Description

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 a noisy communication channel to a another location, where a decoder then forms an estimate of the signal.

![Block diagram](image)

Fig. 1. Schematic illustration of the problem under consideration. The encoder and the decoder are designed to minimize the error. In the nominal case \( P \) represents a fixed time delay but more general dynamics are allowed.

The task is to design the encoder and the decoder such that the estimation error becomes as small as possible. The
estimation has to occur in real-time, as dictated by the transfer function $P$. Besides containing a fixed time delay, $P$ may include general dynamics that the signal passes through before it is to be estimated. In a feed-forward context, $P$ describes the propagation of the disturbance between the measurement and the compensation points.

The communication channel is modeled as a number of parallel Gaussian channels. That is, there is additive channel noise and the total power of the transmitted signal is limited. The relation between the noise variance and the power constraint determines the maximum amount of information that can be communicated.

C. Relations to Earlier Work

A lot of research efforts in the control community have 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 [4] and [12] respectively. Communication channel requirements for stability of feedback systems was given in [18], [14] and [2], among others. Fundamental limitations originating from channel communication and the compensation points.

Equalities and inequalities involving functions of $e^{i\omega}$ are to be interpreted as holding for almost all $\omega$. That is, the set of $\omega$ for which the (in)equality does not hold is of measure zero.

II. Notation

For $1 \leq p \leq \infty$, we define the Lebesgue spaces $L_p$ and the Hardy spaces $H_p$, over the unit circle, in the usual manner.

A transfer matrix $A \in \mathbb{C}^{m \times n}$ is given by $A = UV^*$, where $U \in \mathbb{C}^{m \times r}$, $\Sigma \in \mathbb{C}^{r \times r}$, $V \in \mathbb{C}^{n \times r}$ and $r = \min\{m, n\}$. Moreover, $U^*U = V^*V = I$ and $\Sigma$ is diagonal with diagonal elements $\sigma_k \geq 0$, $k = 1 \ldots r$.

The singular value decomposition of a transfer matrix $X \in L_p$ is defined pointwise and $L_p$ if the elements of $X$ are in $L_p$ and that $X \in H_p$ if the elements of $X$ are in $H_p$. For more details, consult a standard textbook such as [13].

A singular value decomposition of a matrix $A \in \mathbb{C}^{m \times n}$ is given by $A = U\Sigma V^*$, where $U \in \mathbb{C}^{m \times r}$, $\Sigma \in \mathbb{C}^{r \times r}$, $V \in \mathbb{C}^{n \times r}$ and $r = \min\{m, n\}$. Moreover, $U^*U = V^*V = I$ and $\Sigma$ is diagonal with diagonal elements $\sigma_k \geq 0$, $k = 1 \ldots r$.

The singular value decomposition of a transfer matrix $X \in L_p$ is defined pointwise and $L_p \times L_p$ if the elements of $X$ are in $L_p$. A transfer matrix $X \in H_p$ is said to be inner if $X \in H_p$ and $X^*X = I$. A transfer matrix $X \in H_p$ is defined to be outer if the set

$$\{Xq : q \text{ is a vector of polynomials in } z^{-1}\}$$

is dense in $H_p$. We say that $X$ is co-inner (co-outer) if $X^T$ is inner (outer). If $X \in H_p$ there exists an inner-outer factorization $X = X_0X_\omega$ where $X_0$ is inner and $X_\omega \in H_p$ is outer. Similarly there exists a co-inner-co-outer factorization $X = X_{\omega^*}X_i$ where $X_{\omega^*}$ is co-outer and $X_i$ is co-inner.

For any matrix $A \in \mathbb{C}^{m \times n}$ with $r = \min\{m, n\}$, define the Nuclear and the Frobenius norms as

$$\|A\|_e = \text{tr}\sqrt{A^*A} = \sum_{i=1}^{r} \sigma_i$$

$$\|A\|_F = \sqrt{\text{tr}(A^*A)} = \left( \sum_{i=1}^{r} \sigma_i^2 \right)^{1/2}$$

respectively. For transfer matrices, define the norms:

$$\|X\|_1 = \frac{1}{2\pi} \int_{0}^{2\pi} \|X(e^{i\omega})\|_e d\omega$$

$$\|X\|_2 = \left( \frac{1}{2\pi} \int_{0}^{2\pi} \|X(e^{i\omega})\|_F^2 d\omega \right)^{1/2}$$

and the relation

$$\langle X, Y \rangle = \frac{1}{2\pi} \int_{0}^{2\pi} \text{tr}(X^*(e^{i\omega})Y(e^{i\omega})) d\omega$$

To shorten notation, we will omit the argument $e^{i\omega}$ to transfer matrices when it is clear from the context.
variable, and instantaneous transmission power. That is, the elements in these transfer matrices do not need to be rational. Moreover, if $B$ is not identically zero, let $B = B_o$ be an inner-outer factorization and $B_oH = U_oΣV^*$ be a singular value decomposition. Then $C, D \in H_2$ are optimal if and only if $B = DC, \quad \|CH\|_2^2 \leq \alpha^2$

subject to the constraints

$$B = DC, \quad \|CH\|_2^2 \leq \alpha^2$$

is attained. The minimum value is $\frac{1}{\|F\|_2^2} \|BH\|_2^2$. Moreover, if $B = 0$ then the minimum is achieved by $D = 0$ and any function $C \in H_2$ that satisfies $\|CH\|_2^2 \leq \alpha^2$. If $B$ is not identically zero, let $B = B_o$ be an inner-outer factorization and $B_oH = U_oΣV^*$ be a singular value decomposition. Then $C, D \in H_2$ are optimal if and only if $B = DC, \quad \|CH\|_2^2 = \alpha^2, \quad DD^* = \frac{1}{\alpha^2} \|BH\|_2^2$$

Proof: If $B = 0$ the proof is trivial, so we assume from now that $B$ is not identically zero. It follows that neither $C$
nor $D$ are identically zero and that $\beta = \|CH\|_2 > 0$. Now, suppose that $C, D$ are feasible and that $\beta < \alpha$. Then
\[
\hat{C} = \frac{\alpha}{\beta} C, \quad \hat{D} = \frac{\beta}{\alpha} D
\]
are feasible and $\|\hat{D}\|_2 < \|D\|_2$. Hence, a necessary condition for optimality is that $\|CH\|_2 = \alpha$.

The remainder of this proof is divided into three parts. First, the dual problem is considered. Then, it is shown that there is a saddle point and the optimality criteria are derived. Finally, existence is proven by construction of a solution.

**Dual Problem:** In order to avoid dealing with analyticity constraints, we will now relax the search to $C, D \in L_2$ and show later that there are $C, D \in H_2$ that satisfy the optimality criteria. For $\lambda \geq 0$ and $\Phi \in L_\infty$, introduce the Lagrangian
\[
L(C, D, \lambda, \Phi) = \|D\|_2^2 + \lambda \left( \|CH\|_2^2 - \alpha^2 \right)
\]
\[
- \langle \text{Re } F, \text{Re } DC - B \rangle - \langle \text{Im } F, \text{Im } DC - B \rangle
\]
\[
= \|D\|_2^2 + \lambda \left( \|CH\|_2^2 - \alpha^2 \right) - \text{Re } \langle \Phi, DC - B \rangle
\]
\[
= \int_0^{2\pi} \left[ \|D\|_2^2 + \lambda \|CH\|_2^2 - \text{Re } \langle \Phi^* (DC - B) \rangle \right] \frac{d\omega}{2\pi} - \lambda \alpha^2
\]
The integrand in $L$ can be rewritten, by a completion of squares, as
\[
\|D\|_2^2 + \lambda \|CH\|_2^2 - \text{Re } \langle \Phi^* D - F \rangle
\]
\[
= \|D - \frac{1}{\lambda} \Phi C^*\|_2^2 + \lambda \|CH\|_2^2 - \frac{1}{\lambda} \|\Phi C\|_2^2 + \text{Re } \langle \Phi^* B \rangle
\]
\[
= \|D - \frac{1}{\lambda} \Phi C^*\|_2^2 + \text{tr} \left[ C (\lambda HH^* - \frac{1}{\lambda} \Phi^* \Phi) C^* \right] + \text{Re } \langle \Phi^* B \rangle
\]
only the first term in the integrand depends on $D$. The contribution of this term to the integral is minimized if and only if
\[
D = \frac{1}{\lambda} \Phi C^* \tag{7}
\]
If (7) holds, then the integral in $L$ only depends on $C$ through the second term, which has the pointwise infimum
\[
\inf_C \text{tr} \left[ C (\lambda HH^* - \frac{1}{\lambda} \Phi^* \Phi) C^* \right] = \begin{cases} 0 & \text{if } 4\lambda HH^* \geq \Phi^* \Phi \\ -\infty & \text{otherwise} \end{cases}
\]
Moreover,
\[
\text{tr} \langle \Phi^* B \rangle = \text{tr} \langle \Phi^* DC \rangle = \frac{1}{2} \text{tr} (C \Phi^* \Phi C^*) = \frac{1}{2} \|\Phi C^*\|_F^2 \tag{8}
\]
Thus, $\text{tr} \langle \Phi^* B \rangle$ is real and
\[
\inf_{C \in L_2} \int_{\mathbb{C}} \text{tr} \langle \Phi^* B \rangle \frac{d\omega}{2\pi} - \lambda \alpha^2 = \begin{cases} \frac{1}{2} \|\Phi C^*\|_F^2 - \lambda \alpha^2 & \text{if } 4\lambda HH^* \geq \Phi^* \Phi \\ -\infty & \text{otherwise} \end{cases}
\]
so the dual problem is:
\[
\begin{align*}
\text{maximize } & \quad \frac{1}{2\pi} \int_0^{2\pi} \text{tr} \langle \Phi^* B \rangle d\omega - \lambda \alpha^2 \\
\text{subject to } & \quad \Phi^* \Phi \leq 4\lambda HH^* \tag{9}
\end{align*}
\]
If $\lambda, \Phi$ are dual feasible then
\[
\text{tr} \left[ C (4\lambda HH^* - \Phi^* \Phi) C^* \right] = 0
\]
and integration gives
\[
\|\Phi C^*\|_F^2 = 4\lambda \|CH\|_2^2 = 4\lambda \alpha^2. \tag{10}
\]
It follows from (8) and (10) that
\[
\frac{1}{2\pi} \int_0^{2\pi} \text{tr} \langle \Phi^* B \rangle d\omega = \frac{1}{2} \|\Phi C^*\|_F^2 = 2\lambda \alpha^2. \tag{11}
\]
Introduce
\[
\Psi = \frac{1}{2\sqrt{\lambda}} \Phi H^{* -} - ^{\circ}
\]
The constraint (9) can then be written as $\Psi^* \Psi \leq I$ and
\[
\frac{1}{2\pi} \int_0^{2\pi} \text{tr} \langle \Psi^* B \rangle d\omega = \sqrt{\lambda} \frac{1}{2\pi} \int_0^{2\pi} \text{tr} \langle H \Psi^* B \rangle d\omega. \tag{12}
\]
From (11) and (12) we see that
\[
\frac{1}{2\pi} \int_0^{2\pi} \text{tr} \langle \Psi^* BH \rangle d\omega = \sqrt{\lambda} \alpha^2
\]
and thus the dual function can be written as
\[
\frac{1}{2\pi} \int_0^{2\pi} \text{tr} \langle \Psi^* BH \rangle d\omega = \sqrt{\lambda} \alpha^2 = \frac{1}{2\alpha^2} \left( \frac{\sqrt{\lambda}}{\alpha^2} \right)^2
\]
\[
\leq \frac{1}{2\alpha^2} \left( \frac{\sqrt{\lambda}}{\alpha^2} \right)^2 \tag{13}
\]
We will now perform pointwise maximization of the integrand in (13). Recall that $BH = B B_0 V = B U_0 \Sigma V^*$. Assume that $B_0$ has $n$ rows. Then $\Sigma$ is diagonal with diagonal elements $\sigma_k \geq 0$, $k = 1 \ldots n$. Since $B_0 V$ is wide (it has $n_f \geq n$ columns) it follows that $U_0$ is square and thus it is unitary.

Let $U = B B_0 U_0$ and introduce $\bar{\Psi} = \Psi^* \Psi$. Then it follows from $\Psi^* \Psi \leq I$ and $UU^* \leq I$ that
\[
\Psi^* \Psi = V^* \Psi^* U U^* \Psi \Psi V \leq V^* \Psi^* \Psi \leq V^* V = I
\]
Using $\bar{\Psi}$, we can obtain an upper bound for the maximum:
\[
\max_{\Psi^* \Psi \leq I} \text{tr} \langle \Psi^* BH \rangle = \max_{\Psi^* \Psi \leq I} \text{tr} \langle \Psi^* U \Sigma V^* \rangle = \max_{\Psi^* \Psi \leq I} \text{tr} \langle V^* \Psi^* U \Sigma \rangle
\]
\[
\leq \max_{\Psi^* \Psi \leq I} \text{tr} \langle \Psi^* \Psi \rangle = \sum_{k=1}^n \max_{\|\Psi\|_0 = 1} \sigma_k \bar{\Psi}_{kk} = \sum_{k=1}^n \sigma_k
\]
which is achieved if and only if $\bar{\Psi} = I$. Therefore, $\Psi$ is a maximizer if and only if $U^* \Psi V = I$ and $\Psi^* \Psi \leq I$. The solutions can be parametrized as:
\[
\Psi = U V^* + \Psi_0 = B B_0 V^* + \Psi_0
\]
\[
0 = U^* \Psi_0 V = U_0^* B_0^* \Psi_0 V
\]
\[
I \geq \Psi^* \Psi \tag{14}
\]
Pre-multiplying (14) with $U_0$ gives
\[
B_0^* \Psi_0 V = 0. \tag{16}
\]
The upper bound of the maximum is achieved (for example with $\Psi_0 = 0$), so the value of the dual problem is
\[
\max_{\Psi^* \Psi \leq I} \frac{1}{\alpha^2} \left( \frac{1}{2\pi} \int_0^{2\pi} \text{tr} \langle \Psi^* BH \rangle d\omega \right)^2 = \frac{1}{\alpha^2} \|BH\|^2_2
\]
The maximizing dual variables are given by
\[
\Phi = 2\sqrt{\lambda} \Psi H^{* -} = 2\sqrt{\lambda} (B_0 U_0 V^* + \Psi_0) H^{* -} \tag{17}
\]
where \( \Psi_0 \) is such that (15) and (16) hold, and
\[
\lambda = \left( \frac{1}{\alpha^2} \| BH \|_1 \right)^2. \tag{18}
\]

**Saddle Point:** We will now show that there is a saddle point, which implies that the duality gap is zero.

In the following, assume that (15), (16), (17) and (18) hold. Then \( \lambda \) and \( \Phi \) are dual feasible and \( (C, D, \lambda, \Phi) \) is a saddle point if and only if \( C, D \in \mathbf{H}_2 \) are primal feasible,
\[
\lambda \left( \| CH \|_2^2 - \alpha^2 \right) = 0 \tag{19}
\]
and
\[
L(C, D, \lambda, \Phi) = \inf_{C \in \mathbf{H}_2} L(\tilde{C}, D, \lambda, \Phi). \tag{20}
\]

The saddle point conditions imply that \( \| CH \|_2 = \alpha \) since \( \lambda > 0 \) and that \( D = \frac{1}{\lambda} \Phi C^* \) as we have seen earlier that this follows from minimization of the Lagrangian.

Suppose that \( C, D \) satisfy \( B = DC \) and \( D = \Phi C^* \). Then
\[
DD^* = \Phi DC\Phi^* = \frac{1}{\lambda} \Phi B_0 H (V U_o^* B_i^* + \Psi_0^*)
\]
\[
= \sqrt{\lambda} (B_0 U_o^* \Psi_0^* B_i^* + B_0 U_o^* \Psi_0^* B_i^*)
\]
Clearly, \( DD^* \) and \( B_0 U_o^* \Psi_0^* B_i^* \) are Hermitian. Accordingly, \( A = B_0 U_o^* \Psi_0^* \Psi_0^* \) must be Hermitian. Now, by (16),
\[
AB_i = B_0 U_o^* \Psi_0^* B_i^* \Rightarrow 0 = AB = A^* B_i = \Psi_0^* V U_o^* B_i^* = \Psi_0^* V U_o^*
\]
Hence,
\[
DD^* = \sqrt{\lambda} B_0 U_o^* \Psi_0^* B_i^* \tag{21}
\]
Now, suppose instead that \( C, D \in \mathbf{H}_2 \) satisfy \( B = DC \), \( \| CH \|_2 = \alpha \) and (21). Then \( C, D \) are primal feasible and (19) is satisfied. Moreover,
\[
L(C, D, \lambda, \Phi) = \| D \|_2^2 = \frac{\sqrt{\lambda}}{\alpha} \int_0^{2\pi} \text{tr}(B_0 U_o^* \Psi_0^* B_i^*) d\omega
\]
\[
= \sqrt{\lambda} \int_0^{2\pi} \text{tr}(\Sigma) d\omega = \frac{1}{\alpha^2} \| BH \|_1^2,
\]
so (20) holds and thus the saddle point conditions are satisfied. Since these assumptions and the saddle point conditions imply each other, they are equivalent.

To conclude, we have shown that \( (C, D, \lambda, \Phi) \) is a saddle point (and \( C, D \) are thus optimal) if and only if \( C, D \in \mathbf{H}_2 \) satisfy \( B = DC \), \( \| CH \|_2 = \alpha \) and (21).

**Existence of Solution:** We will now construct a solution that satisfies the optimality conditions.

Define \( M = \sqrt{\lambda} U_o^* \Psi_0^* \) which is Hermitian with real diagonal. \( B_o H \) is outer because both \( B_o \) and \( H \) are outer. Since \( U_o \) is unitary it follows that \( \sigma_k > 0 \), \( k = 1 \ldots n \) and that \( M \) is positive definite. From Lemma 3 (in appendix) it follows that \( \log \sigma_k \in \mathbf{L}_1 \) and therefore
\[
\log \det M = \frac{n}{2} \log \lambda + \sum_{k=1}^{n} \log \sigma_k \in \mathbf{L}_1
\]
According to Theorem 2 there is an outer transfer matrix \( D_o \in \mathbf{H}_2 \) such that \( M = D_o D_o^* \). Let \( \tilde{D} = B_o D_o \in \mathbf{H}_2 \) and \( \tilde{C} = D_o^{-1} B_o \). Then
\[
\tilde{C} = D_o^{-1} B_o H H^{-1} = D_o^{-1} U_o \Sigma V^* H^{-1}
\]
\[
= D_o^{-1} U_o \Sigma U_o^* U_o V^* H^{-1} = \frac{1}{\sqrt{\lambda}} D_o^{-1} U_o V^* H^{-1} \in \mathbf{L}_2
\]
Since \( D_o \) is outer it follows that \( \tilde{C} \in \mathbf{H}_2 \).

We can now verify that \( \tilde{C} \) and \( \tilde{D} \) satisfy the optimality conditions:
\[
\tilde{D} \tilde{C} = B_o D_o \tilde{D}_o^{-1} B_o = B_o B_o = B,
\]
\[
\| \tilde{C} \|_2^2 = \| D_o^{-1} B_o H \|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} \text{tr}(H^* B_o D_o^{-1} B_o H) d\omega
\]
\[
= \frac{1}{2\pi} \int_0^{2\pi} \text{tr}(V U_o^* M U_o^* V^*) d\omega
\]
\[
= \frac{1}{2\pi} \int_0^{2\pi} \text{tr}(\Sigma) d\omega = \alpha^2
\]
and
\[
\tilde{D} \tilde{D}^* = B_o D_o \tilde{D}_o^* B_o = \sqrt{\lambda} B_o U_o^* \Psi_0^* B_i^*.
\]
If the rank of \( B \) is smaller than \( n \), then \( \tilde{C} \) and \( \tilde{D} \) are not of the required size. We have that \( \tilde{C} \) is \( n \times n \) and that \( \tilde{D} \) is \( n_o \times n_f \), where \( n \leq \min\{n_o, n_f\} \leq n \). Meanwhile, \( C \) to be \( n_i \times n_f \) and \( D \) to be \( n_e \times n_i \). Therefore, let
\[
D = \begin{bmatrix} \tilde{D} & 0_{n_o \times n_x \times n_f} \\ 0_{n_i \times n_r \times n_f} \end{bmatrix}, \quad C = \begin{bmatrix} \tilde{C} \\ 0_{n_o \times n_x \times n_f} \end{bmatrix}
\]
Noting that \( DC = \tilde{D} \tilde{C} = B \), \( \| CH \|_2 = \| \tilde{C} \|_2 \) and that \( DD^* = \tilde{D} \tilde{D}^* \) we conclude that \( C, D \) are optimal.

Using Lemma 2, we can state the main result of this paper, which shows that the design problem can be solved using convex optimization techniques. The theorem is more or less the same as the corresponding theorem in [7], although the optimality conditions have changed to reflect that the systems are MIMO.

**Theorem 1:** Suppose that \( \alpha > 0 \), that \( F, G, P \in \mathbf{H}_\infty \) and that
\[
\exists \epsilon > 0 \text{ such that } FF^* + GG^* \geq \epsilon I.
\]
Then the minimum
\[
\min_{C,D \in \mathbf{H}_2} ||(P - DC)F||_2^2 + ||DCG||_2^2 + ||D||_2^2 \tag{22}
\]
subject to
\[
\| CF \|_2^2 + \| CG \|_2^2 \leq \alpha^2 \tag{23}
\]
is attained and is equal to the minimum of the convex optimization problem
\[
\min_{B \in \mathbf{H}_2} ||(P - B)F||_2^2 + ||BG||_2^2 + \frac{1}{\alpha^2} \| B \left[ F \right. \left. G \right] \|_1^2 \tag{24}
\]
which is attained by a unique minimizer.

Further, suppose \( B \in \mathbf{H}_2 \) minimizes (24). If \( B = 0 \), then (22) subject to (23) is minimized by \( D = 0 \) and any function
$C \in \mathbf{H}_2$ that satisfies (23). If $B$ is not identically zero, then $C,D \in \mathbf{H}_2$ minimize (22) subject to (23) if and only if

$$B = DC, \quad \|C [F \ G] \|_2^2 = \alpha^2,$$

$$DD^* = \frac{1}{\alpha^2} \|B [F \ G] \|_1^2 B^*_o \Sigma U^*_o B^*_1,$$

where $B_1$ is defined by an inner-outer factorization $B = B_1B_o$ and $U_o$ is given by a singular value decomposition $B_oH = U_o \Sigma V^*$, with $H \in \mathbf{H}_\infty$ satisfying $H^{-1} \in \mathbf{H}_\infty$ and $HH^* = FFF^* + GG^*$.

**Proof:** By Lemma 1 there exists $H \in \mathbf{H}_\infty$ with the specified properties and the channel input constraint (23) will henceforth be written as $\|CH\|_2 \leq \alpha$. Define the sets:

$$\Theta = \{(C,D) : C,D \in \mathbf{H}_2, \|CH\|_2 \leq \alpha\}$$

$$\Theta_B = \{(C,D) : (C,D) \in \Theta, B = DC\}$$

and the functional

$$\psi(C,D) = \|(P - DC)F\|_2^2 + \|D CG\|_2^2 + \|D\|_2^2.$$

The infimum of (22) subject to (23) can be written

$$\inf_{C,D \in \Theta} \psi(C,D) = \inf_{B \in \mathbf{H}_2} \inf_{C,D \in \Theta_B} \psi(C,D)$$

$$= \inf_{B \in \mathbf{H}_2} \|(P - B)F\|_2^2 + \|B G\|_2^2 + \inf_{C,D \in \Theta_B} \|D\|_2^2$$

$$= \inf_{B \in \mathbf{H}_2} \|(P - B)F\|_2^2 + \|B G\|_2^2 + \frac{1}{\alpha^2} \|BH\|_2^2. \quad (25)$$

The first equality comes from the fact that a product of two functions in $\mathbf{H}_2$ is in $\mathbf{H}_1$, and that any function in $\mathbf{H}_1$ can be written as a product of two functions in $\mathbf{H}_2$. In the third equality, Lemma 2 was applied to perform the inner minimization.

We will now show that the minimum in (25) is attained by a unique $B \in \mathbf{H}_2$. To this end, perform a completion of squares:

$$\psi(B) = \|(P - B)F\|_2^2 + \|B G\|_2^2 + \frac{1}{\alpha^2} \|BH\|_2^2$$

$$= \|BH - PFF^*H^{-1}\|_2^2 + \frac{1}{\alpha^2} \|BH\|_2^2 + \text{const.}$$

Let $X = BH \in \mathbf{H}_1$ and $R = PFF^*H^{-1} \in \mathbf{L}_\infty$. Minimizing $\psi(B)$ is then equivalent to minimizing

$$\rho(X) = \|R - X\|_2^2 + \frac{1}{\alpha^2} \|X\|_1^2 \quad (26)$$

over $X \in \mathbf{H}_1$. However, since we want to minimize $\rho(X)$ it is enough to consider $X$ with $\rho(X) \leq \rho(0) = \|R\|_2^2$. Hence,

$$\|X\|_2 \leq \|R - X\|_2 + \|R\|_2 \leq \sqrt{\rho(X)} + \|R\|_2 \leq 2\|R\|_2 \overset{df}{=} r.$$

Now, in the weak topology, $\rho(X)$ is lower semicontinuous on $\mathbf{H}_2$ and the set $\{X : \|X\|_2 \leq r\}$ is compact. This proves the existence of a minimum. Moreover, $\rho(X)$ is strictly convex, and thus the minimum is unique.

Since $\|X\|_2 \leq r$, we can restrict the search to $X \in \mathbf{H}_2$ without loss of generality. Because $H^{-1} \in \mathbf{H}_\infty$ it follows that $B = H^{-1}X \in \mathbf{H}_2$ and that (25) is equal to (24).

Since $\rho(X)$ attains a unique minimum in $\mathbf{H}_2$, so does $\psi(B)$ and hence the minimum (22) subject to (23) is attained, since it is equal to the minimum of $\psi(B)$. The optimality conditions follow from the application of Lemma 2.

The cost function (24) consists of three terms, which can be given the following interpretations: The sum of the first two are equal to the cost in the situation where the channel is noise-free and has unlimited capacity, which is the error variance in the Wiener-Kolmogorov problem (1), if the time delay is replaced by the more general filter $P$. The third term is the error induced by the channel noise. It is interesting to note that the first two terms are 2-norm functions of the decision variable $B$, while the third term is a weighted 1-norm of $B$. Thus, the problem is equivalent to a mixed norm minimization problem with the parameter $\alpha$ determining the relative importance of the two norms.

It was noted earlier that the solution is not unique. To clarify, the optimal $B$ is unique but there are multiple factorizations of $B$ into $C$ and $D$ that achieve the optimal value. For example, a second solution is trivially found by changing the sign of both $C$ and $D$.

**B. Procedure for Numerical Solution**

Working along the lines of Theorem 1, it is possible to numerically solve the design problem, described in Section III, by the following procedure: First, minimize (26). In practice, this is done approximately using a finite basis representation of $X$ and sum approximations of the integrals. This minimization can then be cast a quadratic program with second-order cone constraints.

Then perform a matrix spectral factorization to find $H \in \mathbf{H}_\infty$ with $H^{-1} \in \mathbf{H}_\infty$ that satisfies (6). Perform an inner-outer factorization of $B = XH^{-1}$ to obtain $B = B_1B_o$ and a singular value decomposition to obtain $B_oH = U_o \Sigma V^*$. Let

$$M = \frac{1}{\alpha^2} \|BH\|_1 U_o \Sigma U^*_o$$

and perform a matrix spectral factorization to obtain $D_o \in \mathbf{H}_2$ which is outer and satisfies $D_oD^*_o = M$. Finally let $D = B_1D_o$ and $C = D^*_oB_o$.

It is possible that the obtained $D$ and $C$ are of incorrect size (if the rank of $B$ is lower than the number of channels $n_i$). In this case just add columns of zeros to $D$ and rows of zeros to $C$ until they are of correct size.

**V. CONCLUSIONS**

This paper treats the joint design of optimal linear MIMO encoders and decoders for filtering and transmission of a signal over parallel Gaussian channels 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 [7], we studied the SISO version of this problem and found that it can be formulated as a mixed $\mathbf{H}_1$ and $\mathbf{H}_2$ problem. In this paper, these results have been extended to the MIMO case. Perhaps as expected, the results are very similar, especially the main theorem. The factorization
problem seems however to be more difficult for multi-variable systems, both in theory and in practice.

The assumption (4) may deserve some explanation: If there are too few channels, the maximum rank of the product of $C$ and $D$ may become smaller than the smallest dimension of $B$. Then not all $B$ would be realizable as a product of $C$ and $D$. Thus some rank condition would have to be imposed on $B$ in Theorem 1, which is very difficult to handle.

This work provides several topics for further research, that we plan to investigate in the future:

- If $P$, $F$ and $G$ are rational, will the optimal $C$ and $D$ also be rational (that is, implementable with finite memory)?
- Preliminary results suggest that the answer is negative.
- Are linear solutions optimal? Under what conditions?
- Is the method used in this paper applicable to other structures, such as feedback loops?

**APPENDIX**

**Lemma 3:** Suppose that $m < n$ and that the $m \times n$ transfer matrix $X \in \mathbb{H}_p$ is outer. Then the singular values of $X$ satisfy

$$\log \sigma_k \in L_1, \quad k = 1 \ldots m.$$ 

**Proof:** A co-inner-outer factorization gives $X = X_{co}X_{ci}$, where $X_{co}$ is $m \times m$. It is well-known that if a matrix function $Y \in \mathbb{H}_p$ is square then it is outer if and only if $\det Y$ is outer. Also, if a scalar function is outer, then the logarithm of the absolute value is $L_1$. Thus if $Y$ is square and outer then $\log |\det Y| \in L_1$.

For the singular values of $X$, it holds that

$$\sum_{k=1}^m \log \sigma_k = \frac{1}{2} \log \prod_{k=1}^m \sigma_k^2 = \frac{1}{2} \log \det XX^*$$

$$= \frac{1}{2} \log \det X_{co}X_{ci}^*X_{ci}X_{co}^* = \frac{1}{2} \log \det X_{co}^*X_{co}$$

$$= \log |\det X_{co}| \in L_1.$$ 

Now, $\log \sigma_k < \sigma_k \in L_1$ and so

$$\int_0^{2\pi} \log \sigma_k \, dw < \int_0^{2\pi} \sigma_k \, dw < \infty$$

Since the sum of the logarithms is $L_1$ and every term has an integral bounded from above, it follows that

$$\int_0^{2\pi} \log \sigma_k \, dw > -\infty$$

and hence $\log \sigma_k \in L_1$.

The following theorem, given in [11], is the matrix generalization of a spectral factorization theorem by Szegö [17].

**Theorem 2 (Matrix Spectral Factorization):** Suppose that $Y \in L_1$ is $m \times m$ and positive definite on the unit circle. If $\log \det Y \in L_1$ then there exists an outer $m \times m$ transfer matrix $X \in \mathbb{H}_2$ such that

$$Y = XX^*$$

**ACKNOWLEDGEMENTS**

The authors gratefully acknowledge funding received for this research from the Swedish Research Council through the Linnaeus Center LCCC and the European Union’s Seventh Framework Programme under grant agreement number 224428, project acronym CHAT.

**REFERENCES**


