Proceedings of the 32nd Conference
on Decision and Control
San Antonio, Texas • December 1993

A PROBABILISTIC APPROACH TO MULTIVARIABLE ROBUST FILTERING, PREDICTION AND SMOOTHING*

Kenth Öhrn, Anders Ahlén and Mikael Sternad
Systems and Control Group, Uppsala University
P.O. Box 27, S–751 03 Uppsala, Sweden

Abstract

A new approach to robust filtering, prediction and smoothing of discrete–time signal vectors is presented. Linear time–invariant filters are designed to be insensitive to spectral uncertainty in signal models. The goal is to obtain a simple design method, leading to filters which are not overly conservative. Modelling errors are described by sets of models, parametrized by random variables with known covariances. A robust design is obtained by minimizing the $\mathcal{H}_2$–norm of the estimation error, averaged with respect to the assumed model errors. A polynomial solution, based on an averaged spectral factorization and a unilateral Diophantine equation, is presented. The robust estimator is referred to as a cautious Wiener filter. It turns out to be only slightly more complicated to design than an ordinary Wiener filter.

1. Introduction

For any model–based filter, modelling errors are a potential source of performance degradation. A *cautious Wiener filter* for predicion, filtering or smoothing of discrete–time signal vectors will be proposed here. It can be designed by using a generalization of the polynomial equations methodology pioneered by Kučera, [9]. It is based on a stochastic description of model errors, related to the stochastic embedding concept of Goodwin and co–workers [4], [5]. Our approach is based on the following choices of model and criterion:

- A set of (true) dynamic systems is assumed to be well described by a set of stable, discrete–time linear and time–invariant transfer function matrices
$$\mathcal{F} = \mathcal{F}_o + \Delta\mathcal{F} \quad . \tag{1}$$
We call such a set an *extended design model.* Here, $\mathcal{F}_o$ represents a stable nominal model, while an *error model* $\Delta\mathcal{F}$ describes a set of stable transfer functions, parametrized by stochastic variables. The random variables enter linearly into $\Delta\mathcal{F}$.
- A single robust linear filter is to be designed for the whole class of possible systems. Robust performance is obtained by minimizing the averaged mean square estimation error
$$J = \text{trace}\bar{E}E(\varepsilon(k)\varepsilon(k)^*) \quad . \tag{2}$$
Here, $\varepsilon(k)$ is the estimation error vector, E denotes expectation over noise and $\bar{E}$ is an expectation over the stochastic variables parametrizing the error model $\Delta\mathcal{F}$.

The averaged mean square error has been used previously, see [2] and [6]. We suggest the use of (2), together with a particular description of the set (1): transfer function elements in $\Delta\mathcal{F}$ have stochastic numerators and fixed denominators. Such models can describe nonparametric uncertainty and undermodelling as well as parametric uncertainty.

Most previous suggestions for obtaining robust filters have been based on some type of minimax approach [3], [8], [10]. Minimax design becomes very complex, unless there exists either a saddle point or a boundary point solution. The computational effort involved in minimax design is considerable. Furthermore, in many problems, closed–form solutions do not exist.

Apart from leading to a radical reduction of computational complexity, the approach proposed here avoids two drawbacks of robust minimax design. First, the stochastic variables in $\Delta\mathcal{F}$ need not have compact support. Thus, the descriptions of model uncertainties may have *"soft" bounds.* Secondly, not only the range of the uncertainties, but also their likelihood is taken into account by the criterion (2). Highly probable model errors will affect the estimator design more than do very rare "worst cases". Therefore, the performance loss in the nominal case, the price paid for robustness, becomes smaller than for a minimax design. The conservativeness is thus reduced.

The present paper generalizes the scalar robust design of [12] to multisignal estimation. One of our goals will be to hold the number of design equations to a minimum, without sacrificing numerical accuracy. We use matrix fraction descriptions with diagonal denominators and common denominator forms. This leads to a solution which is, in fact, significantly simpler and more numerically well–behaved,

*The work has been supported by the Swedish Research Council for Engineering Sciences (TFR), under grant 92-775.

0191-2216/93/$3.00 © 1993 IEEE

than the corresponding nominal $\mathcal{H}_2$–design (without uncertainty) presented in [1]. We end up with just two equations for robust estimator design: a polynomial matrix spectral factorization and a unilateral Diophantine equation. This solution provides structural insight; important properties of a robust estimator are evident by direct inspection of the filter expression.

Remarks on the notation: Signals and polynomial coefficients may, in the following, be complex-valued. (This is required in e.g. communications applications.) For any polynomial, $P(q^{-1}) = p_o + p_1q^{-1} + \ldots + p_{np}q^{-np}$ in the backward shift operator q^{-1}, define $P_*(q) \triangleq p_o^* + p_1^*q + \ldots + p_{np}^*q^{np}$, where q is the forward shift operator and p_j^* denotes the complex conjugate of p_j. Rational matrices, or transfer functions, are denoted by boldface script symbols, e.g. $\mathcal{R}(q^{-1})$. Polynomial matrices are denoted by boldface symbols, such as $\boldsymbol{P}(q^{-1})$, while constant matrices are denoted as $\mathbf{P}$. For polynomial or rational matrices, $\boldsymbol{P}_*(q)$ and $\mathcal{R}_*(q)$ means complex conjugate and transpose. Arguments of polynomials and matrices are often omitted, when there is no risk of misunderstanding. The *degree* of a polynomial matrix is the highest degree of any of its polynomial elements. $\boldsymbol{P}(q^{-1})$ is called *stable* if all zeros of $\det \boldsymbol{P}(z^{-1})$ are located in $|z| < 1$. We denote the trace of $\boldsymbol{P}$ by $\mathrm{tr}\boldsymbol{P}$. A rational matrix may be represented by polynomial matrices as a left matrix fraction description (MFD), $\mathcal{G} = \boldsymbol{A}^{-1}\boldsymbol{B}$. It may also be represented in a common denominator form $\mathcal{G} = \boldsymbol{B}/A$. The monic polynomial A is then the least common denominator of all elements in $\mathcal{G}$.

2. Problem formulation

Consider the following extended design model

$$\begin{aligned} y(k) &= \mathcal{G}(q^{-1})u(k) + \mathcal{H}(q^{-1})v(k) \\ u(k) &= \mathcal{F}(q^{-1})e(k) \\ f(k) &= \mathcal{D}(q^{-1})u(k) \end{aligned} \qquad (3)$$

where $\mathcal{G}$, $\mathcal{H}$, $\mathcal{F}$ and $\mathcal{D}$ are stable and causal, but possibly uncertain, transfer functions of dimension $p|s$, $p|r$, $s|n$ and $\ell|s$, respectively. The noise sequences $\{e(k)\}$ and $\{v(k)\}$ are mutually uncorrelated and zero mean stochastic sequences. They have unit covariance matrices. Thus scaling is included in $\mathcal{F}$ and $\mathcal{H}$.

Multisignal estimation

From data $y(k)$ up to time $k+m$, an estimator

$$\hat{f}(k|k+m) = \mathcal{R}(q^{-1})y(k+m) \qquad (4)$$

of $f(k)$ is sought. The estimator may be a predictor ($m < 0$), a filter ($m = 0$), or a fixed lag smoother ($m > 0$). The transfer function $\mathcal{R}$, of dimension $\ell|p$, is required to be stable and causal. It is designed to minimize (2), where

$$\varepsilon(k) = (\varepsilon_1(k)\ldots\varepsilon_\ell(k))^T \triangleq \mathcal{W}(q^{-1})(f(k) - \hat{f}(k|k+m)).$$

Here, $\mathcal{W}(q^{-1})$ is a known stable and causal $\ell|\ell$ rational weighting matrix, with a stable and causal inverse. It may be used to emphasize filtering performance in particular frequency bands. See Figure 1.

The model (3) offers considerable flexibility. For example, when estimating a signal $u(k)$ in coloured noise, we set $\mathcal{G} = \mathcal{D} = \mathbf{I}_s$, giving $f(k) = u(k)$. In deconvolution, or input estimation problems, $\mathcal{G}$ is a dynamic system and $\mathcal{D} = \mathbf{I}_s$. In a state estimation problem, $u(k)$ is the state vector, $\mathcal{G}$ and $\mathcal{D}$ are constant matrices while $\mathcal{H}\,v(k)$ represents measurement noise.

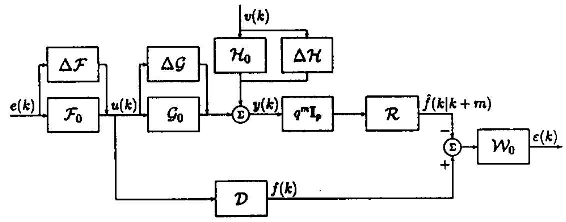

Figure 1: A general linear filtering setup.

Parametrization of the model

We choose to represent $\mathcal{G}$ and $\mathcal{H}$ with left MFD's having *diagonal* denominators[1], while $\mathcal{F}$, $\mathcal{D}$ and $\mathcal{W}$ are represented in common denominator form;[2]

$$\mathcal{G} = \boldsymbol{A}^{-1}\boldsymbol{B} \;\; ; \;\; \mathcal{H} = \boldsymbol{N}^{-1}\boldsymbol{M} \qquad (5)$$

$$\mathcal{F} = \frac{1}{D}\boldsymbol{C} \;\; ; \;\; \mathcal{D} = \frac{1}{T}\boldsymbol{S} \;\; ; \;\; \mathcal{W} = \frac{1}{U}\boldsymbol{V} \; .$$

In (5), $\mathcal{G}$, $\mathcal{H}$ and $\mathcal{F}$ may be uncertain, while $\mathcal{W}$ is assumed to be known. It can be shown that uncertainty in $\mathcal{D}$ does not affect the optimal filter design. Therefore, uncertainty in $\mathcal{D}$ is not introduced. The matrix $\boldsymbol{V}$ is assumed stable, with $\boldsymbol{V}(0)$ nonsingular.

The extended design models, cf. (1) and (3),

$$\mathcal{G} = \mathcal{G}_o + \Delta\mathcal{G} \;\; , \;\; \mathcal{H} = \mathcal{H}_o + \Delta\mathcal{H} \;\; , \;\; \mathcal{F} = \mathcal{F}_o + \Delta\mathcal{F}$$

are now represented in polynomial matrix form. Us-

[1]Note that this is a natural choice, if transfer functions are obtained by means of identification.

[2]We have made these choices to obtain tidy and transparent design equations. Coprime factorizations are avoided , which are numerically sensitive, and also, we obtain an *unilateral* Diophantine equation.

ing $\hat{B}_o = A_1 B_o$, $\hat{B}_1 = A_o B_1$ etc., we define

$$\begin{aligned} \mathcal{G} &= A_o^{-1}B_o + A_1^{-1}B_1\Delta B \\ &= A_o^{-1}A_1^{-1}(\hat{B}_o + \hat{B}_1\Delta B) \quad \triangleq \quad A^{-1}B \\ \mathcal{H} &= N_o^{-1}M_o + N_1^{-1}M_1\Delta M \\ &= N_o^{-1}N_1^{-1}(\hat{M}_o + \hat{M}_1\Delta M) \quad \triangleq \quad N^{-1}M \\ \mathcal{F} &= \frac{1}{D_o}C_o + \frac{1}{D_1}C_1\Delta C \\ &= \frac{1}{D_oD_1}(\hat{C}_o + \hat{C}_1\Delta C) \quad \triangleq \quad \frac{1}{D}C \; . \end{aligned} \tag{6}$$

Above, $\mathcal{G}_o = A_o^{-1}B_o$ represents the nominal model and $\Delta\mathcal{G} = A_1^{-1}B_1\Delta B$ is the error model. The same holds for $\mathcal{H}$ and $\mathcal{F}$. The diagonal polynomial matrices $A = A_oA_1$, $N = N_oN_1$ and the polynomials $D = D_oD_1$, T and U are all assumed to be stable, with causal inverses. In the error models, the polynomial D_1, the diagonal matrices A_1 and N_1 and the matrices C_1, B_1 and M_1 are fixed. They can be used to tailor the error models for specific needs.

The matrices ΔB, ΔC and ΔM contain polynomials, with jointly distributed random variables as coefficients. These coefficients parametrize the class of assumed true systems. One particular modelling error is represented by one particular realization of the random coefficients. Element ij of ΔP is denoted

$$\Delta P^{ij} = \Delta p_o^{ij} + \Delta p_1^{ij}q^{-1} + \ldots + \Delta p_{\delta p}^{ij}q^{-\delta p} \tag{7}$$

where δp is the degree of ΔP. All coefficients have zero means, so the nominal model is the average model in the set. Only the first and second order moments need to be known, since the type of distribution, and higher order moments, will not affect the filter design. The parameter covariances $\bar{E}(\Delta p_r^{ij})(\Delta p_s^{\ell k})^*$ are collected in covariance matrices $\mathbf{P}_{\Delta\mathbf{P}}^{(ij,\ell k)}$, see (10) below.

Error models can be obtained from ordinary identification experiments, provided the model structures match. For SISO systems, error models can be estimated in presence of undermodelling, using a maximum likelihood approach, [4]. They could also be obtained from frequency domain data on system variability, [5]. Even if the statistics is hard to obtain, one could still use the elements of covariance matrices pragmatically, as robustness "tuning knobs". They are then tuned to obtain reasonable performance for the uncertainty set, without degrading the performance in the nominal case too much. See [11] or [12] for a further discussion about the error models.

Covariance matrices

The covariance matrices will be organized as follows. First, (7) is written as

$$\Delta P^{ij}(q^{-1}) = \varphi^T(q^{-1})\bar{p}_{ij} \tag{8}$$

$$\begin{aligned} \varphi^T(q^{-1}) &= (1 \;\; q^{-1} \ldots q^{-\delta p}) \\ \bar{p}_{ij} &= (\Delta p_o^{ij} \;\; \Delta p_1^{ij} \ldots \Delta p_{\delta p}^{ij})^T \; . \end{aligned} \tag{9}$$

Secondly, the cross covariance matrix $\mathbf{P}_{\Delta\mathbf{P}}^{(ij,\ell k)} = \bar{E}\bar{p}_{ij}\bar{p}_{\ell k}^*$ between coefficients of $\Delta P^{ij}(q^{-1})$ and $\Delta P^{\ell k}(q^{-1})$, is defined by

$$\mathbf{P}_{\Delta\mathbf{P}}^{(ij,\ell k)} = \begin{bmatrix} \bar{E}(\Delta p_o^{ij})(\Delta p_o^{\ell k})^* & \ldots & \bar{E}(\Delta p_o^{ij})(\Delta p_{\delta p}^{\ell k})^* \\ \vdots & \ddots & \vdots \\ \bar{E}(\Delta p_{\delta p}^{ij})(\Delta p_o^{\ell k})^* & \ldots & \bar{E}(\Delta p_{\delta p}^{ij})(\Delta p_{\delta p}^{\ell k})^* \end{bmatrix} . \tag{10}$$

Then we can write,

$$\begin{aligned} \bar{E}(\Delta P^{ij}\Delta P_*^{\ell k}) &= \bar{E}(\varphi^T(q^{-1})\bar{p}_{ij}\bar{p}_{\ell k}^*\varphi_*^T(q)) \\ &= \varphi^T\mathbf{P}_{\Delta\mathbf{P}}^{(ij,\ell k)}\varphi_*^T \; . \end{aligned} \tag{11}$$

We collect all matrices of type (10) into one large covariance matrix, $\mathbf{P}_{\Delta\mathbf{P}}$, with ijth block given by

$$[\mathbf{P}_{\Delta\mathbf{P}}]_{ij} = \begin{bmatrix} \mathbf{P}_{\Delta\mathbf{P}}^{(i1,j1)} & \ldots & \mathbf{P}_{\Delta\mathbf{P}}^{(i1,jm)} \\ \vdots & \ddots & \vdots \\ \mathbf{P}_{\Delta\mathbf{P}}^{(im,j1)} & \ldots & \mathbf{P}_{\Delta\mathbf{P}}^{(im,jm)} \end{bmatrix} . \tag{12}$$

If ΔP has dimension $n|m$, then $\mathbf{P}_{\Delta\mathbf{P}}$ is composed of nm by nm covariance matricies $\mathbf{P}_{\Delta\mathbf{P}}^{(ij,\ell k)}$. The structure of $\mathbf{P}_{\Delta\mathbf{P}}$ is useful from a design point of view. If, for example, a multivariable moving average model, or FIR model, is to be identified, then $\mathbf{P}_{\Delta\mathbf{P}}$ is the natural way of representing the covariance matrix. If we instead prefer to use the blocks $\mathbf{P}_{\Delta\mathbf{P}}^{(ij,\ell k)}$ of $\mathbf{P}_{\Delta\mathbf{P}}$ as multivariable "tuning knobs", we can assign a given amount of uncertainty to a specific input–output pair.

The following assumption will be utilized.

A1. The coefficients of all polynomial elements in ΔC are independent of those in ΔB.[3]

3. Design of robust filters

The averaged spectral factorization

An averaged spectral factor $\beta(q^{-1})$ is defined as the numerator polynomial matrix of an averaged innovations model. It constitutes a key element of the robust filter. The average, over the set of models, of the spectral density matrix $\Phi_y(e^{i\omega})$ of the measurement $y(k)$ is given by

$$\bar{E}\{\Phi_y(e^{i\omega})\} = \frac{1}{DD_*}A^{-1}N^{-1}\beta\beta_*N_*^{-1}A_*^{-1} \; .$$

The square polynomial matrix $\beta(z^{-1})$ is given by the stable solution to

$$\beta\beta_* = \bar{E}\{NBCC_*B_*N_* + DAMM_*A_*D_*\} \; . \tag{13}$$

[3] Assumption **A1** could be excluded, but it does simplify the solution and it is also reasonable in most practical cases.

The following results are useful when solving (13).

Lemma 1: Let $\boldsymbol{H}(q,q^{-1})$ be a $m|m$ polynomial matrix with double-sided polynomial elements having stochastic coefficients. Also, let $\boldsymbol{G}(q^{-1})$ be an $n|m$ polynomial matrix with polynomial elements having stochastic coefficients, independent of $\boldsymbol{H}$. Then,

$$\bar{E}[\boldsymbol{G}\boldsymbol{H}\boldsymbol{G}_*] = \bar{E}[\boldsymbol{G}\bar{E}(\boldsymbol{H})\boldsymbol{G}_*] \ . \qquad (14)$$

Proof. See [11].

Now, define the double-sided polynomial matrices

$$\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_* \triangleq \bar{E}(\boldsymbol{C}\boldsymbol{C}_*) \ ; \ \tilde{\boldsymbol{M}}\tilde{\boldsymbol{M}}_* \triangleq \bar{E}(\boldsymbol{M}\boldsymbol{M}_*)$$
$$\tilde{\boldsymbol{B}}_C\tilde{\boldsymbol{B}}_{C*} \triangleq \bar{E}(\boldsymbol{B}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\boldsymbol{B}_*) \ . \qquad (15)$$

Invoking (6) and using the fact that the stochastic coefficients are assumed to be zero mean, gives

$$\begin{aligned} \tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_* &= \hat{\boldsymbol{C}}_o\hat{\boldsymbol{C}}_{o*} + \hat{\boldsymbol{C}}_1\bar{E}(\Delta\boldsymbol{C}\Delta\boldsymbol{C}_*)\hat{\boldsymbol{C}}_{1*} \\ \tilde{\boldsymbol{B}}_C\tilde{\boldsymbol{B}}_{C*} &= \hat{\boldsymbol{B}}_o\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*} + \hat{\boldsymbol{B}}_1\bar{E}(\Delta\boldsymbol{B}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\Delta\boldsymbol{B}_*)\hat{\boldsymbol{B}}_{1*} \\ \tilde{\boldsymbol{M}}\tilde{\boldsymbol{M}}_* &= \hat{\boldsymbol{M}}_o\hat{\boldsymbol{M}}_{o*} + \hat{\boldsymbol{M}}_1\bar{E}(\Delta\boldsymbol{M}\Delta\boldsymbol{M}_*)\hat{\boldsymbol{M}}_{1*} \ . \qquad (16) \end{aligned}$$

Factorizations of $\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*$ etc. need not be performed.

Lemma 2: Let Assumption **A1** hold. By using (15), (16) and invoking Lemma 1, the averaged spectral factorization (13) can be expressed as

$$\boldsymbol{\beta}\boldsymbol{\beta}_* = \boldsymbol{N}\tilde{\boldsymbol{B}}_C\tilde{\boldsymbol{B}}_{C*}\boldsymbol{N}_* + \boldsymbol{D}\boldsymbol{A}\tilde{\boldsymbol{M}}\tilde{\boldsymbol{M}}_*\boldsymbol{A}_*\boldsymbol{D}_* \ . \qquad (17)$$

Proof. See [11].

With a given right-hand side, equation (17) is just an ordinary polynomial matrix left spectral factorization. It is solvable under the following mild assumption

A2. The averaged spectral density matrix $\bar{E}\{\boldsymbol{\Phi}_y(e^{i\omega})\}$ is nonsingular for all ω.

This assumption is equivalent to the right-hand side of (17) being nonsingular on $|z| = 1$. Then, the solution to (17) is unique, up to a right orthogonal factor. Under **A2**, a solution exists, with $\boldsymbol{\beta}$ having a nonsingular leading coefficient matrix $\boldsymbol{\beta}(0)$. Its degree, $n\beta$, will be determined by the maximal degree of the two right-hand terms in (17).[4]

To obtain the right-hand side of (17), averaged polynomial matrices $\bar{E}(\Delta\boldsymbol{P}\boldsymbol{H}\Delta\boldsymbol{P}_*)$ have to be computed, where $\boldsymbol{H}(q,q^{-1}) = \tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*$ or $\mathbf{I}$. It is shown in [11] that the ij'th element of $\bar{E}(\Delta\boldsymbol{P}\boldsymbol{H}\Delta\boldsymbol{P}_*)$ is given by

$$\bar{E}[\Delta\boldsymbol{P}\boldsymbol{H}\Delta\boldsymbol{P}_*]_{ij} = \mathrm{tr}\boldsymbol{H} \begin{bmatrix} \varphi^T & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \varphi^T \end{bmatrix} \times$$

$$\begin{bmatrix} \mathbf{P}_{\Delta\mathbf{P}}^{(i1,j1)} & \cdots & \mathbf{P}_{\Delta\mathbf{P}}^{(im,j1)} \\ \vdots & \ddots & \vdots \\ \mathbf{P}_{\Delta\mathbf{P}}^{(i1,jm)} & \cdots & \mathbf{P}_{\Delta\mathbf{P}}^{(im,jm)} \end{bmatrix} \begin{bmatrix} \varphi_*^T & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \varphi_*^T \end{bmatrix} \qquad (18)$$

where φ^T is defined in (9). The block covariance matrix in (18) constitutes the block-transpose of the ijth block $[\,\cdot\,]$ of $\mathbf{P}_{\Delta\mathbf{P}}$, see (12). Thus, the average factors in (16) are readily obtained by substituting $\Delta\boldsymbol{C}$, $\Delta\boldsymbol{B}$ and $\Delta\boldsymbol{M}$ for $\Delta\boldsymbol{P}$ in (18).

The cautious multivariable Wiener filter

Theorem 1: Assume an extended design model (3), (5), (6), to be given, with known covariance matrices (12). Assume **A1** and **A2** to hold. A realizable estimator of $f(k)$ then minimizes (2), among all linear time-invariant estimators based on $y(k+m)$, if and only if it has the same coprime factors as

$$\hat{f}(k|k+m) = \boldsymbol{\mathcal{R}}\, y(k+m) = \frac{1}{T}\boldsymbol{V}^{-1}\boldsymbol{Q}\boldsymbol{\beta}^{-1}\boldsymbol{N}\boldsymbol{A}\, y(k+m) \ . \qquad (19)$$

Here, $\boldsymbol{\beta}(q^{-1})$ is obtained from (17), while $\boldsymbol{Q}(q^{-1})$ together with $\boldsymbol{L}_*(q)$, both of dimensions $\ell|p$, is the unique solution to the unilateral Diophantine equation

$$q^{-m}\boldsymbol{V}\boldsymbol{S}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*}\boldsymbol{N}_* = \boldsymbol{Q}\boldsymbol{\beta}_* + q\boldsymbol{L}_*UTD\mathbf{I}_p \qquad (20)$$

with generic[5] degrees

$$\begin{aligned} nQ &= \max(nv + ns + n\tilde{c} + m, nu + nt + nd - 1) \\ nL_* &= \max(n\tilde{c} + n\hat{b}_o + nn - m, n\beta) - 1 \end{aligned} \qquad (21)$$

where $ns = \deg \boldsymbol{S}$ etc. When applying (19) on an ensemble of systems, the minimal criterion value becomes

$$\mathrm{tr}\bar{E}E(\varepsilon(k)\varepsilon(k)^*)_{\min} = \mathrm{tr}\frac{1}{2\pi j}\oint_{|z|=1}\Big\{\boldsymbol{L}\boldsymbol{\beta}^{-1}\boldsymbol{\beta}_*^{-1}\boldsymbol{L}_*$$
$$+ \frac{1}{UTDD_*T_*U_*}\boldsymbol{V}\boldsymbol{S}\tilde{\boldsymbol{C}} \ \times$$
$$\Big[\mathbf{I}_n - \tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*}\boldsymbol{N}_*\boldsymbol{\beta}_*^{-1}\boldsymbol{\beta}^{-1}\boldsymbol{N}\hat{\boldsymbol{B}}_o\tilde{\boldsymbol{C}}\Big]\tilde{\boldsymbol{C}}_*\boldsymbol{S}_*\boldsymbol{V}_*\Big\}\frac{dz}{z} \ . \qquad (22)$$

Proof. See [11].

Remarks. The design equations are, (17), (18) and (20). The only new type of computation, as compared to the nominal case described in [1], is the calculation of averaged polynomials, using (18).

Since both $\boldsymbol{V}$ and $\boldsymbol{\beta}$ are stable, the estimator $\boldsymbol{\mathcal{R}}$ will be stable[6]. Furthermore, since $\boldsymbol{V}(0)$ and $\boldsymbol{\beta}(0)$ are assumed to be nonsingular, $\boldsymbol{\mathcal{R}}$ will be causal.

[4] When solving (17), we have utilized an algorithm by Ježek and Kučera, presented in [7]. It provides a solution with an upper triangular full rank leading coefficient matrix.

[5] In special cases, the degrees may be lower.

[6] Stable common factors may exist in (19). They could be detected by calculating invariant polynomials of the involved matrices. If such factors have zeros close to the unit circle, it is advisable to cancel them before the filter is implemented. Otherwise, slowly decaying initial transients may deteriorate the filtering performance.

Note that the diagonal matrix $NA = N_o N_1 A_o A_1$ appears explicitly in the filter (19). Important properties of the robust estimator are therefore evident by direct inspection. For example, assume some diagonal elements of N_1^{-1} or A_1^{-1} in the error models to have resonance peaks, indicating large uncertainty at some frequencies. Then, the filter will have notches, so the filter gain from the corresponding components of $y(k+m)$ will be low at the relevant frequencies.

The model structure (5)–(6) was selected to obtain a few simple design equations. Other choices are possible, but lead to various complications. For example, if stochastic polynomials had been introduced in the denominators, no exact analytical solution could have been obtained. The use of general left MFD representations, instead of forms with diagonal denominators or common denominators, would have led to a solution involving seven coprime factorizations. Such a solution reduces the possibility to obtain physical insight. It would also exhibit worse numerical behaviour, since algorithms for coprime factorization are numerically sensitive.

Robust design improves the numerical properties of the solution. Almost common factors of $\det \beta_*$ and UTD close to $|z|=1$ would make the solution of (20) numerically sensitive. Due to the presence of averaged factors in (16), the risk for this is less than in the nominal case.

4. A design example

Assume that a scalar signal $u(k)$, described by a first order AR–process without uncertainty, is to be estimated:

$$u(k) = \frac{1}{1-0.5q^{-1}} e(k) \quad ; \quad Ee(k)^2 = 1 \ .$$

Thus, $\mathcal{D} = \mathcal{S} = T = D_1 = \hat{C}_o = 1$, $\hat{C}_1 = 0$ and $D = D_o = 1 - 0.5q^{-1}$. This signal is measured by two transducers with nominal models being second order FIR–filters. The transducers are modelled by

$$y(k) = (B_o + A_1^{-1}\Delta B)u(k) + M_o v(k)$$

where

$$\begin{aligned} B_o &= \begin{pmatrix} B_o^{11} \\ B_o^{21} \end{pmatrix} = \begin{pmatrix} 0.100 + 0.080q^{-2} \\ 1 - 1.4q^{-1} + 0.92q^{-2} \end{pmatrix} \\ A_1 &= \begin{pmatrix} 1 & 0 \\ 0 & 1-0.6q^{-1} \end{pmatrix} \\ \Delta B &= \begin{pmatrix} \Delta B^{11} \\ \Delta B^{21} \end{pmatrix} \\ &= \begin{pmatrix} \Delta b_o^{11} + \Delta b_2^{11} q^{-2} \\ \Delta b_o^{21} + \Delta b_1^{21} q^{-1} + \Delta b_2^{21} q^{-2} \end{pmatrix} \\ M_o &= \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix} \ . \end{aligned}$$

Thus, $B_1 = A_o = N = I_2$ and $A = A_1$ are used in (6). In the first transducer B^{11}, there is only a single uncertain parameter. It affects the coefficients Δb_o^{11} and Δb_2^{11} with opposite signs, so they have zero mean, variance $r_1^2 = 0.02$ and cross covariance $-r_1^2$. In the second transducer, the stochastic coefficients are assumed mutually uncorrelated, with zero means and equal variance $r_2^2 = 0.10$. Coefficients of ΔB^{11} and ΔB^{21} are assumed mutually uncorrelated. Now, expression (11) gives

$$\begin{aligned} \bar{E}(\Delta B^{11} \Delta B_*^{11}) &= r_1^2(-q^2 + 2 - q^{-2}) \\ \bar{E}(\Delta B^{21} \Delta B_*^{21}) &= 3r_2^2 \ . \end{aligned}$$

Note that $\bar{E}(\Delta B^{11}\Delta B_*^{11})$ has zeros at $z=1$ and at $z=-1$. Thus, the static gain and the high–frequency gain is assumed to be exactly known. The channel B^{11} has its uncertainty concentrated around the notch at frequency $\omega = \pi/2$, while B^{21} is uncertain mainly at low frequencies. See Figure 2.

The goal is now to design a filter ($m=0$), which estimates $u(k)$ based on the two measurements $y_1(k)$ and $y_2(k)$. No frequency weighting is used ($\mathcal{W} = 1$).

A stable averaged spectral factor, which satisfies (17), with $\beta(0)$ nonsingular, is given by

$$\beta = \begin{pmatrix} 0.1339 - 0.01867q^{-1} + 0.01622q^{-2} & 0.07862 - 0.01488q^{-1} + 0.06905q^{-2} \\ -0.1474q^{-1} + 0.2908q^{-2} - 0.1325q^{-3} & 1.1585 - 2.0327q^{-1} + 1.6219q^{-2} - 0.4765q^{-3} \end{pmatrix} \ .$$

The solution to the Diophantine equation (20), is given by

$$\begin{aligned} Q &= (\ 0.4005 \quad 0.7746\ ) \\ L_* &= (0.0290 + 0.0200q \quad -0.2053 + 0.3224q - 0.1299q^2). \end{aligned}$$

Finally, the robust estimator (19) becomes

$$\mathcal{R} = Q\beta^{-1}A_1 = \frac{1}{R_r}(K_r^{11} \ K_r^{12})$$

where

$$\begin{aligned} K_r^{11} &= 2.9922 - 4.5138q^{-1} + 2.7365q^{-2} - 0.5687q^{-3} \\ K_r^{12} &= 0.4655 - 0.3341q^{-1} - 0.06445q^{-2} + 0.05841q^{-3} \\ R_r &= 1 - 1.8193q^{-1} + 1.6043q^{-2} - 0.6584q^{-3} \\ &\quad + 0.08479q^{-4} + 0.009182q^{-5} \ . \end{aligned}$$

A corresponding nominal estimator is given by $\mathcal{R}_n = (K_n^{11} \ K_n^{12})/R_n$, with

$$\begin{aligned} K_n^{11} &= 0.7419 - 1.0943q^{-1} + 0.3617q^{-2} \\ K_n^{12} &= 0.8792 - 0.3767q^{-1} - 0.03145q^{-2} \\ R_n &= 1 - 1.7786q^{-1} + 1.4269q^{-2} - 0.3938q^{-3} \ . \end{aligned}$$

From Figure 2 the following can be seen. The gains of the nominal estimator are determined exclusively by

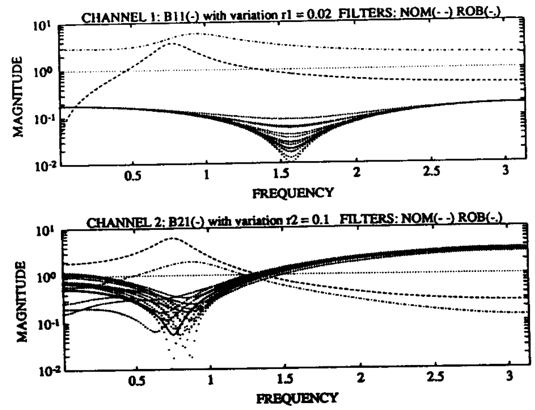

Figure 2: Bode magnitude plots for the nominal models of the two transducers $B^{11}(q^{-1})$ and $B^{21}(q^{-1})$ (solid). The dotted curves show fifteen realizations of possible true systems, assuming Gaussian distributions. Magnitude plots for the gains from $y_1(k)$ (upper) and $y_2(k)$ (lower), are shown for the robust (dash–dotted) and nominal Wiener filter (dashed).

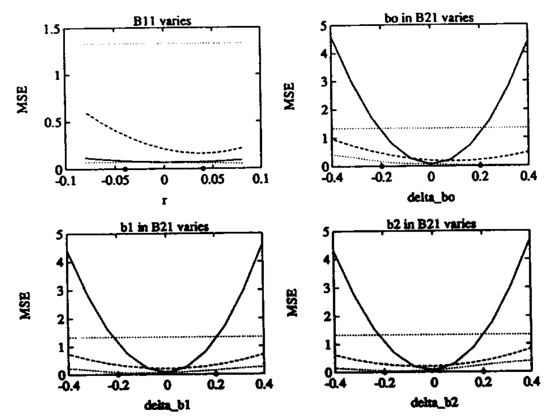

Figure 3: MSE for robust (dashed) and nominal filter (solid). Also shown is the variance of $u(k)$ (upper dotted). The lower dotted curve is the lower bound, achievable with knowledge of the true system. Rings (o) indicate the two standard deviation limits.

the nominal signal to noise ratios, while the gains of the robust estimator are determined by the balance between noise levels and model uncertainties in the two channels. For example, the robust filter "knows" that channel 1 is well known at low and high frequencies. Consequently, a higher gain is used from $y_1(k)$ as compared to the nominal case, and a lower gain from $y_2(k)$. The difference, as compared to nominal design, is largest at low frequencies. There, the dynamics of channel 1 is almost perfectly known, while channel 2 is very uncertain. The nominal filter gain in channel 2 is an approximate inverse of the nominal transducer. Compared to the nominal filter, the robust filter has much lower peak at the (uncertain) notch around $\omega = 0.7$. It utilizes channel 1 more at this frequency.

Figure 3 shows the mean square estimation error, when *one* of the uncertain parameters

$$\Delta b_o^{11} = -\Delta b_2^{11} \triangleq r \quad ; \quad \Delta b_o^{21} \quad ; \quad \Delta b_1^{21} \quad ; \quad \Delta b_2^{21}$$

is varied, while the others are zero. The four parameters span the set of assumed true systems, the extended design model. On average, over the four uncorrelated stochastic coefficients, the MSE is 0.32 for the robust filter and 0.90 for the nominal design.

The robust estimator does, of course, not perform as well as the nominal one in the nominal case. It is evident from Figure 3 that this performance loss is small, compared to the improvement in non–ideal situations.

5. References

[1] A. Ahlén and M. Sternad, "Wiener filter design using polynomial equations," *IEEE Transactions on Signal Processing*, vol 39, pp 2387–2399, 1991.

[2] R. C. Chung and P. R. Bélanger, "Minimum-sensitivity filter for linear time-invariant stochastic systems with uncertain parameters," *IEEE Transactions on Automatic Control*, vol 21, pp 98–100, 1976.

[3] J. A. D'Appolito and C. E. Hutchinson, "A minimax approach to the design of low sensitivity state estimators," *Automatica*, vol 8, pp 599–608, 1972.

[4] G. C. Goodwin, M. Gevers and B. Ninness, "Quantifying the error in estimated transfer functions with application to model order selection," *IEEE Transactions on Automatic Control*, vol 27, pp 913–928, 1992.

[5] G. C. Goodwin and M. E. Salgado, "A stochastic embedding approach for quantifying uncertainty in the estimation of restricted complexity models," *International Journal of Adaptive Control and Signal Processing*, vol 3, pp 333–356, 1989.

[6] M. J. Grimble, "Wiener and Kalman filters for systems with random parameters," *IEEE Transactions on Automatic Control*, vol 29, pp 552–554, 1984.

[7] J. Ježek and V. Kučera, "Efficient algorithm for matrix spectral factorization," *Automatica*, vol. 21, pp 663–669 1985.

[8] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: a survey," *Proceedings of the IEEE*, vol 73, pp 433- 481, 1985.

[9] V. Kučera, *Discrete Linear Control. The Polynomial Equations Approach.* Wiley, New York, 1979.

[10] C. T. Leondes and J. O. Pearson, "A minimax filter for systems with large plant uncertainties," *IEEE Transactions on Automatic Control*, vol 17, pp 266–268, 1972.

[11] K. Öhrn, A. Ahlén and M. Sternad, "A probabilistic approach to multivariable robust filtering and open–loop control," *Report UPTEC 93008R*, Systems and Control Group, Uppsala University, Sweden. Submitted.

[12] M. Sternad and A. Ahlén, "Robust filtering and feedforward control based on probabilistic descriptions of model errors", *Automatica*, vol 29, pp 661–679, 1993.