ROBUST WIENER FILTERING BASED ON PROBABILISTIC DESCRIPTIONS OF MODEL ERRORS

Mikael Sternad and Anders Ahlén *

Final version, August 1993

Abstract

A new approach to robust estimation of signals and prediction of time–series is considered. Possible modelling errors are described by sets of systems, parametrized by random variables, with known covariances. A robust design is obtained by minimizing the squared estimation error, averaged both with respect to model errors and noise. A polynomial solution, based on averaged spectral factorizations and averaged Diophantine equations, is derived. The robust estimator is called a cautious Wiener filter. It turns out to be no more complicated to design than an ordinary Wiener filter. The methodology can be applied to any open loop filtering or control problem.

*Systems and Control Group, Uppsala University, P. O. Box 27 S-751 03 Uppsala, Sweden.
The work has been partially supported by the Swedish Board for Technical Developement, under contract 8701573 and by the Swedish Institute. It was carried out while Anders Ahlén was on leave at The Department of Electrical Engineering and Computer Science, The University of Newcastle, Australia. The paper was presented at the 2nd IFAC workshop on System Structure and Control, Prague, Sept. 3–5 1992. A considerably expanded version of this paper can be found in the IFAC Journal Automatica, May 1993.

1 INTRODUCTION

In *robust filter* synthesis, the ever present model uncertainty, and the whole range of expected system behaviour, is taken into account. We here propose a novel approach to robust design for signal estimation. It is based on a stochastic description of model errors, related to the stochastic embedding concept of Goodwin and Salgado [8]. A single robust filter, for the whole class of possible models, is obtained by minimizing the squared estimation error, averaged both with respect to model errors and the noise.

Most previous suggestions for robust filter design have been based on the minimax approach. Minimax design becomes very complex unless there exists a saddle–point solution. One may then search for a least favourable pair of signal and noise spectra, in prespecified uncertainty classes. The optimal estimator is a filter designed for that pair. See [6], [12], [17], [19], [20], [22], [27], and the survey paper by Kassam and Poor [13]. Uncertainties can be described in a state space framework. See e.g. [10], [18] and [28]. The computational effort involved in minimax–design is considerable. Closed–form solutions do mostly not exist.

Apart from leading to a much simpler design methodology, the approach proposed here avoids two drawbacks of robust minimax design. First, the descriptions of model uncertainties may have statistical, *"soft" bounds*. These are more readily obtainable in a noisy environment than the hard bounds required for minimax design. Secondly, not only the range of uncertainties, but also their likelihood is taken into account; probable model errors will have a greater impact on an estimator design than do very rare "worst cases". The conservativeness is thus reduced.

A polynomial solution, based on averaged spectral factorizations and averaged Diophantine equations, will be presented. The design procedure constitutes a generalization of the polynomial equations approach, which was pioneered by Kučera [14]. Mild solvability conditions guarantee the existence of stable optimal filters. The robust design turns out to be no more complicated than the design of an ordinary Wiener filter. The methodology is here exemplified on a scalar discrete–time deconvolution problem. Robust design for this and other related problems, such as state estimation and feedforward control, is discussed in more detail in [26].

Remarks on the notation. For any complex polynomial in the backward shift operator q^{-1}, of degree np,

$$P(q^{-1}) = p_o + p_1 q^{-1} + \ldots + p_{np} q^{-np}$$

the *conjugate polynomial* is defined as $P_*(q) \triangleq p_o^* + p_1^* q + \ldots + p_{np}^* q^{np}$. In the frequency domain, z is substituted for q. For convenience, polynomial arguments will often be omitted. We call $P(q^{-1})$ *stable* if all zeros of $P(z^{-1})$ are in $|z| < 1$.

2 THE ESTIMATION PROBLEM

A scalar discrete–time deconvolution problem will be considered, to illustrate the design principles. It includes e.g. ordinary output filtering and prediction of ARMA–processes as special cases. It also includes the design of linear recursive equalizers for digital communications. Measurements are described as

$$y(t) = \mathcal{G}(q^{-1})u(t-k) + w(t) \quad . \tag{2.1}$$

The linear, causal and possibly uncertain transfer function $\mathcal{G}(q^{-1})$ may e.g. represent a transducer. If the delay is uncertain, k denotes its minimum value. The input $u(t)$ and the measurement noise $w(t)$ are described by possibly uncertain ARMA–models

$$u(t) = \mathcal{F}(q^{-1})e(t) \quad ; \quad w(t) = \mathcal{H}(q^{-1})v(t) \tag{2.2}$$

$$E|e(t)|^2 = 1 \quad ; \quad E|v(t)|^2 \triangleq \rho \quad .$$

The time–series $e(t)$ and $v(t)$ are assumed mutually uncorrelated. They are stationary white noises or impulse sequences, with zero mean. All transfer functions are assumed time–invariant. Signals may be complex–valued; this is the case eg. in digital communications applications.

A stable, linear and time–invariant estimator of $u(t)$, given $y(t+m)$, is sought:

$$\hat{u}(t|t+m) = \frac{Q}{R}y(t+m) \quad . \tag{2.3}$$

It could be either a predictor ($m < 0$), a filter ($m = 0$) or a fixed lag smoother ($m > 0$).

3 PROBABILISTIC ERROR MODELS

The transfer functions $\mathcal{G}$, $\mathcal{F}$ and $\mathcal{H}$ may be uncertain. An *error model* is a quantification of the model error class. Together with a nominal model, it constitutes an *extended design model* on which a robust design is based. As error models, we will utilize additive transfer functions $\Delta\mathcal{F}, \Delta\mathcal{G}, \Delta\mathcal{H}$, with stochastic numerators and pre–specified denominators. This choice is crucial for obtaining a simple solution to the filtering problem. The extended design models are parametrized as

$$\begin{aligned}
\mathcal{F} &= \frac{C_0}{D_0} + \frac{C_1\Delta C}{D_1} = \frac{C_0D_1 + D_0C_1\Delta C}{D_0D_1} \triangleq \frac{C}{D} \\
\mathcal{G} &= \frac{B_0}{A_0} + \frac{B_1\Delta B}{A_1} = \frac{B_0A_1 + A_0B_1\Delta B}{A_0A_1} \triangleq \frac{B}{A} \\
\mathcal{H} &= \frac{M_0}{N_0} + \frac{M_1\Delta M}{N_1} = \frac{M_0N_1 + N_0M_1\Delta M}{N_0N_1} \triangleq \frac{M}{N}
\end{aligned} \tag{3.1}$$

Above, C_0/D_0 etc. represent nominal models, with degrees nc_0, nd_0 etc. They are assumed known and stable. Stable "error denominators" D_1, A_1 and N_1, of degrees nd_1, na_1 and nn_1, as well as numerator factors C_1, B_1, M_1, of degrees nc_1, nb_1 and nm_1, may be specified by the designer or obtained from data. Coefficients of numerator polynomials

$$\Delta P(q^{-1}) = \Delta p_0 + \Delta p_1 q^{-1} + \ldots + \Delta p_{\delta p} q^{-\delta p} \tag{3.2}$$

are stochastic variables, with zero means. Let $\bar{E}(\cdot)$ denote expectation with respect to coefficients of stochastic polynomials ΔP. Parameter covariances are thus denoted $\bar{E}\Delta p_i \Delta p_j^*$. They are collected in the covariance matrix $\mathbf{P}_{\Delta P}$.

The coefficients of $\Delta C, \Delta B$ and ΔM are constant in time, so they are independent of the time series $e(t)$ and $v(t)$. Except for first and second order moments, their distributions need not be known, since they will not affect the design.

All polynomial degrees are assumed known to (or specified by) the designer [1]. Note also that denominator polynomials are *all assumed stable.* Uncertainty more or less forces us to restrict attention to stable extended design models [2].

In the sequel, we utilize two mild assumptions:

A1. The coefficients of ΔC and of ΔB are independent.

A2. The covariance matrices $\mathbf{P}_{\Delta C}, \mathbf{P}_{\Delta B}$ and $\mathbf{P}_{\Delta M}$ are Hermitian and positive semidefinite.

It is necessary to assure A2 when the covariance matrices are used pragmatically, as "robustness tuning knobs". Design equations could be derived for situations with correlations between ΔC and ΔB. Assumption A1 does, however, simplify the solution, and seems reasonable.

Model error covariances may be obtained from identification experiments, or from frequency domain data on system variability. See [7], [8] and [26]. If a fixed filter is to be designed for a large number of systems, the statistics may be obtained from a representative sample of systems. (Under the name of "statistical quality control", sampling techniques for obtaining means and variances of important properties are becoming increasingly widespread within manufacturing industries.)

Probabilistic error models remain useful also when statistics is hard to obtain. Those who prefer a Bayesian view could then interprete error distributions as subjective probabilities. Others may just use them pragmatically, as robustness

[1] Note that we are talking about an extended design *model*. In practice, it will only be an approximation of a class of possibly infinite dimensional and time–varying true systems.

[2] If unstable poles were exactly known, a finite estimation error could be obtained, by a filter which cancels unstable poles by zeros in the total signal path to the estimation error. Such a strategy is, of course, highly non–robust to mis–modelling of unstable poles. With *uncertain* unstable poles, the design problem becomes unsolvable, in the open–loop context considered here. (Therefore, a general solution involving two coupled Diophantine equations will not be of interest here.)

"tuning knobs". The covariances are then altered until satisfactory spectral properties of the filter are obtained.

If only the signal to noise ratio is uncertain, we may set $\Delta C = \Delta B = \Delta M = 0$, and use a *higher equivalent noise variance.* A model (3.1), with uncertain noise variance but well–defined noise spectrum, is given by $M/N = M_0/N_0 + M_0 \Delta m/N_0 = (M_0/N_0)(1 + \Delta m)$, with a scalar stochastic Δm. It corresponds to regarding the noise as having variance $\rho(1 + \bar{E}|\Delta m|^2)$.

Another special case is the use of FIR- or MA–filters (i.e. no denominators):

$$u(t) = (C_0 + \Delta C)e(t) \quad ; \quad y(t) = q^{-k}(B_0 + \Delta B)u(t) + (M_0 + \Delta M)v(t) \quad . \qquad (3.3)$$

In (3.3), degrees of stochastic polynomials may be set higher than those of the nominal polynomials, $\delta c > nc_0$, etc. This can be used to guard against under–parametrization. However, for systems with long or infinite impulse responses, error models with denominators are more appropriate than FIR–models.

The structure (3.1) covers multiplicative as well as additive descriptions of model errors. A multiplicative error is obtained with e.g. $B_1 = B_0 B_m, A_1 = A_0 A_m$, with B_m arbitrary and A_m stable.

4 DESIGN OF ROBUST FILTERS

We proceed from the model (2.1), (2.2) and (3.1). The coefficients of ΔC, ΔB and ΔM are random variables, whose possible values parametrize a set of systems. We will minimize the *averaged mean square error (MSE) criterion*

$$\bar{E}(E|z(t)|^2) \quad ; \quad z(t) \triangleq \frac{V}{U}(u(t) - \hat{u}(t|t+m)) \quad . \qquad (4.1)$$

Here, E represents expectation over noise and $\bar{E}$ is an expectation over the model error distribution, i.e. over the coefficients of ΔC, ΔB and ΔM. We thus seek a single estimator which provides the best MSE performance, on average, when applied on randomly selected systems within the specified class. Above, V/U represents a frequency dependent weighting, with both V and U being stable polynomials. In the *nominal case* (no uncertainty assumed), (4.1) reduces to an ordinary frequency weighted MSE criterion.

The averaged MSE has been used in connection to other robust filtering formulations, e.g. by Chung and Bélanger [5], Speyer and Gustafson [23] and by Grimble [9]. These works were based on the assumption of small uncertainties and on series expansion of uncertain poles. In our design philosophy, we start from a model structure (3.1), and adjust it to the uncertainty directly. Large uncertainties can then be described in a much better way. (Examples are discussed in [26].)

4.1 The averaged spectral factorization

An *averaged spectral factor* $\beta(q^{-1})$ is defined as the numerator polynomial of an averaged innovations model of the measurement. In the present case, it is given by the stable and monic solution to

$$r\beta\beta_* \triangleq \bar{E}\{CC_*BB_*NN_* + \rho MM_*AA_*DD_*\} \tag{4.2}$$

with scalar r. Define double–sided polynomials

$$\tilde{C}\tilde{C}_* \triangleq \bar{E}(CC_*),\ \ \tilde{B}\tilde{B}_* \triangleq \bar{E}(BB_*),\ \ \ \tilde{M}\tilde{M}_* \triangleq \bar{E}(MM_*)\ \ . \tag{4.3}$$

Then, the use of (3.1) gives

$$\begin{aligned}\tilde{C}\tilde{C}_* &= C_0C_{0*}D_1D_{1*} + D_0D_{0*}C_1C_{1*}\bar{E}(\Delta C\Delta C_*)\\ \tilde{B}\tilde{B}_* &= B_0B_{0*}A_1A_{1*} + A_0A_{0*}B_1B_{1*}\bar{E}(\Delta B\Delta B_*)\\ \tilde{M}\tilde{M}_* &= M_0M_{0*}N_1N_{1*} + N_0N_{0*}M_1M_{1*}\bar{E}(\Delta M\Delta M_*)\ \ .\end{aligned} \tag{4.4}$$

We can now simplify (4.2).

Lemma 1. Let assumption A1 hold. Then, (4.2) can be expressed as

$$r\beta\beta_* = \tilde{C}\tilde{C}_*\tilde{B}\tilde{B}_*NN_* + \rho\tilde{M}\tilde{M}_*AA_*DD_* \tag{4.5}$$

■

Proof: The coefficients of a polynomial ΔP are zero mean stochastic variables. Coefficients of $\Delta P\Delta P_*$ will also be stochastic variables, having expected values given by (4.7) below. The coefficients of (ΔB, ΔC), are independent, and so are the coefficients of $\Delta B\Delta B_*$, $\Delta C\Delta C_*$. Using independence for complex parameters, the right–hand side of (4.2) becomes

$$\bar{E}(CC_*)\bar{E}(BB_*)NN_* + \rho\bar{E}(MM_*)AA_*DD_*$$

■

The averaged factors in (4.4) can be evaluated as follows. For a stochastic error model numerator $\Delta P(q^{-1})$, as in (3.2), let the Hermitian parameter covariance matrix be

$$\mathbf{P}_{\Delta P} = \begin{bmatrix} \bar{E}|\Delta p_0|^2 & \dots & \bar{E}(\Delta p_0\Delta p^*_{\delta p}) \\ \vdots & \ddots & \vdots \\ \bar{E}(\Delta p_{\delta p}\Delta p^*_0) & \dots & \bar{E}|\Delta p_{\delta p}|^2 \end{bmatrix} \tag{4.6}$$

Denote the sum of the diagonal elements h_0, the sum of elements in the i'th super–diagonal h_i, and the sum of elements in the i'th subdiagonal h_{-i}. Note that $h_{-i} = h_i^*$. Then it becomes evident, by direct multiplication of $\Delta P(q^{-1})\Delta P_*(q)$, and taking expectations, that

$$\bar{E}(\Delta P\Delta P_*) =$$

$$h^*_{dp}q^{-dp} + \ldots + h^*_1q^{-1} + h_0 + h_1q + \ldots + h_{dp}q^{dp}\ \ . \tag{4.7}$$

Thus, the averaged factors in (4.4) are readily obtained. Above, $dp \leq \delta p$ (the degree of $\Delta P(q^{-1})$). If coefficients are uncorrelated, $dp = 0$.

In (4.4), $\tilde{C}\tilde{C}_*$ will contain powers up to $q^{\pm n\tilde{c}}$, where $n\tilde{c} = \max\{nc_0+nd_1, nd_0+nc_1+\delta c\}$, with analogous expressions for $n\tilde{b}, n\tilde{m}$. Since $N = N_0N_1$ etc, the averaged spectral factor in (4.5) has degree

$$n\beta = \max\{n\tilde{c} + n\tilde{b} + nn_0 + nn_1, n\tilde{m} + na_0 + na_1 + nd_0 + nd_1\}$$

The factorization (4.5) is solvable with respect to a unique stable $\beta(z^{-1})$ iff its right–hand side is positive on $|z| = 1$. Introduce the assumptions

A3. C_0, $C_1\bar{E}(\Delta C\Delta C_*)$, ρM_0 and $\rho M_1\bar{E}(\Delta M\Delta M_*)$ have no common zeros on $|z| = 1$

A4. B_0, $B_1\bar{E}(\Delta B\Delta B_*)$, ρM_0 and $\rho M_1\bar{E}(\Delta M\Delta M_*)$ have no common zeros on $|z| = 1$.

Lemma 2. Let D, A and N be stable and assumption A2 hold. Then, a unique stable spectral factor β, satisfying (4.5), exists, if and only if both of assumption A3 and A4 are true ■

Proof: See Appendix A.

The conditions A3 and A4 are mild. They will almost always be fulfilled, even if C_0, B_0 and M_0 have zeros on the unit circle. In fact, the conditions are more relaxed than for the nominal case, due to the presence of averaged factors $\bar{E}(\cdot)$.

4.2 The cautious Wiener filter

Theorem 1. Assume an extended design model (2.1)–(3.1) to be given, with known covariances of the stochastic polynomial coefficients. Assume A1–A4 to hold. An estimator of $u(t)$ then minimizes (4.1), among all linear time–invariant estimators based on $y(t+m)$, if and only if it has the same coprime factors as

$$\hat{u}(t|t+m) = \frac{Q}{R}y(t+m) \quad ; \quad \frac{Q}{R} = \frac{Q_1N_0N_1A_0A_1}{\beta V} \tag{4.8}$$

Here, $\beta(q^{-1})$ is obtained from (4.5), while $Q_1(q^{-1})$, together with $L_*(q)$, is the unique solution to

$$q^{-m+k}V\tilde{C}\tilde{C}_*B_{0*}A_{1*}N_{0*}N_{1*} = r\beta_*Q_1 + qUD_0D_1L_* \tag{4.9}$$

with polynomial degrees

$$\begin{aligned} nQ_1 &\leq \max(nv + n\tilde{c} - k + m \ , \ nu + nd_0 + nd_1 - 1) \\ nL &\leq \max(n\tilde{c} + nb_0 + na_1 + nn_0 + nn_1 + k - m \ , \ n\beta) - 1 \ . \end{aligned} \tag{4.10}$$

For an ensemble of systems, the minimal criterion value becomes

$$\bar{E}E|z(t)|^2_{\min} = \frac{1}{2\pi i}\oint_{|z|=1} \frac{VV_*}{UU_*}\left[\frac{LL_*}{r\beta\beta_*} + \rho\frac{\tilde{C}\tilde{C}_*\tilde{M}\tilde{M}_*AA_*}{r\beta\beta_*} + \right.$$

$$\left. + \frac{\tilde{C}\tilde{C}_*\tilde{C}\tilde{C}_*\bar{E}(\Delta\mathcal{G}\Delta\mathcal{G}_*)AA_*NN_*}{DD_*r\beta\beta_*}\right]\frac{dz}{z} \tag{4.11}$$

■

Proof: See Appendix B. For a derivation of (4.11), see [26]. (The expression derived in [26] is slightly different, due to the absence of a weighting V/U.)

Remarks. The equations for minimizing (4.1) are (4.5), (4.7) and (4.9). The corresponding equations for the nominal case (no uncertainty) can be found in e.g. [1] or [26]. In a robust design, the only new type of computation, as compared to a nominal solution, is trivial: summation of covariance matrix elements, diagonal-wise.

Note that the "error denominators" N_1 and A_1 affect the filter (4.8) directly. If $1/N_1$ or $1/A_1$ in the error models have resonance peaks, indicating large uncertainty, the filter (4.8) will have low gain at those frequencies. With increasing model uncertainty, the zeros of β are moved inward in the unit circle. Resonance peaks of the estimator are lowered and broadened.

Equation (4.9) will have a unique solution, with degrees (4.10). Note that $\beta_*(z)$ (unstable) and $U(z^{-1})D_0(z^{-1})D_1(z^{-1})$ (stable) have no common factors.

The averaged estimation error (4.11) consists of three terms. Term 1 represents the effect of finite smoothing lag m. It can be shown that $L_*(q) \to 0$ when $m \to \infty$. The second term mainly represents the effect of noise. It vanishes for $\rho = 0$. Finally, the third term represents degradation caused by errors $\Delta\mathcal{G} = B_1\Delta B/A_1$ in the transducer model. It vanishes only when $\Delta\mathcal{G} = 0$.

In situations with little noise and sufficiently large smoothing lag m, term 3 in (4.11) will dominate the error. This is not surprising; a deconvolution smoother then essentially inverts $\mathcal{G}$. This operation is sensitive to model errors there.

4.3 Analytical expressions for performance evaluation

Theorem 2. Let a nominal estimator Q_0/R_0 be designed based on a nominal model, as in eg. [1],[2] Applying it, instead of (4.8), on an ensemble of systems results in an increase, compared to (4.11), of the mean MSE $\bar{E}E|z(t)|^2$. The increase is given by

$$\bar{E}E|z(t)|^2_0 - \bar{E}E|z(t)|^2_{\min} = \frac{r}{2\pi i}\oint_{|z|=1}\left|\frac{V}{U}\right|^2\left|\frac{\beta}{DAN}\right|^2\left|\frac{Q_0}{R_0} - \frac{Q}{R}\right|^2\frac{dz}{z} \tag{4.12}$$

where r, β are defined by (4.2)-(4.5) and Q/R is the optimal robust filter, given by (4.8). ■

Proof: To obtain (4.12), the nominal filter Q_0/R_0 is expressed as $Q/R+(Q_0/R_0-Q/R)$. The optimality of Q/R implies that any modification of it gives an orhtogonal contribution to the criterion. This, and the use of (4.2), gives (4.12). Mixed terms vanish, due to the orthogonality. ∎

Theorem 3. Let a robust estimator Q/R be designed by (4.5)–(4.9). When applying it on a system equal to the nominal model, the increased MSE, compared to the minimum, is

$$E|z(t)|^2 - E|z(t)|_0^2 = \frac{r_0}{2\pi i}\oint_{|z|=1}\left|\frac{V}{U}\right|^2\left|\frac{\beta_0}{D_0A_0N_0}\right|^2\left|\frac{Q}{R}-\frac{Q_0}{R_0}\right|^2\frac{dz}{z} \tag{4.13}$$

where $\beta_0/D_0A_0N_0$ is the nominal innovations model, and where r_0 and β_0 are obtained from the spectral factorization of the nominal model

$$r_0\beta_0\beta_{0*} = C_0C_{0*}B_0B_{0*}N_0N_{0*} + M_0M_{0*}A_0A_{0*}D_0D_{0*}$$

∎

Proof: Analogous to Theorem 2, by expressing Q/R as $Q_0/R_0+(Q/R-Q_0/R_0)$ ∎

Remarks. The expression in (4.12) can be used for arbitrary linear estimators Q_0/R_0, for example minimax–designs. We thus do not have to evaluate the mean performance of alternative designs by Monte–Carlo simulation. The filter magnitude $|V/U|^2$, together with the magnitude of the mean innovations model β/DAN in (4.12), and the nominal innovations model $\beta_0/D_0A_0N_0$ in (4.13), can be seen as weighting functions. In frequency regions where their magnitude is large, differences between the two estimators will have a large impact on the performance.

If the variance of broad–band measurement noise is increased, the gains of both the nominal and the robust filters decrease. If the noise level is high, performance differences between nominal and robust solutions tend to be small.

5 A NUMERICAL EXAMPLE

Consider an extended design model (3.3), with a FIR nominal model given by

$$C_0(q^{-1}) = 1-0.95q^{-1} \quad ; \quad B_0(q^{-1}) = 0.5-0.4q^{-1} \quad ; \quad M_0(q^{-1}) = 1-0.8q^{-1}$$
$$k=1 \quad , \quad \rho = 0.001 \quad . \tag{5.1}$$

The covariance matrices of $\Delta C, \Delta B$ and ΔM are

$$\mathbf{P}_{\Delta C} = \begin{bmatrix} 0 & 0 \\ 0 & 0.040 \end{bmatrix} \; ; \; \mathbf{P}_{\Delta B} = \begin{bmatrix} 0.0025 & 0 \\ 0 & 0.0225 \end{bmatrix} \; ; \; \mathbf{P}_{\Delta M} = \begin{bmatrix} 0 & 0 \\ 0 & 0.010 \end{bmatrix}$$

This corresponds to standard deviations 0.20 for the coefficient c_1, 0.05 for b_0, 0.15 for b_1 and 0.10 for m_1. We would like to obtain a robust filter $\hat{u}(t|t) = (Q/R)y(t)$.

This estimator should, essentially, perform a one step prediction to obtain $u(t)$, since the transducer $q^{-k}B(q^{-1})$ has a one step delay. Using (4.7), we obtain

$$\bar{E}(\Delta C \Delta C_*) = 0.040 \quad , \quad \bar{E}(\Delta B \Delta B_*) = 0.0250 \quad , \quad \bar{E}(\Delta M \Delta M_*) = 0.010 \quad . \tag{5.2}$$

Inserting (5.2) into (4.5) gives

$$r\beta\beta_* = (1.9425 - 0.95(q+q^{-1}))(0.435 - 0.2(q+q^{-1})) \; + \; 0.001(1.65 - 0.8(q+q^{-1})).$$

By solving for the stable monic $\beta(q^{-1})$ and r, we obtain

$$\beta(q^{-1}) = 1 - 1.4659q^{-1} + 0.5315q^{-2} \quad , \quad r = 0.3575 \quad .$$

Proceed to calculate the filter polynomial $Q_1(q^{-1})$ from (4.9), in which $U = V = A_1 = N_0 = N_1 = D_0 = D_1 = 1$. With degrees $nQ_1 = 0, nL = 2$, we obtain, with $Q_1(q^{-1}) = c$,

$$q(1.9425 - 0.95(q + q^{-1}))(0.5 - 0.4q) \; =$$

$$0.3575(1 - 1.4659q + 0.5315q^2)c + q(\ell_0 + \ell_1 q + \ell_2 q^2 \;) \quad .$$

Equating for different powers of q gives

$$Q_1(q^{-1}) = -1.3288 \quad ; \quad L_*(q) = 0.6549 - 0.9995q + 0.380q^2 \quad .$$

The robust estimator (4.6) is $\hat{u}(t|t) = (Q_1/\beta)y(t)$, or

$$\hat{u}(t|t) = \frac{-1.3288}{1 - 1.4659q^{-1} + 0.5315q^{-2}} y(t) \quad . \tag{5.3}$$

It has poles in $z = 0.8086$ and $z = 0.6573$. The nominal filter polynomials are

$$Q_{nom}(q^{-1}) = -1.8413 \quad ; \quad R_{nom}(q^{-1}) = 1 - 1.7206q^{-1} + 0.7365q^{-2} \tag{5.4}$$

with estimator poles in $z = 0.9206$ and $z = 0.80$. The robust estimator has decreased the gain and moved the poles inward. It has become more cautious.

The performance of nominal and robust estimators is exemplified in Fig. 1. The MSE is much lower for the robust filter, for most parameter values. For a wide range of parameter variations, the performance of the robust estimator is close to that which could be obtained if the true parameters were known.

In Fig. 2, the mean performance (4.1) was calculated, for nominal and robust estimators, as a function of the standard deviations of some of the parameters. As expected, the mean performance is much better for the robust filter for large parameter deviations. In situations with small model errors, we might just as well use the nominal estimator. This is also true when the noise level is high; when $Ev(t)^2 \geq 0.1$, the difference between robust and nominal filters is very small.

Figure 1. Variation of one parameter, while the others are nominal. The MSE $Ez(t)^2$, as a function of the zero $-c_1$ of the signal model (left) and of $-b_1$ in the transducer (right). Performance of a nominal design (solid), of the robust filter (4.8) (dashed) is compared to the lower bound (dotted), achievable with the knowledge of the true parameter value.

Figure 2. Mean performance $\bar{E}Ez(t)^2$, when standard deviations of c_1 (left) and either b_0 or b_1 (right) are varied, while other coefficients are nominal. Nominal filters (solid) are compared to robust filters, designed for the corresponding parameter variance (dashed).

In Fig. 3, we compare with minimax–designs, assuming two distributions, both with the same variances as above.

1) rectangular distribution, with hard bound at $= \sqrt{3}$standard deviations

2) 5–point distribution for each parameter, with bound ± 0.5 for $c_1, b_1, m_1, \pm 0.2$ for b_0.

In the minimax design, filters were constructed for the worst case (giving highest

MSE). When these filters were applied to other systems in the class, the MSE was never higher than in the design case. Thus, a saddle point condition was fulfilled, and the minimax solution had been found[3]. The worst case for the rectangular distribution was

$$C(q^{-1}) = 1 - 1.3q^{-1} \quad , \quad B(q^{-1}) = 0.41 - 0.66q^{-1} \quad , \quad M(q^{-1}) = 1 - 0.63q^{-1}$$

and for the 5–point distribution

$$C(q^{-1}) = 1 - 1.45q^{-1} \quad , \quad B(q^{-1}) = 0.3 - 0.9q^{-1} \quad , \quad M(q^{-1}) = 1 - 0.3q^{-1} \quad .$$

Fig. 3 clearly reveals the weaknesses of minimax designs: for wide model error distributions, with unlikely remote values, the filter performance for the nominal case deteriorates. More reasonable filters are obtained when the most remote value is close to the standard deviation of the distribution. However, even the assumption of a rectangular distribution results in a more conservative design than our cautious Wiener filter.

The simplest way of robustifying an estimator is to just increase the noise variance ρ used in a nominal design. The performance of this technique is also illustrated in Fig. 3. The result around the nominal case is not as good as for our robust filter. The difference can be expected to be even larger in more high–order examples. The use of just one single robustification parameter mostly provides insufficient degrees of freedom. It can only vary the spectral factor β along a single root locus trajectory

An extensive simulation study of the performance of robust filters is presented in [26].

Figure 3. Error variance $Ez(t)^2$ for nominal (solid) and robust (dashed) filters, as in Figure 1. We also illustrate a minimax–design assuming rectangular distributions (dashed–dotted), a minimax–design based on 5–point distributions (crosses) and also a nominal filter, designed using $\rho = 0.1$ (dotted).

[3]On closer investigation, this turned out to be a fortunate coincidence. In many similar examples, minmax $\neq$ maxmin, so the above technique cannot be used. A minimax MSE design then involves an exhaustive numerical search, and not even the polynomial degrees of the optimal filter can be determined a priori.

6 CONCLUSIONS

Estimation, based on imperfectly known linear discrete–time models, has been considered. Model errors were represented as additive transfer functions, with random numerators. A robust design was obtained by minimizing the squared estimation error, averaged both with respect to model errors and noise. This allows large but unlikely model errors to be taken into account, without dominating the design. The resulting filter becomes cautious, but not conservative.

With the presented polynomial equations approach, the robust filter design becomes simple and straightforward: just sum elements along diagonals of covariance matrices. Then, solve somewhat modified, “averaged”, spectral factorizations and Diophantine equations.

The methodology can be applied to any open loop problem. Robust state estimation, feedforward regulation, servo feedforward design and model matching are discussed in [26]. Application to the related problem of decision feedback equalization, discussed in [24], is straightforward. See [25]. Singular and non–singular continuous–time problems are presently under investigation. Multivariable design is discussed in [29].

With very large system variations, the performance of even a robust linear filter will be unsatisfactory. An approach analogous to gain scheduling in feedback control can then be of use. A bank of filters is designed, with each filter attuned to a subset of the total system class. By using the output or auxiliary information, the most likely subset is selected, and the corresponding filter is used. See e.g. [16] or [21]. Robust design is a complement to this appoach. By utilizing robust filters, which give acceptable behaviour for large model subsets, the number of filters in the filter bank may be reduced.

Robust design could also complement adaptation. Adaptive robust filtering, based on on–line estimation of nominal models and also of error model parameters, is a challenging subject. It is a main goal motivating our present research.

APPENDICES:

A Proof of Lemma 2.

Introduce a vector

$$f(\omega) \triangleq (\ 1\ e^{i\omega} \dots\ e^{i\omega dp}\ )^T \quad .$$

Then,

$$f^*(\omega)\mathbf{P}_{\Delta P}f(\omega) = g_{dp}^* e^{-i\omega dp} + \dots + g_1^* e^{-i\omega} + g_0 + g_1 e^{i\omega} + \dots + g_{dp}e^{i\omega dp} \quad .$$

This real–valued scalar is the polynomial $\bar{E}(\Delta P \Delta P_*)$ from (4.7), evaluated on $z = e^{i\omega}$. Thus, since $\mathbf{P}_{\Delta C}, \mathbf{P}_{\Delta B}$ and $\mathbf{P}_{\Delta M}$ are all assumed positive semidefinite, the corresponding polynomials from (4.7) will be non–negative on the unit circle. This is a sufficient condition for the expressions (4.4) to be non–negative on the unit circle. If $M_1\bar{E}(\Delta M \Delta M_*)$ has no zeros on the unit circle, the same will then be true for $\tilde{M}\tilde{M}$: it can only have zeros on $|z| = 1$ which are common to $M_0 M_{0*}$ and $M_1 M_{1*}\bar{E}(\Delta M \Delta M_*)$. Such zeros are zeros also of $\tilde{C}\tilde{C}_*$ or $\tilde{B}\tilde{B}_*$ in (4.5) only if either of A3 or A4 is violated. This holds for the special case $\rho = 0$ as well.

B Proof of Theorem 1.

A technique for simple constructive derivation of polynomial design equations for Wiener filters is presented in [2], [3] and [4]. It is based on the evaluation of orthogonality in the frequency domain. This technique will be utilized here, and be shown to be applicable to the averaged criterion (4.1).

With (2.1)–(3.1), the estimation error $z(t)$ in (4.1) is

$$z(t) = \frac{V}{U}\left(1 - q^{m-k}\frac{Q}{R}\frac{B}{A}\right)\frac{C}{D}e(t) - q^m \frac{V}{U}\frac{QM}{RN}v(t) \quad . \tag{B.1}$$

All admissible alternatives to a proposed weighted estimate, given by (2.3), can be described by

$$\hat{d}(t) = \frac{V}{U}\hat{u}(t) + n(t) \quad ; \quad \hat{u}(t) = \frac{Q}{R}y(t+m) \quad ; \quad n(t) = \mathcal{M}y(t+m) \quad .$$

Here, $\mathcal{M}(q^{-1})$ is a causal, stable but otherwise arbitrary rational function. Optimality of (2.3) is obtained if no perturbation $n(t)$ can improve the mean estimator performance. This occurs if and only if the corresponding error $z(t)$ is orthogonal to any admissible estimator variation $n(t)$, i.e. $\bar{E}Ez(t)n^*(t) = \bar{E}Ez(t)^*n(t) = 0$. Then, the perturbed criterion value reduces to

$$\begin{aligned} \bar{E}E\left|\frac{V}{U}u(t) - \hat{d}(t)\right|^2 &= \bar{E}E\left(\frac{V}{U}(u(t) - \hat{u}(t)) - n(t)\right)\left(\frac{V}{U}(u(t) - \hat{u}(t)) - n(t)\right)^* \\ &= \bar{E}E(\ |z(t)|^2 - z(t)n^*(t) - n(t)z^*(t) + |n(t)|^2\ ) \end{aligned}$$

$$= \quad \bar{E}E(\ |z(t)|^2 + |n(t)|^2\ ) \tag{B.2}$$

which is obviously minimized by $n(t) = 0$. Since all systems included in the extended design model (2.1)–(2.3) are assumed stable, both $z(t)$ and $n(t)$ will be stationary. Parsevals formula can then be used, to express $\bar{E}Ez(t)n^*(t)$ as

$$\bar{E}E\left\{\frac{V}{U}\left[\left(1-q^{m-k}\frac{QB}{RA}\right)\frac{C}{D}e(t)-q^m\frac{QM}{RN}v(t)\right]\right\}\left\{q^{m-k}\mathcal{M}\frac{BC}{AD}e(t)+q^m\mathcal{M}\frac{M}{N}v(t)\right\}^*$$

$$= \bar{E}\frac{1}{2\pi i}\oint_{|z|=1}\frac{V}{U}\left[\left(z^{k-m}-\frac{QB}{RA}\right)\frac{C}{D}\frac{B_*C_*}{A_*D_*}-\rho\frac{QMM_*}{RNN_*}\right]\mathcal{M}_*\frac{dz}{z}$$

$$= \frac{1}{2\pi i}\oint\frac{\bar{E}\{z^{k-m}VCC_*B_*N_*NAR - VQ(CC_*BB_*NN_* + \rho MM_*AA_*DD_*)\}}{UDNARN_*A_*D_*}\mathcal{M}_*\frac{dz}{z}\ . \tag{B.3}$$

The denominator has no zeros on the unit circle. We are allowed to move the expectation $\bar{E}$ inside the integration, since, for any particular realization of $\Delta C, \Delta B$ and ΔM, the integrand is Riemann integrable on the unit circle. See e.g. [11] Theorem 3.8. The expectation $\bar{E}$ operates on the numerator, since stochastic variables are present there only. Using the spectral factorization (4.1), and the fact that neither of V, N, A, R or Q contain stochastic variables, (B.3) can be expressed as

$$\frac{1}{2\pi i}\oint\frac{(z^{k-m}VNN_*AR\bar{E}(CC_*B_*) - VQr\beta\beta_*)}{UDNARN_*A_*D_*}\mathcal{M}_*\frac{dz}{z}\quad . \tag{B.4}$$

Now, $\bar{E}Ez(t)n^*(t) = 0$ is fullfilled if all poles in $|z| < 1$ of the integrand are cancelled by zeros. Since $\mathcal{M}$ and NAD are stable, $(1/N_*A_*D_*)\mathcal{M}_*$ will have poles only in $|z| > 1$. All other poles are in $|z| < 1$. Thus, we require

$$z^{k-m}VNN_*AR\bar{E}(CC_*B_*) - VQr\beta\beta_* = zL_*UDNAR$$

for some polynomial $L_*(z)$ or, equivalently,

$$(z^{k-m}VN_*\bar{E}(CC_*B_*) - zL_*UD)NAR = Qr\beta_*\beta V\quad . \tag{B.5}$$

The right–hand side of (B.5) must contain R as a factor. Since R must be stable, its factors cannot include factors of β_*. Thus, $\beta V = RH$ for some stable $H(z^{-1})$. Now, cancel R in (B.5). Observe that NA must be factor of QH, i.e. $QH = Q_1NA$. The filter $Q/R = (Q_1NA/H)/(\beta V/H)$ is (4.8). Cancel NA and exchange q for z in (B.5), to obtain

$$q^{k-m}VN_*\bar{E}(CC_*B_*) = r\beta_*Q_1 + qUDL_*\quad . \tag{B.6}$$

This is (4.9), since by using independence and $\bar{E}(\Delta B) = 0$,

$$\bar{E}(CC_*B_*) = \bar{E}(CC_*)\bar{E}(B_0A_1 + A_0B_1\Delta B)_* = \tilde{C}\tilde{C}_*B_{0*}A_{1*}\quad . \tag{B.7}$$

The "only if"–part of the result follows because choices of Q/R other than (4.8) correspond to $n(t) \neq 0$, which, according to (B.2), increase the criterion value.

Remark on the degrees (4.10). Diophantine equations in general have an infinite number of solutions. In (4.9), however, causality requires Q_1 to be a polynomial only in q^{-1}, while optimality requires L_* to be a polynomial in q. (If L_* were allowed to have negative powers of z as arguments, poles at the origin would be introduced in the integrand of (B.4). The path integral would then not vanish.)

The generic degrees (4.10) are then uniquely determined by the requirement that the highest occuring powers of q^{-1} and q, respectively, must be covered by the variables in (4.9). This gives an equal number of equations and unknowns. The linear system of equations is nonsingular, since β_* and UD have no common factors. A more general discussion of these points can be found in [3], Section IV of [2] and in [14] [15].

References

[1] A. Ahlén and M. Sternad, "Optimal deconvolution based on polynomial methods," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol 37, (1989) 217–226.

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

[3] A. Ahlén and M. Sternad, "Optimal filtering problems", in K. Hunt, ed. *Polynomial methods in optimal control and filtering.* Control Engineering Series, Peter Peregrinus, London, 1993.

[4] A. Ahlén and M. Sternad, "Derivation and design of Wiener filters using polynomial equations," in C. T. Leondes ed. *Advances in Control and Dynamic Systems: Digital Signal Processing and Applications*, Academic Press, New York, 1994, to appear.

[5] 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, (1976) 98–100.

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

[7] 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, (1992) 913–928.

[8] 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, (1989) 333–356.

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

[10] W. F. Haddad and D. S. Bernstein, "Robust, reduced–order, nonstrictly proper state estimation via the optimal projection equations with guaranteed cost bounds," *IEEE Transactions on Automatic Control*, 33, (1988) 591–595.

[11] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

[12] S. A. Kassam and T. L. Lim, "Robust Wiener filters," *Journal of The Franklin Institute*, (1977) 171–185.

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

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

[15] V. Kučera, *Analysis and Design of Discrete Linear Control Systems.* Academia, Prague and Prentice Hall International, London, 1991.

[16] D. G. Lainiotis, "Partitioning: a unifying framework for adaptive systems. Part I: estimation," *Proceedings of the IEEE*, vol 64, (1976) 1126–1143.

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

[18] C. J. Martin and M. Mintz, "Robust filtering and prediction for linear systems with uncertain dynamics: a game–theoretic approach," *IEEE Transactions on Automatic Control*, vol 28, (1983) 888–896.

[19] G. Moustakides and S. A. Kassam, "Robust Wiener filters for random signals in correlated noise," *IEEE Transactions on Information Theory*, vol 29, (1983) 614–619.

[20] G. Moustakides and S. A. Kassam, "Minimax equalization for random signals," *IEEE Transactions on Communications*, vol 33, (1985) 820–825.

[21] R. A. Padilla and A. H. Haddad, "On the estimation of uncertain signals using an estimation–detection scheme," *IEEE Transactions on Automatic Control*, vol 21, (1976) 509–512.

[22] H. V. Poor, "On robust Wiener filtering," *IEEE Transactions on Automatic Control*, vol 25, (1980) 531–536.

[23] J. L. Speyer and D. E. Gustafson, "An approximation method for estimation in linear systems with parameter uncertainty," *IEEE Transactions on Automatic Control*, vol 20, (1975) 354–359.

[24] M. Sternad and A. Ahlén, "The structure and design of realizable decision feedback equalizers for IIR channels with colored noise," *IEEE Transactions on Information Theory*, vol 36, (1990) 848–858.

[25] M. Sternad, A. Ahlén and E. Lindskog, "Robust decision feedback equalization," *ICASSP'93, International Conference on Acoustics, Speech and Signal Processing*, Minneapolis, MN, April 26-30, 1993.

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

[27] K. S. Vastola and H. V. Poor, "Robust Wiener–Kolmogorov theory," *IEEE Transactions on Information Theory*, vol 30, (1984) 316–327.

[28] S. D. Wang, T.S. Kuo and C. F. Hsu, "Optimal observer design for linear dynamical systems with uncertain parameters," *International Journal of Control*, vol 45, (1987) 701–711.

[29] K. Öhrn, A. Ahlén and M. Sternad, "A polynomial approach to multivariable robust filtering, prediction and smoothing." To appear at the 32nd IEEE Conference on Decision and Control, San Antonio, TX, December 15–17, 1993.