WIENER FILTER DESIGN USING POLYNOMIAL EQUATIONS

Anders Ahlén and Mikael Sternad *

October 1990

Abstract

A simplified way of deriving of realizable and explicit Wiener filters is presented. Discrete time problems are discussed, in a polynomial equation framework. Optimal filters, predictors and smoothers are calculated by means of spectral factorizations and linear polynomial equations. A new tool for obtaining these equations, for a given problem structure, is described. It is based on evaluation of orthogonality in the frequency domain, by means of cancelling stable poles with zeros. Comparisons are made to previously known derivation methodology such as "completing the squares" for the polynomial systems approach and the classical Wiener solution. The simplicity of the proposed derivation method is particularly evident in multisignal filtering problems. To illustrate, two examples are discussed: a filtering and a generalized deconvolution problem. A new solvability condition for linear polynomial equations appearing in scalar problems is also presented.

EDICS no. 4.2.2.

Keywords: Wiener filters; digital filter design; optimal estimation; smoothing; polynomial equations.

*Automatic Control and Systems Analysis Group, Department of Technology, Uppsala University, P. O. Box 534, S-751 21 Uppsala, Sweden. This work has been supported by the Swedish National Board for Technical Development, under contract 87-01573.

1 INTRODUCTION

Wiener filtering has been a classical tools in signal processing and communication since the 1950's. Despite of this, there are still ways to improve the estimator design technique. One such development is the theme of the present paper.

The concept of orthogonality will be utilized in a novel way, within the polynomial equations approach to linear filtering problems. The process of deriving estimator design equations, for a given problem structure, is then simplified significantly, compared to the techniques now in use.

The problem of interest is the optimization of realizable IIR–filters. They are used for prediction, filtering or smoothing of signals or signal vectors. The minimization of mean square error criteria by linear estimators is considered. Stochastic models of possibly complex–valued signals in discrete time are assumed known. Apart from realizability (internal stability and the use of finite smoothing lags), no restrictions are placed on the estimator structures and degrees. In the extensive literature on such problems, three basic methodologies can be distinguished.

1. In the classical Wiener filtering approach, variational arguments are utilized. Frequency functions are obtained, whose causal parts are sought. See, for example, [1]–[6]. These causal parts are then evaluated in a rather cumbersome way, using partial fraction expansions and residue calculus. While this is suitable for solving simple specific examples, an efficient general technique for determining the filters as rational transfer functions would be valuable.

2. The problem may be transformed to state–space form. A (stationary) Kalman filter can then be designed [33], [28]. In contrast to Wiener estimators, Kalman filters can be designed also when measurements are nonstationary. When only a few of the states need to be estimated from stationary data, Kalman filters are, however, unnecessarily complex. The relationship between Kalman and Wiener filters has been studied by e. g. Shaked [32].

3. Within the control systems field, the polynomial approach to linear quadratic optimization problems has been developed in a general way by Kučera, [9]–[14]. It provides a systematic way of evaluating the causal factor of the Wiener–Hopf solution. Transfer functions in the signal models are represented by polynomial fractions (or by polynomial matrix fractions in multivariable problems). Optimal filters are designed by solving spectral factorizations and linear polynomial equations. For a given problem structure, these equations are usually derived using a method of "completing the squares", [10]–[17].

The polynomial systems approach to the design of IIR–filters is well suited to many applications such as adaptive filtering and control. A drawback with the derivation technique based on "completing the squares" is that it often leads to rather long and tedious calculations. The same is true for an alternative approach, based on differentiation of the criterion [23], [24]. A new and simpler methodology is presented in this paper. It is based on the evaluation of orthogonality, to obtain the required polynomial equations. The technique can be utilized for solving estimation and control problems, in discrete time as well as in continuous time.

Control problems are discussed in [40]. In this paper, we will focus on discrete time estimation problems with stationary signals. In the following section, we present the technique in general. The design of filters, predictors and smoothers for scalar signals illustrates the approach in Section III. For comparison, the solutions derived by means of "completing the squares" and by using the conventional Wiener approach are discussed in Appendix A and B, respectively. It is pointed out that a linear polynomial equation provides a systematic way of calculating the causal part of the Wiener solution. In Section IV, a new solvability condition is presented for the type of linear polynomial equations that are utilized. Multi–channel filter design is illustrated in Section V by two examples: an estimation problem, previously discussed by Roberts and Newmann [15] and a more complicated generalized deconvolution problem. The latter result, for coloured signals and noises, is believed to be new. It generalizes the scalar estimator of Ahlén and Sternad [18], and the multivariable smoothers for white signals and noises of Deng [21] and Moir [22]. However, the different estimation problems discussed are not the main point of the paper. They have been included merely to clarify different aspects of the reasoning used in the derivation technique. A numerical example illustrating a filter calculation can be found in Section VI.

2 THE OPTIMIZATION PROCEDURE

The derivation technique is outlined in this section, using a minimum of notations. Specific notations are introduced in Sections III and V, when required.

Consider a linear discrete–time system, which is stable and time–invariant. It is driven by a vector of stationary white noises, with zero means

$$e(t) = (e_1(t) \dots e_n(t))^T \quad .$$

The system generates a stationary *measurement* vector sequence

$$y(t) = (y_1(t) \dots y_p(t))^T$$

and a vector sequence of stationary *desired responses*

$$f(t) = (f_1(t) \dots f_\ell(t))^T$$

see Figure 1.

The system is parametrized by transfer functions and ARMA–models, using the backward shift operator q^{-1}, where $q^{-1}v(t) \triangleq v(t-1)$. Signals and transfer function coefficients are allowed to be complex–valued. The superscript "$*$" denotes complex conjugate–transpose for signal vectors. Rational functions and matrices are denoted by symbols like $\mathcal{R}$.

Figure 1. The estimation problem, where $\hat{f}(t)$ is the estimate of $f(t)$, while $n(t)$ is a variational term.

Our aim is to optimize a linear estimator of $f(t)$

$$\hat{f}(t|t+m) = \mathcal{F}(q^{-1})y(t+m) \tag{2.1}$$

where $\mathcal{F}(q^{-1})$ is an $\ell|p$–matrix, having causal and stable transfer functions as elements. Depending on m, the estimator constitutes a predictor ($m<0$), a filter ($m=0$) or a fixed lag smoother ($m>0$). Denote the trace of a matrix P by tr$P$ and introduce the quadratic criterion

$$J = \mathrm{tr}E(\varepsilon(t)\varepsilon^*(t)) = \sum_{i=1}^{\ell} E|\varepsilon_i(t)|^2 \tag{2.2}$$

where

$$\varepsilon(t) = (\varepsilon_1(t)\ldots\varepsilon_\ell(t))^T \triangleq f(t) - \hat{f}(t|t+m) \quad . \tag{2.3}$$

The criterion (2.2) is to be minimized, under the constraint of realizability (stability and causality) of the filter $\mathcal{F}(q^{-1})$. Since all signals are assumed to be stationary and stability of $\mathcal{F}(q^{-1})$ is required, $\varepsilon(t)$ is stationary.

We will use variational arguments in order to minimize (2.2). For this purpose, introduce an *alternative estimator*

$$\hat{d}(t|t+m) = \mathcal{F}(q^{-1})y(t+m) + n(t) \tag{2.4}$$

where the stationary signal $n(t)$ represents a modification of the estimate (2.1). All admissible variations can be represented by $n(t) = \delta\mathcal{G}(q^{-1})y(t+m)$, where δ

is a scalar and $\mathcal{G}(q^{-1})$ is an arbitrary, but stable and causal, rational $\ell|p$–matrix. The use of (2.4) results in the criterion

$$\begin{aligned}\overline{J} &= \mathrm{tr}E\{f(t) - \hat{d}(t|t+m)\}\{f^*(t) - \hat{d}^*(t|t+m)\} \\ &= \mathrm{tr}\{E\varepsilon(t)\varepsilon^*(t) - E\varepsilon(t)n^*(t) - En(t)\varepsilon^*(t) + En(t)n^*(t)\} \ . \qquad (2.5)\end{aligned}$$

If the *mixed terms in (2.5) are zero*, then $n(t) \equiv 0$ evidently minimizes $\overline{J}$, since $\mathrm{tr}En(t)n^*(t) > 0$ for $n(t) \neq 0$. Then, the estimator (2.1) is optimal. Orthogonality between the error $\varepsilon(t)$ and any admissible linear function of the measurements, $n(t)$, guarantees optimality. (This well–known condition is also obtained by differentiating with respect to δ, and requiring $\partial\overline{J}/\ \partial\delta|_{\delta=0} = 0$.)

For symmetry reasons, it is sufficient to consider $E\varepsilon(t)n^*(t)$. Use Parseval's formula to convert the required orthogonality, $E\varepsilon(t)n^*(t) = 0$, into the frequency–domain relation

$$E\varepsilon(t)n^*(t) = \frac{1}{2\pi j}\oint_{|z|=1} \phi_{\varepsilon n^*}\frac{dz}{z} = 0 \ . \qquad (2.6)$$

The ij'th element of the $\ell|\ell$ cross–spectral density matrix $\phi_{\varepsilon n^*}$ can always be expressed as

$$\phi_{\varepsilon n^*} = \frac{T^{ij}(z,z^{-1})}{S_+^{ij}(z^{-1})S_-^{ij}(z)} \qquad (2.7)$$

where T^{ij}, S_+^{ij} and S_-^{ij} are polynomials. While S_+^{ij} has all zeros in $|z| < 1$, S_-^{ij} has all zeros in $|z| > 1$. The relation (2.6) is fulfilled if, in *every* element of the integrand, *all poles inside the integration path* $|z| = 1$ *are cancelled by zeros.* The use of this requirement to obtain orthogonality is the key idea of the derivation technique. Thus, we require that

$$\frac{T^{ij}(z,z^{-1})}{S_+^{ij}(z^{-1})}\frac{1}{z} = L^{ij}(z) \qquad i = 1\ldots\ell \ , \ j = 1\ldots\ell \qquad (2.8)$$

where $L^{ij}(z)$ are polynomials in z.

The relations (2.8) determine the estimator $\mathcal{F}(q^{-1})$. As will be seen in the sequel, (2.7) can be simplified using a *spectral factorization*, derived by expressing $y(t)$ in innovations form. Using the polynomial matrix fraction description, discussed in Section V, the relations (2.8) can be evaluated collectively, rather than individually, when $\ell > 1$. They then reduce to one *linear polynomial (matrix) equation.* Let us summarize the procedure.

1. Parametrize the system by rational transfer functions, described as polynomial (matrix) fractions. Define a polynomial spectral factorization from the spectral density of $y(t)$.

2. Define the estimation error $\varepsilon(t)$ and introduce a variation $n(t)$ of the estimate. Express $E\varepsilon(t)n^*(t)$ in the frequency domain by means of Parseval's formula and simplify it, using the spectral factorization.

3. Fulfill the orthogonality requirement $E\varepsilon(t)n^*(t) = 0$ by cancelling all poles in $|z| < 1$, in every element of the integrand, by zeros. This leads to a linear polynomial equation, which determines the estimator.

It can be shown that the introduction of a weighting matrix $\Psi > 0$ with *constant* elements into the criterion, $J = \text{tr}E(\varepsilon(t)\Psi\varepsilon^*(t))$, does not affect the optimal solution in any way. The use of frequency shaped weighting is more interesting. We could, for example, use $J = \text{tr}E(\varepsilon_w(t)\varepsilon_w^*(t))$ where $\varepsilon_w(t) \triangleq \mathcal{H}(q^{-1})\varepsilon(t)$, with $\mathcal{H}(q^{-1})$ being a diagonal matrix with stable transfer functions as diagonal elements. Weighted signals $\varepsilon_w(t)$ and $n_w(t)$ are then substituted for $\varepsilon(t)$ and $n(t)$ in (2.5).

The reasoning above describes a systematic *and constructive* derivation technique. In a previously known methodology, (2.4) and (2.6) were sometimes used to demonstrate optimality by contradiction, for estimators or regulators derived by other means. Examples of such non–constructive proofs can be found in [27] and in [18],[19],[20],[41].

3 ESTIMATION OF A SCALAR SIGNAL

The methodology introduced above is best clarified by applying it to a simple scalar example. For comparison, the "completing the squares" approach is discussed in Appendix A. An extension of the conventional Wiener formulation, using rational transfer functions, is presented in Appendix B. First, some notations are introduced. For any polynomial

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

define the *conjugate polynomial*

$$P_*(q) = p_o^* + p_1^* q + \ldots + p_{np}^* q^{np}$$

where q is the forward shift operator ($qv(t) \triangleq v(t+1)$) and p_j^* is the conjugate of the (possibly complex) coefficient p_j. In the frequency domain, the complex variable z is substituted for q. Polynomials with positive powers of q or z as arguments are always denoted by a "$*$"–subscript. For convenience, the polynomial arguments are often omitted. We call $P(q^{-1})$ *stable* (or strictly Schur) if all zeros of $P(z^{-1})$ are in $|z| < 1$. Note that whenever P is stable, all zeros of P_* are in $|z| > 1$.

Now, assume that the signal $s(t) = [C(q^{-1})/D(q^{-1})]e(t)$ is to be estimated from noisy measurements

$$y(t) = s(t) + \frac{M(q^{-1})}{N(q^{-1})}v(t) \tag{3.1}$$

up to time $t + m$, using a stable and causal estimator $\mathcal{F}(q^{-1}) = Q(q^{-1})/R(q^{-1})$. See Figure 2.

Figure 2. The scalar output filtering, prediction or smoothing problem. The signal $\{f(t)\} = \{s(t)\}$ is to be estimated from $\{y(t+m)\}$.

Here, $e(t)$ and $v(t)$ are mutually independent and white stationary sequences. They have zero means and variances $\lambda_e > 0$ and $\lambda_v \geq 0$, respectively. The ARMA–models C/D and M/N are stable, causal and have no common zeros on the unit circle. All model polynomials, with degree nc, nd, etc, are monic. The measurements $\{y(t)\}$ can also be described by the innovations model

$$y(t) = \frac{\beta(q^{-1})}{D(q^{-1})N(q^{-1})}(\sqrt{\lambda_\eta}\eta(t)) \tag{3.2}$$

where the innovations sequence $\sqrt{\lambda_\eta}\eta(t)$ has variance λ_η. The monic and stable polynomial $\beta(q^{-1}) = 1 + \beta_1 q^{-1} + \ldots + \beta_{n\beta} q^{-n\beta}$ is the (polynomial) *spectral factor*. Let us optimize $\mathcal{F} = Q/R$, following the procedure introduced in Section II.

1. Set the spectral densities $\Phi_y(e^{j\omega})$ equal in (3.1) and (3.2). This gives the spectral factorization equation

$$r\beta\beta_* = CC_*NN_* + \rho MM_*DD_* \tag{3.3}$$

   where $r = \lambda_\eta/\lambda_e$, $\rho \triangleq \lambda_v/\lambda_e$ and $\beta(z^{-1})$ is stable.

2. Use the error signal

$$\varepsilon(t) = \left(1 - q^m \frac{Q}{R}\right) \frac{C}{D} e(t) - q^m \frac{Q}{R}\frac{M}{N} v(t)$$

   and the estimator variation $n(t) = \mathcal{G}(q^{-1})y(t+m)$, where $\mathcal{G}(q^{-1})$ is *any* stable and causal transfer function. The first mixed term in (2.5) is then

$$E\varepsilon(t)n^*(t) =$$

$$= E\left(\frac{(R - q^m Q)C}{RD} e(t)\right)\left(\mathcal{G}q^m \frac{C}{D} e(t)\right)^* - E\left(q^m \frac{QM}{RN} v(t)\right)\left(\mathcal{G}q^m \frac{M}{N} v(t)\right)^*$$

$$= \frac{\lambda_e}{2\pi j} \oint \frac{(R - z^m Q) z^{-m} CC_*NN_* - \rho QMM_*DD_*}{RDD_*NN_*} \mathcal{G}_* \frac{dz}{z}$$

$$= \frac{\lambda_e}{2\pi j} \oint \frac{(z^{-m} RCC_*NN_* - Qr\beta\beta_*)}{RDD_*NN_*} \mathcal{G}_* \frac{dz}{z} \tag{3.4}$$

where Parseval's formula (see Appendix C) and (3.3) were utilized.

3. The stable polynomials R, D and N have zeros in $|z| < 1$, while the poles of $\mathcal{G}_*$ and the zeros of $D_* N_*$ are in $|z| > 1$. Thus, all poles inside $|z| = 1$ of the integrand of (3.4) are eliminated if (and only if)

$$z^{-m} R C C_* N N_* - Q r \beta \beta_* = z R D N L_*$$

for some polynomial $L_*(z)$, cf. (2.8). Now, N must be a factor of $Qr\beta\beta_*$. Set $Q = Q_1 N$. [1] Cancel N and substitute q for z:

$$R(q^{-m} C C_* N_* - q D L_*) = Q_1 r \beta \beta_* \quad . \tag{3.5}$$

Evidently, R must be a factor of $Q_1 r \beta \beta_*$. Since β_* is unstable, while Q_1 is part of the estimator numerator, set $R = \beta$. Now $Q_1(q^{-1})$, together with $L_*(q)$, can be found as the *unique* solution to the linear polynomial equation

$$q^{-m} C C_* N_* = r \beta_* Q_1 + q D L_* \quad . \tag{3.6}$$

The solvability and solution of equations like (3.6) is discussed in the next section. With $Q = Q_1 N$ and $R = \beta$, the optimal estimator is

$$\hat{s}(t|t+m) = \frac{Q_1(q^{-1}) N(q^{-1})}{\beta(q^{-1})} y(t+m) \quad . \tag{3.7}$$

The estimator is obtained by solving (3.3) for β (and r) and (3.6) for Q_1 (and L_*). If the innovations model (3.2) is known in addition to (3.1), there is no need to solve (3.3). The IIR-filter is internally stable, since β is stable. It may contain stable common factors. For a detailed solved example of an equation similar to (3.6), see [18], Example 1.

Note that the arbitrary rational function $\mathcal{G}$ does not affect the result in any way. We could have derived (3.6) with $n(t) = y(t+m)$, i.e. by just requiring orthogonality to the *last* measurement. This is the case in general, if the measurements are stationary.

For a derivation of (3.7) by the "completing the squares" method, see Appendix A. That derivation requires significantly more calculations.

Comparison with the Wiener solution, reformulated as in Appendix B, provides the following insights:

- The solution of the spectral factorization equation (3.3) corresponds to the design of a whitening filter (the inverse of the innovations model (3.2)).

[1] The polynomial β_* is unstable, and N and β have no common factors, in general. If they have, $Q\beta = Q_1 \beta_1 N$ must hold. This implies $R = \beta_1$. Eq. (3.6) remains unaffected, and the filter $Q/R = Q_1 \beta_1 N / \beta \beta_1$ is still given by (3.7).

- The linear equation (3.6) represents a calculation of the causal part $\{\cdot\}_+$ of the Wiener filter. (Readers who are used to, and prefer, the classical Wiener formulation could utilize this relation, by deriving their filters in the usual way and then evaluate the causal bracket by solving a polynomial equation.)

The optimality requirement (2.8) determines the *structure* and *degree* of the estimator. The methodology cannot be utilized for optimizing filters with a prespecified restricted complexity and degree. In such problems, all poles inside $|z| = 1$ of $\phi_{\varepsilon n^*}$ cannot be eliminated. Instead, the orthogonality (2.6) is fulfilled at an optimum because the residues corresponding to all poles in $|z| \leq 1$ cancel. While a well–known closed–form expression exists for optimal FIR–filters [8], no corresponding expression exists for IIR–filters of fixed degree.

The scalar variant of the derivation technique may be applied when the desired response is scalar ($\ell = 1$) but the measurements are multiple ($p > 1$), if the number of signal and noise sources n equals p. Optimization of decision feedback equalizers [19] is such a case. Multiple scalar variations, $n_i(t) = \mathcal{G}_i y_i(t + m)$, $i = 1, \ldots, p$, are then utilized in (2.4) and orthogonality with respect to each, $E\varepsilon(t)n_i^*(t) = 0$, $i = 1, \ldots, p$, is required. When $n > p > 1$, multivariable spectral factorizations become an integral part of the solution. Such problems can be handled by the multivariable polynomial formalism, to be discussed in Section V.

4 REMARKS ON THE SOLVABILITY

With stable transfer functions without common zeros on $|z| = 1$, the right hand side of the polynomial spectral factorization (3.3) is positive on $|z| = 1$. A stable $\beta(q^{-1})$ and a scale factor r, which satisfy this equation, thus exist. Efficient algorithms for polynomial spectral factorization can be found in [10] and [37].

A *Diophantine equation* like (3.6) can easily be written as a linear systems of equations, $\mathbf{AX} = \mathbf{B}$, where $\mathbf{A}$ is a Sylvester matrix containing polynomial coefficients [18]. It can be solved directly. An alternative, more computationally efficient solution procedure is based on the Euclidean algorithm [36]. The most critical numerical property is the occurence of almost common factors of A and B, which are not factors of C, in a polynomial equation $AX + BY = C$.

While Diophantine equations in general have an infinite number of solutions, equations arising from linear quadratic design problems mostly have one unique solution. This is a consequence of two requirements.

1. Filter causality requires Q_1 to be a polynomial only in q^{-1}.

2. Optimality restricts L_* to be a polynomial only in q. If powers of q^{-1} were allowed in L_* in (3.6), the integrand of (3.4) would have poles at the origin, resulting in a non–vanishing integral.

For polynomial equations with these properties, the following result can be established.

Lemma 1. The linear polynomial equation

$$A(q,q^{-1})X(q^{-1}) + B(q,q^{-1})Y(q) = C(q,q^{-1}) \tag{4.1}$$

where

$$\begin{aligned} A(q,q^{-1}) &\triangleq a_{na1}q^{na1} + \ldots + a_{-na2}q^{-na2} \neq 0 \\ B(q,q^{-1}) &\triangleq b_{nb1}q^{nb1} + \ldots + b_{-nb2}q^{-nb2} \neq 0 \\ C(q,q^{-1}) &\triangleq c_{nc1}q^{nc1} + \ldots + c_{-nc2}q^{-nc2} \end{aligned}$$

has a *unique* solution $X(q^{-1}) = x_o + x_1q^{-1} + \ldots + x_{nx}q^{-nx}$, $Y(q) = y_o + y_1q + \ldots + y_{ny}q^{ny}$ if and only if common factors of A and B are also factors of C and

$$nb1 + na2 - d = 1 \tag{4.2}$$

where d is the number of linearly dependent equations in the corresponding system of linear equations $\mathbf{AX} = \mathbf{B}$. □

Proof. Let g_1 and g_2 denote the highest powers of q and q^{-1} respectively, present anywhere in (4.1). The degree of $Y(q)$ must be selected such that the highest power of q in $B(q,q^{-1})Y(q)$ equals the highest power of q in any of the two other terms. (To increase ny above this value would be useless; with no matching terms, the superfluous coefficients would become zero.) Thus,

$$g_1 = nb1 + ny = \max\{na1, nc1\} \quad . \tag{4.3}$$

For similar reasons, the degree of $X(q^{-1})$, nx, must satisfy

$$g_2 = na2 + nx = \max\{nb2, nc2\} \quad . \tag{4.4}$$

Equation (4.1) corresponds to g_1+g_2+1 linear simultaneous equations. A solution exists only if factors of A and B are also factors of C. (Write $A = TA_1$ and $B = TB_1$, where T is the greatest common factor of A and B in (4.1). Since T is a factor of the left–hand side of (4.1), $T(A_1X + B_1Y)$, it must also be a factor of C.) A unique solution then exists if and only if the number of *linearly independent* equations equals the number of unknowns (coefficients of X and Y):

$$g_1 + g_2 + 1 - d = nx + 1 + ny + 1 \quad . \tag{4.5}$$

The use of (4.3) and (4.4) in (4.5) gives (4.2). If the left–hand side of (4.2) is < 1, the solution is non–unique. If it is > 1, no solution exists. □

Ježek [42] has studied the special case $nb1 = na2 = 0$ of equation (4.1). Note that, regardless of the solvability, we are forced to choose the degrees nx, ny according to (4.4),(4.3). *The structure of the equation determines the degrees uniquely.* Non–solvability of (4.1) could occur for two different reasons:

1. Existence of common factors of A and B, which are *not* factors of C. This would lead to algebraic equations which contradict each other.

2. The number of unknowns might be insufficient. High $nb1$ or $na2$ in (4.3), (4.4) lead to correspondingly small ny, nx.

A reduction of the number of linearly independent equations, i.e. $d > 0$, may occur because of common factors of A and B, of degree k, which *are* factors of C. They reduce the number of linearly independent equations by k. (Factors $q^{\pm k}$ do not count; multiplication of (4.1) by $q^{\pm k}$ leaves the linear simultaneous equations unaffected.) Another possibility is that some equations may be zero. [2]

When the number of equations equals the number of unknowns ($g_1 + g_2 + 1 = nx + ny + 2$) and no common polynomial factors occur, the Sylvester matrix $\mathbf{A}$ has full rank [31], which implies $d = 0$.

This is precisely the situation in a typical estimation problems, in particular in the estimator design equation (3.6). With degrees derived from (4.3) and (4.4), the number of unknowns and number of equations will coincide. Furthermore, D (with zeros in $|z| < 1$, since D is stable) and β_* (with zeros in $|z| > 1$, since β is stable) cannot have common factors. Thus, $d = 0$. With $nb1 = 1$ because of the free q–factor and $na2 = 0$, the condition (4.2) is satisfied. Consequently, a unique solution to (3.6) always exists[3]. The solution has polynomial degrees, obtained from (4.4) and (4.3)

$$
\begin{aligned}
nQ_1 &= \max(nc + m, nd - 1) \\
nL &= \max(nc + nn - m, n\beta) - 1 \quad .
\end{aligned}
\tag{4.6}
$$

The highest–degree coefficients of Q_1 and/or L_* may become zero when specific problems are solved. The degrees are then less than (4.6). Note that in smoothers ($m > 0$), the required degree of Q_1 grows with the smoothing lag m.

The choice $R = \beta$ in (3.5) is unique: any other choice, $\beta = \beta_2\beta_3$, $R = \beta_2$, $\deg \beta_3 > 0$, would result in a polynomial equation $q^{-m}CC_*N_* = Q_1 r \beta_3 \beta_* + qDL_*$, which would be unsolvable in general. (In (4.2), $nb1 = 1$ and $na2 = \deg \beta_3$.)

[2] Consider $(2q + q^{-k})X(q^{-1}) + (-q - q^{-k})Y(q) = q$. No common factors, except q^j, occur. From (4.3) and (4.4), $nx = ny = 0$. The algebraic equations consist of $g_1 + g_2 + 1 = k + 2$ equations, with only two unknowns. Of the equations, k are identically zero. Thus, $d = k$. With $nb1 = 1$ and $na2 = k$, the requirement (4.2) is fulfilled. A unique solution ($X = 1, Y = 1$) exists. (This example was suggested by V. Kučera.)

[3] We prefer to use polynomial equations with both q and q^{-1} as arguments, since they are related to the optimization in a direct way. Equation (3.6) could be transformed into an equation with argument q^{-1} only, by multiplication by q^{-nL-1}. The polynomial degrees (4.6) then correspond to the *least degree solution with respect to* $\overline{L}(q^{-1}) \triangleq q^{-nL}L_*(q)$ of such an equation. (It can be shown that the reason for this is that the free q–factor is positioned in the term qDL_*. If the relation (4.2) had instead been satisfied by a free q^{-1}–factor in $\beta_* Q_1$, the solution would have corresponded to the least degree solution with respect to $Q_1(q^{-1})$.)

In contrast to the classical Wiener formulation, approaches based on polynomial equations can be used when signal and/or noise–generating processes are unstable. This includes, for example, models of random walk signals (ARIMA–processes) and models of deterministic signals and disturbances such as sinusoids. In the derivation technique, the stationarity of $n(t)$ and $\varepsilon(t)$ must then be ascertained. See [44] for a detailed example.

Solvability problems would occur only in very unrealistic situations, of little practical interest: when D contains *strictly unstable* factors (with zeros in $|z| > 1$), which are *also* factors of $C_* N_*$, and consequently also of β_*. With common factors, the solution of (3.6), with degrees (4.6), becomes non–unique. The requirement that $\varepsilon(t)$ must be stationary defines a second Diophantine equation. In these rare cases, this equation would have to be used in combination with (3.6) to determine the filter uniquely. See [13] and [43].

5 MULTIVARIABLE ESTIMATION

In this section, multivariable systems will be described by means of fractions of polynomial matrices (MFD's). The methodology presented in Section II is exemplified by two estimation problems. In Subsection B, a prediction, filtering or smoothing problem is discussed. The signal and noise models are expressed in "common denominator"–form. This problem is considered for pedagogical reasons, because of its simplicity. The filter case ($m = 0$) has been derived previously e.g. by Roberts and Newmann [15], using the method of "completing the squares". In Subsection C, a more difficult problem is solved: the derivation of design equations for a general deconvolution estimator. This is a multivariable generalization of a problem discussed in [18].

Some definitions and some concepts from the theory of multivariable linear systems will be needed. We begin by introducing these prerequisites below.

A. Preliminaries

A *polynomial matrix* $P(q^{-1})$ is a matrix with all elements being polynomials in the backward shift operator. Alternatively, it can be expressed as a matrix polynomial

$$P(q^{-1}) \triangleq P_o + P_1 q^{-1} + \ldots + P_{np} q^{-np}$$

where P_j are constant matrices. Let $P_*(q)$ denote the complex conjugate transpose of $P(q^{-1})$. The i,j–th polynomial element of P_* is then simply the conjugate (defined in Section III) of the j,i–th element of P. With the *degree* of $P(q^{-1})$, denoted $\deg P$ or np, we mean the highest degree occurring in any element of P. (In other words, $P_{np} \neq 0$). With $rank P(z^{-1}) = r$, we mean the normal rank; $\mathrm{rank} P(z^{-1}) = r$ for almost all z. If $P(q^{-1})$ is square and has full rank, it is nonsingular, and the inverse $P(q^{-1})^{-1}$ exists. In general, the inverse will be a *rational* matrix. All elements of $P(q^{-1})^{-1}$ are *causal* if and only if P_o is nonsingular. $P(q^{-1})^{-1}$ is then said to be *causal* or *proper*.

A square polynomial matrix $P(z^{-1})$ is called *stable* (or strictly Schur) if its determinant polynomial, denoted $\det P(z^{-1})$, has all zeros in $|z| < 1$. A rational matrix $\mathcal{R}(z^{-1})$ is stable if all its elements have stable denominator polynomials.[4] Note that if P is stable, elements of $P_*^{-1} = \mathrm{Adj} P_* / \det P_*$ have poles only in $|z| > 1$.

Two polynomial matrices are said to be left (right) *coprime* if every common left (right) divisor is a *unimodular* matrix U ($\det U$=constant). Any rational matrix $\mathcal{R}(q^{-1})$, of dimension $p|r$, can be represented as a *matrix fraction description* (MFD), either left or right: $\mathcal{R}(q^{-1}) = A_1^{-1}(q^{-1})B_1(q^{-1}) = B_2(q^{-1})A_2^{-1}(q^{-1})$. The polynomial matrices (A_1, B_1) and (B_2, A_2) can be chosen left and right coprime, respectively. The MFD's are then called *irreducible*. If A_1 is square and stable, all elements of the rational matrix $\mathcal{R} = A_1^{-1}B_1 = (\mathrm{Adj} A_1)B_1 / \det A_1$ will, of course, be stable. For a more extensive discussion of MFD's see, for example, [30], [31].

B. Estimation of signals in coloured noise

Assume a signal $\{s(t)\}$ and noisy measurement $\{y(t)\}$, both with p elements, to be stationary stochastic vector–sequences described by

$$\begin{aligned} s(t) &= A^{-1}Be(t) \\ y(t) &= s(t) + A^{-1}Cv(t) \ . \end{aligned} \tag{5.1}$$

Here, (A, B, C) are polynomial matrices in the backward shift–operator q^{-1}, of dimensions $p|p$, $p|k$ and $p|r$, respectively. The matrix A is a common left denominator (not necessarily the least one) of the signal and noise models; A and $[B \ \ C]$ need not be left coprime. The white noises $\{e(t)\}$ and $\{v(t)\}$ are zero mean and mutually independent vector–sequences. They have covariance matrices $\phi \geq 0$ and

[4] The concepts of poles and zeros of rational matrices, defined via the Smith–McMillan form, will not be utilized.

$\psi \geq 0$ of dimensions $k|k$ and $r|r$, respectively. Let the matrices (A, B, C, ϕ, ψ) be known. Given data up to time $t + m$, we seek the optimal estimator

$$\hat{s}(t|t+m) = \mathcal{F}(q^{-1})y(t+m) \tag{5.2}$$

of the signal $s(t)$, such that the criterion (2.2) is minimized. See Figure 3. Note that we include singular filtering problems, which are difficult to handle with Kalman filtering algorithms. The noise covariance matrix ψ need not be strictly positive definite. We make the following assumptions.

Assumption 1. The polynomial matrix $A(q^{-1})$ is stable, with A_o nonsingular.

Assumption 2. The spectral density matrix, $\Phi_y(e^{j\omega})$, is nonsingular for all ω. [5]

The spectral density matrix is given by $\Phi_y(e^{j\omega}) = A^{-1}(B\phi B_* + C\psi C_*)A_*^{-1}$.

Following the scheme in Section II, we define the left spectral factorization

$$DD_* = B\phi B_* + C\psi C_* \quad . \tag{5.3}$$

A $p|p$ polynomial spectral factor D, with $\det D(z^{-1}) \neq 0$ in $|z| \geq 1$ and D_o nonsingular[6], can always be found under Assumption 2, [28], [10], [11], [39]. With a D such that D^{-1} is stable and causal, we continue with the orthogonality requirement (2.6).

Figure 3. A multivariable estimation problem. The vector sequence $\{f(t)\} = \{s(t)\}$ is to be estimated from measurements $\{y(t)\}$, up to time $t + m$.

[5] Two conditions on the polynomial matrices appearing in (5.1) are, together, sufficient for this. (1): The matrix $[B\phi \quad C\psi]$ has full (normal) row rank p and (2): The greatest common left divisor of $B\phi$ and $C\psi$ has nonzero determinant on $|z| = 1$. While (1) is a condition for existence of a spectral factor, (2) provides a spectral factor D such that $\det D \neq 0$ on $|z| = 1$.

[6] The spectral factor D is unique, up to a right orthogonal matrix. (If $VV_* = I$, $DD_* = (DV)(V_*D_*)$.) It is important to choose a (stable) spectral factor D such that D_o is nonsingular. Otherwise, the resulting filter $\mathcal{F}$ would be noncausal. Algorithms providing this exists [38], [39].

Let $\varepsilon(t) = s(t) - \hat{s}(t|t+m)$ and $n(t) = \mathcal{G}(q^{-1})y(t+m)$, where $\mathcal{G}(q^{-1})$ is an arbitrary, stable, causal and rational $p|p$–matrix. The first mixed term in (2.5) becomes

$$E\varepsilon(t)n^*(t) =$$

$$\begin{aligned}
&= E[(I - q^m\mathcal{F})A^{-1}Be(t) - q^m\mathcal{F}A^{-1}Cv(t)][\mathcal{G}q^m(A^{-1}Be(t) + A^{-1}Cv(t))]^* \\
&= \frac{1}{2\pi j}\oint[(I - z^m\mathcal{F})A^{-1}B\phi B_*A_*^{-1}z^{-m} - \mathcal{F}A^{-1}C\psi C_*A_*^{-1}]\mathcal{G}_*\frac{dz}{z} \\
&= \frac{1}{2\pi j}\oint[z^{-m}A^{-1}B\phi B_* - \mathcal{F}A^{-1}DD_*]A_*^{-1}\mathcal{G}_*\frac{dz}{z} \quad . \qquad (5.4)
\end{aligned}$$

In the last equality, the spectral factorization (5.3) was inserted. In order to fulfil the orthogonality requirement (2.6), $\mathcal{F}$ should be chosen such that, in *every* element of the integrand in (5.4), all poles inside the unit circle are cancelled by zeros. The elements of the rational matrix A^{-1} have poles only inside the unit circle, since A is stable. Elements of D contribute with poles in the origin, since they are polynomials in z^{-1}. The factor $A^{-1}D$ in the second term of (5.4) can be eliminated directly by $\mathcal{F}$, while A^{-1} in the first term has to be factored out to the left, to be cancelled later. Thus, set

$$\mathcal{F} = A^{-1}YD^{-1}A \qquad (5.5)$$

where $Y(z^{-1})$, of dimension $p|p$, is undetermined. Insert (5.5) into (5.4), to obtain

$$E\varepsilon(t)n^*(t) = \frac{1}{2\pi j}\oint A^{-1}(z^{-m}B\phi B_* - YD_*)A_*^{-1}\mathcal{G}_*\frac{dz}{z} \quad . \qquad (5.6)$$

Since the polynomial matrix A is stable and the rational matrix $\mathcal{G}$ is causal and stable, all elements of $A_*^{-1}\mathcal{G}_*$ have poles in $|z| > 1$. The requirements (2.8) will be fulfilled, collectively, if (and only if) there exists a polynomial matrix $L_*(z)$, such that

$$A^{-1}(z^{-m}B\phi B_* - YD_*) = zL_* \quad .$$

Rearranging and exchanging q for z, we obtain the equation

$$YD_* + qAL_* = q^{-m}B\phi B_* \quad . \qquad (5.7)$$

The unknowns $Y(q^{-1})$ and $L_*(q)$ have degrees

$$nY \le \max(na-1, nb+m) \quad ; \quad nL \le \max(nd, nb-m) - 1 \quad . \qquad (5.8)$$

In the estimator (5.5), we recognize the whitening filter $D^{-1}A$ of the Wiener solution. Since A and D are stable, all elements of $\mathcal{F}$ will be stable transfer functions. Since A_o and D_o are nonsingular, $\mathcal{F}$ will be causal.

The derivation above is simpler than the one for the filter ($m = 0$) presented by Roberts and Newmann in [15], but the result is the same. To compare the results, introduce the number

$$g \triangleq \max(nd, nb - m) \tag{5.9}$$

and denote $\overline{D} \triangleq q^{-g} D_*$, $\overline{B} \triangleq q^{-m-g} B_*$ and $Z \triangleq q^{-g+1} L_*$. Multiply (5.7) by q^{-g} to obtain

$$Y\overline{D} + AZ = B\phi\overline{B} \tag{5.10}$$

which, with $m = 0$, is equivalent to (27) in [15]. Hence, (5.5) is (26) in [15].

Equation (5.10) is a *bilateral* polynomial matrix equation. (The unknowns Y and Z in (5.10) appear on opposite sides of the two lefthand–side terms.) Note that A and D are stable. Thus, $\det A$ and $\det \overline{D}$ have no common factors. This implies that the invariant polynomials of A are coprime with all those of $\overline{D}$ and a solution *always exists*. (See Lemma 1 of [15].) In particular, there exists a least degree solution with respect to Z, with $\deg Z = \deg \overline{D} - 1 = g - 1$. This corresponds to a solution $(Y^o(q^{-1}), L_*^o(q))$ of (5.7) with degrees (5.8). Every solution to (5.7) can be expressed as $(Y, L_*) = (Y^o - qAX, L_*^o + XD_*)$, where the polynomial matrix X is undetermined, cf [10]. Since Y is required to be a polynomial matrix in q^{-1}, while L_* is required to be a polynomial matrix in q, $X = 0$ is the only choice. We conclude that the solution to (5.7) is *unique*.

C. A generalized deconvolution problem

We will now consider a more complicated problem. In many areas, it is of interest to estimate the input to a linear system, or a filtered version of it. See, for example, [18], [21], [22], [34], [35] and the references therein. Let the noise–corrupted measurement $y(t)$ and the input $u(t)$ be described by

$$\begin{aligned} y(t) &= A^{-1}Bu(t) + N^{-1}Mv(t) \\ u(t) &= D^{-1}Ce(t) \ . \end{aligned} \tag{5.11}$$

Here, (A, B, N, M, D, C) are polynomial matrices in the backward shift operator q^{-1}, of dimensions $p|p$, $p|s$, $p|p$, $p|r$, $s|s$ and $s|k$, respectively. The matrix B need not be stably invertible. It may not even be square. The noises $\{e(t)\}$ and $\{v(t)\}$ are stationary. They have zero means and covariance matrices $\phi \geq 0$ and $\psi \geq 0$, of dimensions $k|k$ and $r|r$, respectively. From data $y(t)$ up to time $t+m$, an estimator

$$\hat{f}(t|t+m) = \mathcal{F}(q^{-1})y(t+m) \tag{5.12}$$

of a filtered version of the input $u(t)$

$$f(t) = T^{-1}Su(t)$$

is sought. See Figure 4. The quadratic estimation error (2.2) is to be minimized. In this generalized deconvolution problem, the filter $T^{-1}S$, with T and S of dimensions $\ell|\ell$ and $\ell|s$, represents additional dynamics in the problem description (cf [41],[44]), or a frequency shaping weighting filter (cf [18]).

Assumption 1: The polynomial matrices $A(q^{-1})$, $N(q^{-1})$, $D(q^{-1})$ and $T(q^{-1})$ are all stable, with nonsingular leading matrices A_o, N_o, D_o and T_o.

Assumption 2: The spectral density matrix, $\Phi_y(e^{j\omega})$, is nonsingular for all ω.

.

Figure 4. The generalized deconvolution problem. The vector sequence $\{f(t)\}$ is to be estimated from the measurements $\{y(t)\}$, up to time $t+m$.

Compared to Subsection V.B, it is here slightly more difficult to express the spectral density matrix, Φ_y, using a *polynomial* matrix spectral factorization.[7] *Coprime factorizations*, which represent a kind of commutation operation for MFD's, have to be introduced. Define the following coprime factorizations

$$
\begin{aligned}
\tilde{D}^{-1}\tilde{B} &= BD^{-1} \\
\tilde{N}^{-1}\tilde{P} &= \tilde{D}AN^{-1}
\end{aligned}
\tag{5.13}
$$

with polynomial matrices $\tilde{D}$, $\tilde{N}$ and $\tilde{P}$ of dimension $p|p$, while $\tilde{B}$ has dimension $p|s$. The factorizations constitute the calculation of irreducible left MFD's from right MFD's. Thus, no unstable common factors are introduced. Since D and N are stable, $\tilde{D}$ and $\tilde{N}$ will be stable. Using (5.13), inverse matrices in the expression for Φ_y can be factored out to the left and right, leaving a polynomial matrix in the middle. We obtain

$$
\Phi_y = A^{-1}BD^{-1}C\phi C_* D_*^{-1}B_* A_*^{-1} + N^{-1}M\psi M_* N_*^{-1} = \alpha^{-1}\beta\beta_*\alpha_*^{-1} \tag{5.14}
$$

where

$$
\beta\beta_* = H\phi H_* + \tilde{P}M\psi M_*\tilde{P}_* \tag{5.15}
$$

and

$$
\alpha \triangleq \tilde{N}\tilde{D}A \quad ; \quad H \triangleq \tilde{N}\tilde{B}C \quad .
$$

From Assumption 2, a stable $p|p$ spectral factor β, with $\det \beta(z^{-1}) \neq 0$ in $|z| \geq 1$ and β_0 nonsingular, can always be found.

Now, the optimal estimate can be derived as in the previous subsection. Let $\varepsilon(t) = u(t) - \hat{u}(t|t+m)$ and $n(t) = \mathcal{G}(q^{-1})y(t+m)$. We obtain,

$$
E\varepsilon(t)n^*(t) =
$$

[7] It is preferable to avoid the numerically difficult task of performing spectral factorization of rational matrices, and instead use factorization of polynomial matrices. For this, there exist efficient numerical algorithms [39].

$$= E[(T^{-1}S - q^m \mathcal{F} A^{-1}B)D^{-1}Ce(t) - q^m \mathcal{F} N^{-1}Mv(t)][\mathcal{G} q^m (A^{-1}BD^{-1}Ce(t) + N^{-1}Mv(t))]^*$$

$$= \frac{1}{2\pi j} \oint \{ z^{-m} T^{-1} S D^{-1} C \phi C_* D_*^{-1} B_* A_*^{-1}$$

$$- \mathcal{F}[A^{-1}BD^{-1}C\phi C_* D_*^{-1} B_* A_*^{-1} + N^{-1} M \psi M_* N_*^{-1}]\} \mathcal{G}_* \frac{dz}{z} \quad . \tag{5.16}$$

The use of (5.13) and (5.14) in (5.16) gives

$$E\varepsilon(t)n^*(t) = \frac{1}{2\pi j} \oint \{ z^{-m} T^{-1} S D^{-1} C \phi C_* \tilde{B}_* \tilde{N} - \mathcal{F} \alpha^{-1} \beta \beta_* \} \alpha_*^{-1} \mathcal{G}_* \frac{dz}{z}. \tag{5.17}$$

Since A, $\tilde{D}$ and $\tilde{N}$ are stable, all elements of $\alpha^{-1} = A^{-1}\tilde{D}^{-1}\tilde{N}^{-1}$ have poles only in $|z| < 1$. Elements of β contribute poles at the origin, since they are polynomials in z^{-1}. These factors can be cancelled directly by $\mathcal{F}$. Moreover, introduce the additional coprime factorization

$$\hat{D}^{-1}\hat{S} = SD^{-1} \tag{5.18}$$

with a stable $\hat{D}$ of dimension $\ell|\ell$ and $\hat{S}$ of dimension $\ell|s$. Apply it in the integrand of (5.17). If $\mathcal{F}$ contains $T^{-1}\hat{D}^{-1}$ as a left factor, $T^{-1}\hat{D}^{-1}$ can be factored out to the left. We thus set

$$\mathcal{F} = T^{-1}\hat{D}^{-1} Q_1 \beta^{-1} \alpha \tag{5.19}$$

where $Q_1(z^{-1})$, of dimension $\ell|p$, is undetermined. With (5.19) inserted, and using $C_* \tilde{B}_* \tilde{N}_* = H_*$, (5.17) becomes

$$E\varepsilon(t)n^*(t) = \frac{1}{2\pi j} \oint T^{-1}\hat{D}^{-1} \{ z^{-m} \hat{S} C \phi H_* - Q_1 \beta_* \} \alpha_*^{-1} \mathcal{G}_* \frac{dz}{z} \quad .$$

All poles of every element of $\alpha_*^{-1}\mathcal{G}_*$ are located outside $|z| = 1$, since α is stable and $\mathcal{G}$ is causal and stable. In order to fulfill (2.8) collectively, we require

$$z^{-m} \hat{S} C \phi H_* = Q_1 \beta_* + z \hat{D} T L_* \quad . \tag{5.20}$$

Here, $Q_1(z^{-1})$ and $L_*(z)$ are polynomial matrices, of dimension $\ell|p$, with degrees

$$\begin{aligned} nQ_1 &\leq \max(nc + n\hat{s} + m, n\hat{d} + nt - 1) \\ nL &\leq \max(nh - m, n\beta) - 1 \quad . \end{aligned} \tag{5.21}$$

With β and $\hat{D}T$ stable, $\det \beta_*$ and $\det \hat{D}T$ will have no common factors. With the same reasoning as in Section V.B, equation (5.20) is found to have a unique solution, with degrees fulfilling (5.21). The derived result is formalized below.

Theorem 1

Let the system and input models be described by (5.11). Introduce the coprime factorizations (5.13), (5.18) and the spectral factorization (5.15). Under Assumptions 1. and 2, a H_2–optimal deconvolution estimator, of dimension $\ell|p$, is given by

$$\hat{f}(t|t+m) = T^{-1}\hat{D}^{-1}Q_1\beta^{-1}\tilde{N}\tilde{D}A\, y(t+m) \tag{5.22}$$

where $Q_1(q^{-1})$ together with $L_*(q)$, of dimensions $\ell|p$, is given by the *unique* solution to the bilateral polynomial matrix equation

$$q^{-m}\hat{S}C\phi C_*\tilde{B}_*\tilde{N}_* = Q_1\beta_* + q\hat{D}TL_* \quad . \tag{5.23}$$

The minimal criterion value is, with $H = \tilde{N}\tilde{B}C$,

$$\frac{1}{2\pi j}\oint \mathrm{tr}\{L_*\beta_*^{-1}\beta^{-1}L + T^{-1}SD^{-1}C(\phi - \phi H_*\beta_*^{-1}\beta^{-1}H\phi)C_*D_*^{-1}S_*T_*^{-1}\}\frac{dz}{z} \tag{5.24}$$

$\square$

Proof: The proof follows directly from the discussion above. Choosing $\mathcal{F}$ according to (5.19), $E\varepsilon(t)n^*(t) = (1/2\pi j)\oint L_*\alpha_*^{-1}\mathcal{G}_*dz = 0$, for admissible variations $n(t)$. (No poles are present in $|z| \leq 1$, in any element.) Consequently, the determined estimator minimizes $\mathrm{tr}E\varepsilon(t)\varepsilon^*(t)$. The minimal criterion value is obtained by inserting (5.22), (5.14), (5.18) and (5.23) (in this order) into $J = \mathrm{tr}E\varepsilon(t)\varepsilon^*(t)$ $\square$

Remarks. By making appropriate substitutions, it is easy to see that the problems and solutions presented in Section III and Subsection V.B are special cases of the result presented above. When S is square and S, D are both diagonal, $\hat{S} = S$ and $\hat{D} = D$. The coprime factorization (5.18) is then superfluous.

When $p = s = \ell$ and A, B, D, S and N are all diagonal, the equations (5.15), (5.22) and (5.23) become a direct generalization of those obtained for scalar problems in Ahlén and Sternad [18]. In this rather special case, without any need for coprime factorizations, $\tilde{D} = \hat{D} = D, \tilde{B} = B, \tilde{N} = N$, $\tilde{P} = DA$ and $\hat{S} = S$. Then, (5.15) gives $\beta\beta_* = NBC\phi C_*B_*N_* + DAM\psi M_*A_*D_*$ while (5.23) becomes $q^{-m}SC\psi C_*B_*N_* = Q_1\beta_* + qDTL_*$. See [18] for a comparison.

An alternative to introducing frequency weighting $T^{-1}S$ as in Figure 4, is to use $f(t) = u(t)$ and a weighted criterion $J_w = \mathrm{tr}E(T^{-1}S\varepsilon(t))(T^{-1}S\varepsilon(t))^*$. In such problems, the polynomial matrix S must be square ($s = \ell$) and stable, with S_o nonsingular. By simple loop transformations, it is easily verified that all design equations then remain unchanged, except the filter expression (5.19), which becomes $\mathcal{F} = S^{-1}\hat{D}^{-1}Q_1\beta^{-1}\alpha$. This estimator equals the H_2–estimator of [45], where it was presented as applicable only for square systems with stably invertible B. Weighted H_2–problems can be used as tools in H_∞ filter design, see [45] $\square$

6 A NUMERICAL EXAMPLE

The numerical feasibility of the polynomial equations approach will now be illustrated. Consider the estimation problem discussed in Section V.B, in the filtering case $m = 0$, with $p = 2$ measurements, $r = 2$ disturbance sources and one signal source ($k = 1$). Let

$$\begin{aligned} A &= \begin{bmatrix} 1-0.9q^{-1} & 0 \\ -(1-0.5q^{-1}) & 2 \end{bmatrix} \\ B &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 \\ -1 & 1+0.7q^{-1} \end{bmatrix} \\ \phi &= 15 \qquad\quad \psi = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \end{aligned} \tag{6.1}$$

Thus, the measurements (5.1) are given by

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} \frac{1}{1-0.9q^{-1}} \\ \frac{1-0.5q^{-1}}{2(1-0.9q^{-1})} \end{bmatrix} e(t) + \begin{bmatrix} \frac{1}{1-0.9q^{-1}} & 0 \\ \frac{0.4}{2(1-0.9q^{-1})} & \frac{1}{2}(1+0.7q^{-1}) \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} \tag{6.2}$$

By inspecting (6.1), we conclude that Assumptions 1. and 2. in Subsection V.B are fulfilled. Calculating the right hand side of (5.3), we obtain

$$DD_* = \begin{bmatrix} 16 & -1 \\ -1 & 1.4q^{-1} + 3.98 + 1.4q \end{bmatrix} \tag{6.3}$$

A stable left spectral factor, with D_o nonsingular, is

$$D = \begin{bmatrix} 4 & 0 \\ -\frac{1}{4} & 1.82447 + 0.76735q^{-1} \end{bmatrix} \tag{6.4}$$

This is easily verified, by computing DD_*. In order to determine the optimal filter (5.5), the polynomial matrix equation (5.7) has to be solved. From (5.8), we get $nY = nL = 0$. Writing the polynomial matrices as matrix polynomials, we have to solve

$$Y_0(D_0^T + D_1^T q) + q(A_0 + A_1 q^{-1})L_0 = B_0 \phi B_0^T \tag{6.5}$$

where M^T denotes transpose. Evaluation for equal powers of q gives

$$\begin{aligned} q^1 : \quad & Y_0 D_1^T + A_0 L_0 &= 0 \\ q^0 : \quad & Y_0 D_0^T + A_1 L_0 &= B_0 \phi B_0^T \end{aligned}$$

or, with numerical values inserted,

$$\begin{pmatrix} y_1 & y_2 \\ y_3 & y_4 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0.76735 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 \\ \ell_3 & \ell_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \tag{6.6a}$$

$$\begin{pmatrix} y_1 & y_2 \\ y_3 & y_4 \end{pmatrix} \begin{pmatrix} 4 & -0.25 \\ 0 & 1.82447 \end{pmatrix} + \begin{pmatrix} -0.9 & 0 \\ 0.5 & 0 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 \\ \ell_3 & \ell_4 \end{pmatrix} = \begin{pmatrix} 15 & 0 \\ 0 & 0 \end{pmatrix} \tag{6.6b}$$

The matrix equations (6.6a) and (6.6b) constitute a system of eight scalar equations in eight unknowns. Solving (6.6a) and (6.6b) for element (1,1) and (2,1) gives $\ell_1 = 0, \ell_3 = 0, y_1 = 3.75, y_3 = 0$, while the elements (1,2) and (2,2) can be written as

$$\begin{aligned} 0.76735 y_2 + \ell_2 &= 0 \\ 0.76735 y_4 - \ell_2 + 2\ell_4 &= 0 \\ -0.25 y_1 + 1.82447 y_2 - 0.9\ell_2 &= 0 \\ -0.25 y_3 + 1.82447 y_4 + 0.5\ell_2 &= 0 \ . \end{aligned} \tag{6.7}$$

From (6.7), we obtain $y_2 = 0.37275$, $y_4 = 0.07839$, $\ell_2 = -0.28603$, $\ell_4 = -0.17309$. Thus, the solution of (6.5) becomes

$$Y_0 = \begin{pmatrix} 3.75 & 0.37275 \\ 0 & 0.07839 \end{pmatrix} \quad L_0 = \begin{pmatrix} 0 & -0.28603 \\ 0 & -0.17309 \end{pmatrix}$$

Having calculated Y_0, the optimal filter is found to be

$$\mathcal{F} = A^{-1} Y_0 D^{-1} A = \frac{1}{1 - 0.479q^{-1} - 0.378q^{-2}}$$

$$\times \begin{bmatrix} 0.746 - 0.359q^{-1} - 0.355q^{-2} & 0.409 \\ 0.353 - 0.338q^{-1} - 0.0963q^{-2} + 0.0887q^{-3} & 0.247 - 0.141q^{-1} \end{bmatrix}$$

Obviously, the estimator is stable and causal. It is interesting to note that both of the measurements are utilized in the estimation of each signal.

Remark on numerical algorithms. In the process of calculating the optimal estimator, polynomial matrix spectral factorization and solution of a bilateral Diophantine equation were performed, as well as polynomial matrix inversions and multiplications. From a computational point of view, it is attractive to diagonalize (6.5) by means of transformation to the Smith form [10], [15]. This would result in a system of $p^2 = 4$ scalar decoupled equations. The solution can be obtained by solving these equations, followed by back–transformation. Efficient algoritms for calculating the Smith form and inverses of polynomial matrices are given in [10], chapter 7. Spectral factorization of polynomial matrices is discussed in [39] and [10], where efficient algorithms are given. See also [38]. Algorithms for coprime factorizations, required in Section V.C, are included in [10]. See also Appendix G–6 of [31]. (Some of the algorithms have to be modified slightly, if the polynomial coefficients are allowed to be complex–valued.)

7 CONCLUSIONS

We have presented a simple methodology for obtaining design equations for predictors, filters and fixed lag smoothers. The minimal estimation error should be orthogonal to any admissible variation of the estimate. Evaluation of the orthogonality in the frequency domain, by cancelling stable poles by zeros, is the main idea behind the method. In contrast to the classical Wiener formulation, the estimators are explicitly parametrized in terms of rational transfer functions.

In scalar estimation problems, the design equations consist of a polynomial spectral factorization and one linear Diophantine equation. The spectral factorization represents the calculation of a whitening filter. The Diophantine equation constitutes a linear system of simultaneous equations. It can be seen as a convenient way to perform a partial fraction expansion, to calculate the causal part of a realizable Wiener filter.

In multi–signal estimation problems, signal and noise–models are expressed by polynomial matrix fractions. The orthogonality requirement is fulfilled by elementwise cancellation of stable poles by zeros. The collective fulfillment of these relations define polynomial matrix equations. Only a few lines of calculation will usually be needed to obtain these equations, a major advantage with the suggested methodology. The difficult task of performing spectral factorization of rational matrices is avoided, in the polynomial equations approach. Instead, spectral factorization of polynomial matrices are utilized, sometimes in combination with coprime factorizations. Furthermore, a bilateral Diophantine equation has to be solved. Efficient numerical procedures exist for these operations.

Acknowledgements

We thank Bengt Carlsson and Vladimir Kučera for valuable comments on an early version of this paper.

A THE "COMPLETING THE SQUARES" APPROACH

The "completing the squares" method in the frequency domain has been used, for example, in [15] and [17]. A time–domain variant is used in [10]. In the example of Section III, the use of (3.1) in the criterion (2.2) gives, with a rational filter $\mathcal{F}$,

$$\begin{aligned} J &= E|s(t) - \mathcal{F}y(t+m)|^2 \\ &= E\left|(1-q^m\mathcal{F})\frac{C}{D}e(t)\right|^2 + E\left|q^m\mathcal{F}\frac{M}{N}v(t)\right|^2 \\ &= \frac{\lambda_e}{2\pi j}\oint\left((1-z^m\mathcal{F})(1-z^{-m}\mathcal{F}_*)\frac{CC_*}{DD_*} + \rho\mathcal{F}\mathcal{F}_*\frac{MM_*}{NN_*}\right)\frac{dz}{z} \\ &= \frac{\lambda_e}{2\pi j}\oint\left(\frac{CC_*}{DD_*} - z^m\mathcal{F}\frac{CC_*}{DD_*} - \frac{CC_*}{DD_*}z^{-m}\mathcal{F}_* + \mathcal{F}\mathcal{F}_*\frac{r\beta\beta_*}{DD_*NN_*}\right)\frac{dz}{z} \quad . \end{aligned}$$

In the third equality, Parseval's formula was used and in the last, the spectral factorization (3.3) was inserted. By completing the square, we obtain

$$\begin{aligned} J &= \frac{\lambda_e}{2\pi j}\oint r\left(\frac{\beta}{DN}\mathcal{F} - \frac{z^{-m}CC_*N_*}{r\beta_*D}\right)\left(\frac{\beta_*}{D_*N_*}\mathcal{F}_* - \frac{z^mC_*CN}{r\beta D_*}\right)\frac{dz}{z} \\ &+ \frac{\lambda_e}{2\pi j}\oint\left(\frac{CC_*}{DD_*} - \frac{CC_*CC_*NN_*}{r\beta\beta_*DD_*}\right)\frac{dz}{z} \triangleq J_1 + J_2 \quad . \end{aligned} \tag{A.1}$$

The first term in (A.1), J_1, depends on $\mathcal{F}$ while the second term, J_2, does not. If $\mathcal{F}$ were not restricted to be realizable (stable and causal), the problem could have been solved by choosing $\mathcal{F}$ such that $J_1 = 0$.

A realizable $\mathcal{F}$ can only eliminate the causal parts of the integrand of J_1. Since $(\beta/DN)\mathcal{F}$ is causal, it remains to partition $(z^{-m}CC_*N_*)/(r\beta_*D)$. Let

$$\frac{z^{-m}CC_*N_*}{r\beta_*D} = \frac{Q_1}{D} + \frac{zL_*}{r\beta_*} \tag{A.2}$$

for some polynomials Q_1 and L_*. The term $Q_1(z^{-1})/D(z^{-1})$ represents the causal part and $(zL_*(z)/(r\beta_*(z))$ the noncausal part. Equivalently, (A.2) can be written

$$z^{-m}CC_*N_* = r\beta_*Q_1 + zDL_* \tag{A.3}$$

which is (3.6), with z exchanged for q. Using (A.2), J_1 may be expressed as

$$J_1 = \frac{\lambda_e}{2\pi j}\oint r\left(\frac{\beta}{DN}\mathcal{F} - \frac{Q_1}{D} - \frac{zL_*}{r\beta_*}\right)\left(\frac{\beta}{DN}\mathcal{F} - \frac{Q_1}{D} - \frac{zL_*}{r\beta_*}\right)_*\frac{dz}{z} \quad .$$

Expanding the integrand, J_1 becomes a sum of four terms

$$\begin{aligned}
V_1 &= \frac{\lambda_e}{2\pi j}\oint r\left(\frac{\beta}{DN}\mathcal{F}-\frac{Q_1}{D}\right)\left(\frac{\beta_*}{D_*N_*}\mathcal{F}_*-\frac{Q_{1_*}}{D_*}\right)\frac{dz}{z} \\
V_2 &= -\frac{\lambda_e}{2\pi j}\oint \left(\frac{\beta}{DN}\mathcal{F}-\frac{Q_1}{D}\right)\frac{z^{-1}L}{\beta}\frac{dz}{z} \\
V_3 &= -\frac{\lambda_e}{2\pi j}\oint z\frac{L_*}{\beta_*}\left(\frac{\beta_*}{D_*N_*}\mathcal{F}_*-\frac{Q_{1_*}}{D_*}\right)\frac{dz}{z} \\
V_4 &= \frac{\lambda_e}{2\pi j}\oint \frac{LL_*}{r\beta\beta_*}\frac{dz}{z} \quad .
\end{aligned}$$

For any causal and stable choice of the rational filter $\mathcal{F}$, all poles of the integrand of V_3 will be located outside the unit circle, since β, D and N are stable. Hence, $V_3 = 0$. (Note that it is crucial that zL_*/β_* is strictly noncausal, starting with a free z. This z cancels the pole at the origin of V_3.) For symmetry reasons, $V_2 = 0$. The term V_4 does not depend on $\mathcal{F}$. Thus, the criterion J_1 is minimized by eliminating V_1:

$$\frac{\beta}{DN}\mathcal{F}-\frac{Q_1}{D}=0$$

which gives

$$\mathcal{F}=\frac{Q_1N}{\beta}$$

where Q_1, together with L_*, is the solution to (A.3) and β is the stable polynomial spectral factor. The minimal criterion value is $J_{\min} = J_2 + V_4$.

This derivation should be compared to steps 2. and 3. in the derivation in Section III.

$\square$

B THE CLASSICAL WIENER SOLUTION, IN A POLYNOMIAL SYSTEMS FRAMEWORK

Wiener filters are designed by first whitening the measurements and then using the cross spectral density, ϕ_{fw}, between desired response and whitened mesurement. See, for example, [2]–[6]. For the problem depicted in Figure 1, the causal Wiener filter is $\hat{f} = \{\phi_{fw}\}_+ w$, where $w = \mathcal{W}(z^{-1})y$ is the whitened measurement. Thus,

$$\mathcal{F}(z^{-1}) = \{\phi_{fw}\}_+ \mathcal{W}(z^{-1}) = \{\phi_{fy}\mathcal{W}_*(z)\}_+ \mathcal{W}(z^{-1}) \quad . \tag{B.1}$$

The whitening filter is denoted $\mathcal{W}(z^{-1})$ and $\mathcal{W}_*(z)$ is its conjugate–transpose, while ϕ_{fy} is the desired signal–measurement cross–spectral density. The bracket $\{\cdot\}_+$ represents the use of only the causal part of the weighting function.

The expression (B.1) is simple. While explicit in terms of factored components of ϕ_{fw}, it is however not explicit in terms of polynomial coefficients of rational transfer functions of the signal and noise models. It is not, as it stands, parametrized by a finite number of parameters. The polynomial systems framework is of help here.

In the scalar example discussed in Section III, the whitening filter is the inverse of the innovations model (3.2). Since $v(t)$ and $e(t)$ are mutually independent and the measurement is $y(t+m)$, we obtain, with $f(t) = s(t)$,

$$\phi_{fy} = \phi_{s(t)s(t+m)} = \frac{C}{D} z^{-m} \frac{C_*}{D_*} \lambda_e \quad .$$

Thus, (B.1) becomes, with $r \triangleq \lambda_\eta / \lambda_e$ and q exchanged for z,

$$\mathcal{F}(q^{-1}) = \left\{ \frac{C}{D} q^{-m} \frac{C_*}{D_*} \lambda_e \frac{D_* N_*}{\sqrt{\lambda_\eta}\beta_*} \right\}_+ \frac{DN}{\sqrt{\lambda_\eta}\beta} = \left\{ q^{-m} \frac{C C_* N_*}{D r \beta_*} \right\}_+ \frac{DN}{\beta} \quad . \tag{B.2}$$

Extraction of the causal part $\{\cdot\}_+$ of the double–sided weighting function corresponds to a partial fraction expansion of the rational function. Let

$$q^{-m} \frac{C(q^{-1}) C_*(q) N_*(q)}{D(q^{-1}) r \beta_*(q)} = \frac{Q_1(q^{-1})}{D(q^{-1})} + \frac{\tilde{L}_*(q)}{r\beta_*(q)} \tag{B.3}$$

for some polynomials $Q_1(q^{-1})$ and $\tilde{L}_*(q)$. Terms without delay are included in the causal part, so the noncausal part starts with a free q–term. Thus, let $\tilde{L}_*(q) \triangleq q L_*(q)$. (This avoids the occurrence of an error pointed out by Chen [7].) Multiplication of both sides of (B.3) by $Dr\beta_*$ then gives

$$q^{-m} C C_* N_* = r \beta_* Q_1 + q D L_* \quad .$$

This is precisely the linear polynomial equation (3.6). Thus, the causal Wiener estimator is

$$\mathcal{F}(q^{-1}) = \left\{ \frac{Q_1}{D} + \frac{qL_*}{r\beta_*} \right\}_+ \frac{DN}{\beta} = \frac{Q_1 DN}{D\beta} \tag{B.4}$$

which is (3.7), if the stable common factor D is cancelled. (Of course, unstable systems are not allowed in the classical Wiener formulation.) Grimble [17], [26] and Sderstrm [25] have noted the link between partial fraction expansions, such as (B.3), and Diophantine equations. This link is also a key part of the "completing the squares"–reasoning. (See (A.2) in Appendix A.) □

C PARSEVAL'S FORMULA

Consider Parseval's formula for the covariance between two complex–valued signals $x(t) = \mathcal{G}(q^{-1})e(t)$ and $w(t) = \mathcal{H}(q^{-1})e(t) = (\sum_0^\infty H_j q^{-j})e(t)$. When $e(t)$ is white noise with covariance matrix $Ee(t)e^*(t) = \Lambda$, we have, cf [29],

$$Ex(t)w^*(t) = \frac{1}{2\pi j}\oint_{|z|=1} \mathcal{G}(z)\Lambda\mathcal{H}^*(\frac{1}{z^*})\frac{dz}{z}$$

which in our notation, with $q \Leftrightarrow z$ and the definition of $\mathcal{H}_*$, is

$$\begin{aligned}\frac{1}{2\pi j}\oint_{|z|=1} \mathcal{G}(z^{-1})\Lambda\mathcal{H}^*(z^*)\frac{dz}{z} &= \frac{1}{2\pi j}\oint_{|z|=1} \mathcal{G}(z^{-1})\Lambda(H_o + H_1 z^* + \ldots)^*\frac{dz}{z} \\ &= \frac{1}{2\pi j}\oint_{|z|=1} \mathcal{G}(z^{-1})\Lambda\mathcal{H}_*(z)\frac{dz}{z} \quad .\end{aligned}$$

□

References

[1] N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series*, The Technology Press and Wiley, New York, 1950.

[2] H. W. Bode and C. E. Shannon, “A Simplified Derivation of Linear Least Square Smoothing and Prediction Theory”, *Proceedings of the I.R.E.*, vol 38, pp 417–425, April 1950.

[3] Y. W. Lee, *Statistical Theory of Communication*, Wiley, New York, 1960.

[4] H. L. van Trees, *Detection Estimation and Modulation Theory*, Part I, Wiley, New York, 1968.

[5] T. Kailath, *Lectures on Wiener and Kalman Filtering*, Springer, Wien, 1981.

[6] J. A. Cadzow, *Foundations of Digital Signal Processing and Data Analysis*, Macmillan, New York, 1987.

[7] C.–T. Chen, “On digital Wiener filters”, *Proc IEEE*, vol 64, pp 1736–1737, December 1976.

[8] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, New Jersey, 1986.

[9] V. Kučera, “Transfer–function solution to the Kalman–Bucy filtering problem”, *Kybernetika (Prague)*, vol 14, pp 110–122, 1978.

[10] V. Kučera, *Discrete Linear Control. The Polynomial Equation Approach*, Wiley, Chichester, 1979.

[11] V. Kučera, “Stochastic multivariable control: a polynomial equation approach”, *IEEE Transaction on Automatic Control*, vol AC–25, pp 913–919, October 1980.

[12] V. Kučera, “New results in state estimation and regulation”, *Automatica*, vol 17, pp 745–748, September 1981.

[13] V. Kučera, “The LQG control problem: a study of common factors”, *Problems of Control and Information Theory*, vol 13, pp 239–251, no 4, 1984.

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

[15] A. P. Roberts and M. M. Newmann, “Polynomial approach to Wiener filtering”, *International Journal of Control*, vol 47, pp 681–696, March 1988.

[16] J. F. Barret and T. J. Moir, “A unified approach to multivariable discrete–time filtering based on the Wiener theory”, *Kybernetika (Prague)*, vol 23, pp 177–197, 1987.

[17] M. J. Grimble, “Polynomial systems approach to optimal linear filtering and prediction”, *International Journal of Control*, vol 41, pp 1545–1564, June 1985.

[18] A. Ahlén and M Sternad, “Optimal deconvolution based on polynomial methods”, *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol 37, pp 217-226, February 1989.

[19] M. Sternad and A. Ahlén, “The structure and design of realizable decision feedback equalizers for IIR–channels with coloured noise”, *IEEE Transactions on Information Theory*, vol IT–36, pp 848–858, July 1990.

[20] M. Sternad and T. Sderstrm, “LQG–optimal feedforward regulators”, *Automatica*, vol 24, pp 557–561, July 1988.

[21] Z. L. Deng, “White–noise filter and smoother with application to seismic data deconvolution”, in *Preprints 7th IFAC/IFORS Symp. Identification Syst. Parameter Estimation*, York, U.K., pp 621–624, July 1985.

[22] T. J. Moir, “Optimal deconvolution smoother”, *Proc. IEE*, pt D, vol 133, no 1, pp 13–18, Jan. 1986.

[23] M. Sternad, *Optimal and Adaptive Feedforward Regulators*, Ph. D. thesis, Uppsala University, Uppsala, Sweden, 1987.

[24] E. Trulsson, “Uniqueness of local minima for linear quadratic control design”, *System and Control Letters*, vol 5, pp 295–302, April 1985.

[25] T. Sderstrm, “Some optimization problems for stochastic systems”, *Report UPTEC 8530K*, Institute of Technology, Uppsala University, Uppsala, Sweden, 1985.

[26] M. J. Grimble and M. A. Johnsson, *Optimal Control and Stochastic Estimation*, Wiley, Chichester, 1988.

[27] K. J. strm and B. Wittenmark, *Computer–Controlled Systems*, Prentice–Hall, Englewood Cliffs, New Jersey, 1984.

[28] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice–Hall, Englewood Cliffs, New Jersey, 1979.

[29] A. V. Oppenheim and R. W. Shafer, *Digital Signal Processing*, Prentice–Hall, Englewood Cliffs, New Jersey, 1975.

[30] T. Kailath, *Linear Systems*, Prentice–Hall, Englewood Cliffs, New Jersey, 1980.

[31] C.–T. Chen, *Linear System Theory and Design*, Holt, Rinehart and Winston, New York, 1984.

[32] U. Shaked, “A generalized transfer function approach to linear stationary filtering and steady–state optimal control problems”, *International Journal of Control*, vol 29, pp 741–770, December 1976.

[33] R. E. Kalman, “A new approach to linear filtering and prediction problems”, Trans. AMSE, Journal of Basic Engineering, vol 82D, pp 35-45, March 1960.

[34] G. Demoment and R. Reynaud, “Fast minimum variance deconvolution”, *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol ASSP–33, pp 1324–1326, October 1985.

[35] J. M. Mendel, *Optimal Seismic Deconvolution. An Estimation–Based Approach.* Academic Press, New York, 1983.

[36] J. Ježek, “New algorithm for minimal solution of linear polynomial equations”, *Kybernetika (Prague)*, vol 18, pp 505–516, 1982.

[37] Z. Vostrý, "New algorithm for polynomial spectral factorization with quadratic convergence", *Kybernetika (Prague)*, vol 11, pp 415–422, 1975.

[38] G. T. Wilson, "The factorization of matricial spectral densities", *SIAM Journal of Applied Mathematics*, vol 23, pp 420–426, December 1972.

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

[40] M. Sternad and A. Ahlén, "A novel derivation methodology for polynomial–LQ controller design", Report UPTEC 90058R, Institute of Technology, Uppsala University, Uppsala, Sweden. Submitted.

[41] B. Carlsson, A. Ahlén and M. Sternad, "Optimal differentiation based on stochastic signal models", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol ASSP–39, February 1991.

[42] J. Ježek, "Conjugated and symmetric polynomial equations. II: discrete–time systems", *Kybernetika (Prague)*, vol 19, pp 196–211, 1983.

[43] M. J. Grimble, "Single versus double diophantine equation debate: comments on 'A polynomial approach to Wiener filtering' ", *International Journal of Control*, vol 48, pp 2161–2165, November 1988.

[44] B. Carlsson, M. Sternad and A. Ahlén, "Digital differentiation of noisy data, measured through a dynamical system". Report UPTEC 90087R, Institute of Technology Uppsala University, Uppsala, Sweden. Submitted.

[45] M. J. Grimble and A. ElSayed, "Solution of the H_∞ optimal linear filtering problem for discrete–time systems", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol ASSP–38, pp 1092–1104, July 1990.