SPIE Annual Symposium, San Diego, USA, July 1991

ADAPTIVE DECONVOLUTION BASED ON SPECTRAL DECOMPOSITION

Anders Ahlén and Mikael Sternad

Automatic Control and Systems Analysis Group,
Department of Technology, Uppsala University.
P. O. Box 27, S-751 03 Uppsala, Sweden.

ABSTRACT

An adaptive algorithm for estimating the input to a linear system is presented. This explicit self–tuning filter is based on the identification of an innovations model. From that model, input and measurement noise ARMA–descriptions are decomposed, using second order moments. Identifiability results guarantee a unique decomposition. Main tools in the algorithm are the solution of two linear systems of equations. The filter design is based on the polynomial approach to Wiener filtering.

1. INTRODUCTION

The need to restore signals, observed through linear systems and contaminated by noise, arises frequently. In order to find the desired signal, some kind of inverse filtering is needed. Such filtering is known as *deconvolution* or *input estimation.* The problem has been discussed by many authors[3,9,10,11,14,15].

In many applications, it is reasonable to assume some crucial part of the system to be known *a priori* or determined in advance by experiments. We will assume the measurement system (the transducer) to be known here. The input and measurement noise properties may, however, be unknown, and vary with time. Ahlén has suggested an adaptive approach based on an innovations model, in order to estimate the input[1,4]. In this paper, we will generalize this approach to the estimation of filtered inputs. A new algorithm is also suggested for cases with ARMA–noise, with both MA–and AR–parts unknown. Moir *et.al.* have developed two related self–tuning algorithms, which turn out to be special cases of ours. One concerns signal estimation with known noise spectrum and no observation system[16]. The other treats adaptive deconvolution with white input and noise[15].

2. PRELIMINARIES

2.1 The deconvolution problem

Consider the linear, causal, discrete–time system

$$y(k) = \frac{B(q^{-1})}{A(q^{-1})}u(k-\tau) + w(k) \tag{1}$$

where q^{-1} denotes the backward shift operator ($q^{-1}v(k) = v(k-1)$), while q is the forward operator ($qv(k) = v(k+1)$). The input $u(k)$ and measurement noise $w(k)$ are assumed to be accurately described by the ARMA–processes

0-8194-0693-7/91/$4.00

$$u(k) = \frac{C(q^{-1})}{D(q^{-1})}e(k) \quad ; \quad w(k) = \frac{M(q^{-1})}{N(q^{-1})}v(k) \tag{2}$$

$$Ee(k)^2 = \lambda_e \qquad Ev(k)^2 = \lambda_v \qquad \rho = \lambda_v / \lambda_e \ .$$

Here, $e(k)$ and $v(k)$ are mutually uncorrelated. They are stationary white and zero mean stochastic sequences. All polynomials, except $B(q^{-1})$, are monic. Using measurements of the output $y(k)$, up to time $k+m$, we seek the stable, linear and time–invariant estimator

$$\hat{d}(k|k+m) = \frac{Q(q^{-1})}{R(q^{-1})}y(k+m) \tag{3}$$

of the possibly filtered input signal

$$d(k) = \frac{S(q^{-1})}{T(q^{-1})}u(k+\ell) \, . \tag{4}$$

See Figure 1. The estimator is to minimze $Ez(k)^2 \triangleq E(d(k) - \hat{d}(k|k+m))^2$. Depending on the sign of m, it is a predictor ($m < 0$), a filter ($m = 0$) or a fixed lag smoother ($m > 0$). The filter $q^\ell S(q^{-1})/T(q^{-1})$ in (4) is stable and possibly noncausal ($\ell > 0$). It is assumed to be specified by the user, and thus to be exactly known. It can be used to reduce the estimator gain outside a restricted interesting frequency range[3]. This type of filter also appears e.g. in digital differentiation problems[7,8]. There, it represents a discrete–time approximation of the derivative operator.

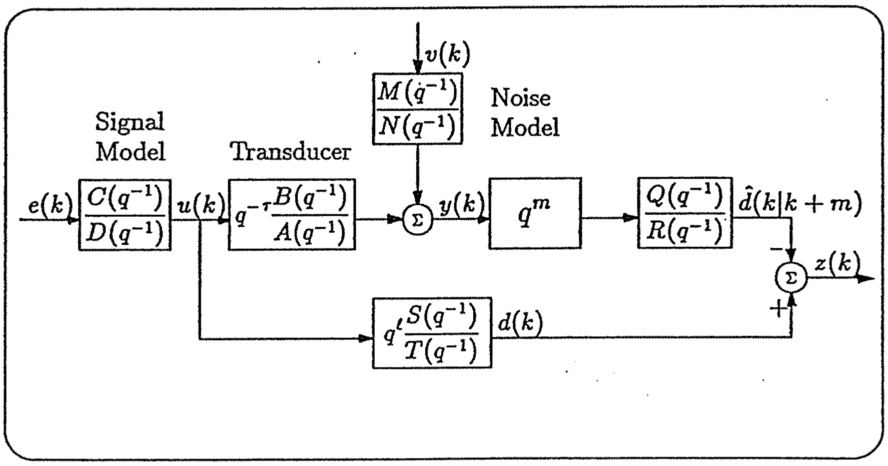

Figure 1: A single channel deconvolution problem, with coloured measurement noise. The filtered input $d(k)$ is to be estimated, from measurements of the output $y(k+m)$.

2.2 Wiener filter design, using polynomial equations

A procedure for optimizing input estimators in the transfer function form (3) has been derived[3]. Compared to Kalman filtering, this approach is simpler for scalar signals, and is well suited for self-tuning applications. It is based on the following assumptions:

- The signal $y(k)$ and input $u(k)$ can be described by linear models with structure (1), (2). All polynomials in (1) and (2), with degrees na, nb etc, are known.
- The denominators $D(z^{-1})$, $A(z^{-1})$ and $N(z^{-1})$ may have zeros inside or on, but not outside, the stability limit. The polynomials $C(z^{-1})B(z^{-1})N(z^{-1})$ and $M(z^{-1})A(z^{-1})D(z^{-1})$ have no common factors with zeros on $|z| = 1$.

For any polynomial $P(q^{-1}) = p_o + p_1 q^{-1} + \ldots + p_{np} q^{-np}$, define the *conjugate polynomial* $P_*(q) = p_o + p_1 q + \ldots + p_{np} q^{np}$. Also, define $\overline{P}(q^{-1}) = q^{-np} P_*(q)$. In the frequency domain, the complex variable z is substituted for q. Polynomial arguments will often be omitted.

The MSE–optimal linear deconvolution estimator is given by

$$\hat{d}(k|k+m) = \frac{Q_1 N A}{T\beta} y(k+m) \tag{5}$$

where $\beta(q^{-1})$, of degree $n\beta$, is the stable monic solution to a polynomial spectral factorization equation

$$r\beta\beta_* = CBNC_*B_*N_* + \rho MADM_*A_*D_* \tag{6}$$

with r being a scalar. The polynomial $Q_1(q^{-1})$, together with a polynomial $L_*(q)$, is the unique solution to the linear Diophantine equation

$$q^{\ell+\tau-m} SCC_*B_*N_* = r\beta_* Q_1 + qDTL_* \tag{7}$$

with polynomial degrees

$$\begin{aligned} nQ_1 &= \max(nc + ns - \ell - \tau + m, nd + nt - 1) \\ nL &= \max(nc + nb + nn + \ell + \tau - m, n\beta) - 1 \ . \end{aligned} \tag{8}$$

The minimal mean square error is

$$Ez(k)^2_{min} = \frac{\lambda_e}{2\pi i} \oint \frac{LL_*}{r\beta\beta_*} + \rho \frac{SS_*}{TT_*} \frac{CMAC_*M_*A_*}{r\beta\beta_*} \frac{dz}{z} \ . \tag{9}$$

This solution can be generalized to complex–valued signals and multivariable problems[5]. A new constructive techique for deriving polynomial design equations, such as those above, has recently been developed[5]. The IIR-filter (5) is stable, since $T(z^{-1})\beta(z^{-1})$ is stable. It may contain stable common factors. Note, that when the noise model has resonances (zeros of N close to the unit circle), the filter has notches at the corresponding frequencies.

The linear equation (7) corresponds to a linear system of equations in the coefficients of $Q_1(q^{-1})$ and $L_*(q)$. By setting the maximal degrees in q^{-1} and in q, respectively, equal on both sides of (7), the degrees (8) of the unknown polynomials are obtained. The number of equations will equal the number of unknowns. The system of equations is nonsingular, since $\beta_*(z)$ (unstable) and $D(z^{-1})T(z^{-1})$ (stable) cannot have any common factors. Consequently, a unique solution exists.

The estimator constitutes a Wiener filter, optimized by using polynomial equations. Comparison with a classical Wiener solution reveals that the spectral factorization corresponds to the calculation of a *whitening filter*. The calculation of the *causal part* $\{\cdot\}_+$ of the Wiener filter is performed by solving the Diophantine equation (7)[5].

2.3 Identifiability

From a practical point of view, we can hardly expect all polynomials to be known *a priori*. It is reasonable, in many applications, to assume $q^{-\tau}B(q^{-1})/A(q^{-1})$ to be known. This will be assumed in the following. Under what conditions is it then possible to determine the ARMA–models (2), from output data only? The identifiability properties of the input estimation problem have been investigated[2], under the following asssumptions:

- A, B, τ and polynomial degrees nc, nd, nm, nn are known *a priori.*
- All polynomials in (1),(2) are asymptotically stable, except B, which may be unstable.
- The polynomial pairs

$$(A,B),(C,D),(M,N),(B,D)$$
$$(\overline{B},A),(\overline{B},D),(A,N),(N,D) \tag{10}$$

  are all coprime.
- The only measurable information is $\phi_y(e^{i\omega})$, the spectral density of the output.

We then have the following result[2]. Introduce the numbers

$$\Delta D \triangleq nd - nc \quad ; \quad \Delta N \triangleq nn - nm \tag{11}$$

$$g \triangleq \max(na + \Delta D, nb + \Delta N) \quad ; \quad h \triangleq 2\min(nn, nd) \ .$$

For unique identification of $\{C, D, M, N, \lambda_e, \lambda_v\}$, the condition

$$g \geq 1 \tag{12}$$

is necessary, while

$$g > h \tag{13}$$

is sufficient. If the N–polynomial is known *a priori*, the number of unknown parameters decreases by nn. We thus set $N(q^{-1}) = 1$ and $nn = 0$ in (10)–(13). Then, the rather weak requirement $g \geq 1$ is both necessary and sufficient for parameter identifiability.

3. A SELF–TUNING DECONVOLUTION ALGORITHM

When designing an optimal estimator from (7), we find that S, C, B, N, β, D and T are required. Of these, only B, S and T are assumed to be known *a priori*. We thus have to estimate (C, D, N, β) in some way, using the output measurements only. Consider an *innovations model* of the output $y(k)$. Assuming A, D and N to be stable, it is given by

$$y(k) = \frac{\beta}{ADN}\tilde{y}(k) \quad .$$

Here, β is the stable spectral factor from (6), and $\tilde{y}(k)$ is the innovation sequence. The A–polynomial is known and can be filtered out. Thus, β/DN may be estimated. Unfortunately, C and M are related to β through the nonlinear equation (6). Note, however, that only the *product* CC_* is needed in (7). A key idea is to formulate (6) as a *linear* system of equations in the coefficients of CC_* and MM_*. Since we have an estimate of β, the system can be solved. For now, assume N to be known *a priori* or equal to unity. This leads to the algorithm below.

3.1 Algorithm for moving average–noise

Algorithm 1.

Assume $S(q^{-1}), T(q^{-1}), A(q^{-1}), B(q^{-1}), N(q^{-1}), \ell, \tau, n\beta, nd, nc$ and nm to be known, such that parameter identifiability holds. For each data,

1. Generate the signal $x(k) = A(q^{-1})N(q^{-1})y(k)$.
2. Update recursive estimates of the β and D in the ARMA–model $x(t) = (\hat{\beta}/\hat{D})\tilde{y}(t)$. Use a recursive prediction error method (RPEM) or extended least squares (ELS)[13].
3. Using $\hat{\beta}$ and $\hat{D}$, the equation (6)
$$AA_*\hat{D}\hat{D}_*\left(\frac{\rho}{r}MM_*\right) + BB_*NN_*\left(\frac{1}{r}CC_*\right) = \hat{\beta}\hat{\beta}_* \qquad (14)$$
is now seen as a linear polynomial equation in $(\rho/r)MM_*$ and $(1/r)CC_*$. It corresponds to an overdetermined linear system of equations. Check the condition number of its Sylvester matrix $\mathsf{S}(AA_*\hat{D}\hat{D}_*, BB_*NN_*)$. If OK, solve it, in the least squares sense.
4. Solve (7), using the estimates $\hat{\beta}$, $\hat{D}$ and $(1/r)CC_*$ from step 3.
$$q^{\ell+\tau-m}S\left(\frac{1}{r}\hat{C}\hat{C}_*\right)B_*N_* = \hat{\beta}_*Q_1 + q\hat{D}T\left(\frac{1}{r}L_*\right) \qquad (15)$$
This gives $Q_1(q^{-1})$ and $(1/r)L_*(q)$, with the degrees from (8).
5. Perform the filtering (3) or, alternatively,
$$\hat{d}(k|k+m) = \frac{Q_1(q^{-1})}{T(q^{-1})\hat{\beta}(q^{-1})}x(k+m) \quad .$$

3.2 Comments and interpretations

1. We assume the system to be parameter identifiable, i.e. (10) and (13) to hold. With a known $N(q^{-1})$, the chances to have parameter identifiability are good. The correct polynomial degrees are assumed known, but this restrictive assumption can be avoided. Over–parametrized models are discussed in Section 3.4.

2. For time–invariant systems, the polynomials $\hat{\beta}$ and $\hat{D}$ in the innovations model can always be correctly estimated asymptotically, using the prediction errror method[13]. Time–variable parameters may be tracked by means of a forgetting factor.

3. The estimate $\hat{\beta}$ is monic. Note that the noise ratio $\rho = \lambda_v / \lambda_e$ is assumed unknown.

4. There is no need to perform any spectral factorization like (6), since the innovations model is estimated directly. Instead, the linear system of equations in Step 3 must be solved. In the transient phase, when $\hat{\beta}$ and $\hat{D}$ approach β and D, Step 3 computes the least squares solution. Asymptotically, when $\hat{\beta} = \beta$ and $\hat{D} = D$, there exists a unique and exact solution.

5. To reduce the computational requirements, the linear system defined by (14) can be transformed into a minimal order system with $n\beta + 1$ equations and $nc + nm + 2$ unknowns. This is achieved by eliminating rows and columns which are redundant due to symmetry[1].

6. Although the system is known to be identifiable when (10) and (13) hold, it may happen that the estimate $\hat{D}$ and BN sometimes have almost common factors as $\hat{D}$ converges towards D. This will cause a rank deficiency of $\mathrm{S}(AA_*\hat{D}\hat{D}_*, BB_*NN_*)$ in Step 3. The linear system becomes under–determined. The singular values of S must be therefore be checked. (When singular value decomposition is used for solving the LS problem, this requires no extra computations.) If the condition number is large, the filter should not be updated.

The computational complexity of Algorithm 1 is presented in Table 1. The approximate number of mult–add operations required per sample has been estimated. All degrees are assumed equal, $na, nb, \ldots$ etc $= n$, and $S = T = 1$. In Step 3, the least squares solution is assumed to be computed by singular value decomposition, using the Golub-Reinsch algorithm[12]. Redundancies due to symmetry in the linear system are eliminated. With $n = 1$ and $m = 2$, approximately 700 floating point mults+adds are required per sample. It should be noted that there is hardly any need to recalculate the filter (steps 3 and 4) at each sample.

Table 1. The computational complexity of Algorithm 1.

1:	Prefiltering	$2n$
2:	Identification	$30n^2 + 23n$
3:	LS solution	$56n^3 + 152n^2 + 136n$
4:	Linear system (2.7)	$\frac{1}{3}(4n + m + 2)^3 + \frac{3}{2}(4n + m + 2)^2$
5:	Filtering	$6n + m$

3.3 An extended algorithm

If $N(q^{-1})$ is unknown, and the noise cannot be described with sufficient accuracy by a low order MA–model, adaptive deconvolution becomes somewhat more complex. For general unknown ARMA–noise, we suggest the following algorithm. It is applicable when the degree of either $B(q^{-1})$ or $A(q^{-1})$ is sufficiently high.

Algorithm 2.

Assume $S(q^{-1}), T(q^{-1}), A(q^{-1}), B(q^{-1}), \ell, \tau, n\beta, nd, nc, nm$ and nn to be known, such that parameter identifiability holds. Assume that (18) below holds. For each data,

1 Generate the signal $x(k) = A(q^{-1})y(k)$.

2 Do recursive identification, as in Algorithm 1. Estimate $\hat{\beta}/\widehat{DN}$.

3 Solve the over-determined linear system of equations in the coefficients of $(\rho/r)DMD_*M_*$ and $(1/r)CNC_*N_*$

$$AA_*\left(\frac{\rho}{r}DMD_*M_*\right) + BB_*\left(\frac{1}{r}CNC_*N_*\right) = \hat{\beta}\hat{\beta}_* \tag{16}$$

with the least squares method.

4 Consider equation (7), multiplied by N:

$$q^{\ell+\tau-m}SB_*\left(\frac{1}{r}CNC_*N_*\right) = \beta_*(NQ_1) + qDNT\left(\frac{1}{r}L_*\right) \quad . \tag{17}$$

Solve it, with respect to $Q_2(q^{-1}) \triangleq N(q^{-1})Q_1(q^{-1})$ and $(1/r)L_*(q)$. Use the known S, B and T, the estimated $\{\hat{\beta}, \widehat{DN}\}$, and $(1/r)CNC_*N_*$ from Step 3.

5 Perform the filtering

$$\hat{d}(k|k+m) = \frac{Q_2(q^{-1})}{T(q^{-1})\hat{\beta}(q^{-1})}x(k+m) \quad .$$

□

For (16) to be solvable, its number of equations must not be smaller than the number of unknown coefficients:

$$2n\beta + 1 \geq 2(nd+nm) + 1 + 2(nc+nn) + 1 \quad .$$

This is equivalent to $n\beta > nd+nm+nc+nn$. Since $n\beta = \max\{nc+nb+nn, nm+na+nd\}$, this condition holds if either of

$$\begin{aligned} nb &> nd+nm \\ na &> nc+nn \end{aligned} \tag{18}$$

holds.

Note, that in Step 4, $\beta_*(z)$ (unstable) and $D(z^{-1})N(z^{-1})T(z^{-1})$ (stable) cannot have any common factors. Compared to (7), the number of equations and the number of unknown coefficients have both been increased by nn, and are thus still equal. Consequently, a unique solution $\{Q_2, (1/r)L_*\}$ exists. Compared to the case of known N, the numbers of unknowns in (16),(17) have increased, compared to (14),(15). Furthermore, the chances to have parameter identifiability are reduced. This is the price to be paid for less *a priori* information.

3.4 Improving the robustness

Two modifications of the Algorithms 1 and 2 are required, in order to obtain a safe behaviour.

- Common factors in the innovations model $\hat{\beta}/\hat{D}$ (or $\hat{\beta}/\widehat{ND}$) should be eliminated by model reduction[17]. *Overparametrization* can thus be handled, and the requirement of known polynomial degrees can be relaxed.
- Rank deficiency in Step 3, caused by (nearly) common factors in $\hat{D}$ and B or $\hat{D}$ and N, must be detected. No such problems can occur in (16) in Algorithm 3, since A and B are known. If the stability of β and the innovations model denominator are monitored during the identification, the solvability of (15) and (17) pose no problems.

With model reduction implemented, we only need to know the *relative* degrees ΔD and ΔN, defined in (11). For reduced models, it is possible to check the conditions (12),(13) for parameter identifiability.

It should be mentioned that β, being an ARMA model numerator, is sometimes rather difficult to estimate. The estimates are often noisy when exponential forgetting is used. The variability of the estimator coefficients (Q_1,$\hat{\beta}$) can be reduced substantially by using low–pass filtered $\hat{\beta}$–polynomial coefficients in the design calculations.

4. A NUMERICAL EXAMPLE

Example 1. Assume the input generation and the measurement system to be given by

$$\frac{C(q^{-1})}{D(q^{-1})} = \frac{1+0.5q^{-1}}{1-0.7q^{-1}} \quad ; \quad q^{-\tau}\frac{B(q^{-1})}{A(q^{-1})} = q^{-1}\frac{1+0.25q^{-1}}{1-0.9q^{-1}}$$

with $M(q^{-1}) = N(q^{-1}) = 1$ and $\lambda_v = 1.563\lambda_e$. An optimal one–lag smoother ($m = 1$) is to be designed. The system is parameter identifiable according to (10)–(13). Assume correct model orders and the identification to give consistent estimates, i.e. $\hat{\beta} = \beta$ from (6) and $\hat{D} = D$. Thus,

$$\begin{aligned} \hat{\beta}(q^{-1}) &= 1 - 0.4314q^{-1} + 0.1739q^{-2} \\ AA_*\hat{D}\hat{D}_* &= 0.63q^2 - 2.608q + 3.9569 - 2.608q^{-1} + 0.63q^{-2} \\ BB_* &= 0.25q + 1.125 + 0.25q^{-1} \quad . \end{aligned}$$

According to Step 3, with $(\rho/r)MM_* = \overline{m}_o$ and $(1/r)CC_* \triangleq \overline{c}_1 q + \overline{c}_o + \overline{c}_1 q^{-1}$, we have to solve (14):

$$\begin{bmatrix} 0.63 & | & 0.25 & 0 & 0 \\ -2.608 & | & 1.125 & 0.25 & 0 \\ 3.9569 & | & 0.25 & 1.125 & 0.25 \\ -2.608 & | & 0 & 0.25 & 1.125 \\ 0.63 & | & 0 & 0 & 0.25 \end{bmatrix} \begin{bmatrix} \overline{m}_o \\ \overline{c}_1 \\ \overline{c}_o \\ \overline{c}_1 \end{bmatrix} = \begin{bmatrix} 0.1739 \\ -0.5064 \\ 1.2164 \\ -0.5064 \\ 0.1739 \end{bmatrix}$$

The least squares solution, in this case exact, is

$$[\overline{m}_o \;\; \overline{c}_1 \;\; \overline{c}_o \;\; \overline{c}_1]^T = [0.2449 \;\; 0.0784 \;\; 0.1959 \;\; 0.0784]^T \; .$$

(From these parameters, $C(q^{-1}) = 1 + 0.5q^{-1}$, $r = \lambda_{\bar{y}}/\lambda_e = 6.38$ and $\rho = \lambda_v/\lambda_e = 1.563$ could be obtained.) As the left–hand side of (15), we obtain

$$q^{1-1}(1/r)CC_*B_*N_* = 0.0196q^2 + 0.1273q + 0.2155 + 0.0784q^{-1} \; .$$

Since $nQ_1 = nc = 1$ and $nL = 1$ according to (8), the polynomial equation in Step 4 is found to be

$$0.0196q^2 + 0.1273q + 0.2155 + 0.0784q^{-1} =$$
$$(0.1739q^2 - 0.4314q + 1)(Q_o + Q_1 q^{-1}) + q(1 - 0.7q^{-1})\left(\frac{1}{r}\right)(\ell_1 q + \ell_o) \; .$$

Multiply both sides by q^{-2} and evaluate for equal-powers of q^{-1}. This gives

$$\begin{bmatrix} 1 & 0 & 0.1739 & 0 \\ -0.7 & 1 & -0.4314 & 0.1739 \\ 0 & -0.7 & 1 & -0.4314 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \ell_1/r \\ \ell_o/r \\ Q_o \\ Q_1 \end{bmatrix} = \begin{bmatrix} 0.0196 \\ 0.1273 \\ 0.2155 \\ 0.0784 \end{bmatrix}$$

The solution is

$$\begin{aligned} Q_1(q^{-1}) &= 0.4322 + 0.0784q^{-1} \\ \frac{1}{r}L_1(q^{-1}) &= 0.2613 - 0.0556q^{-1} \; . \end{aligned}$$

Thus, the optimal one–lag smoothing input estimator (3) is

$$\hat{u}(k|k+1) = \frac{0.4322 - 0.3106q^{-1} - 0.0705q^{-2}}{1 - 0.4314q^{-1} + 0.1739q^{-2}} y(k+1) \; .$$

The corresponding MSE is found to be $Ez(k)^2 = 0.82$. Use of the inverse system as estimator, $Q/R = A/B$, would have resulted in $Ez(k)^2 = 1.56$.

5. SIMULATIONS

Consider Algorithm 1, described in Section 3.1. Three examples will be presented in order to illustrate its behaviour. In all examples described below, one future data was used ($m = 1$).

The initial parameter estimates were $\hat{D} = 1$ and $\hat{\beta} = 1$. The covariance matrix was initialized as a unit matrix.

Example 2: The system is described in Example 1. Correct model orders are assumed and SNR $= 100$. Identification with ELS is used from $k = 0-200$ and RPEM from $k = 201-1000$. A forgetting factor $\lambda(k)$, starting with $\lambda(0) = 0.95$ and increasing to unity, is used.

Example 3: Same as in Example 2, up to $k = 500$. At $k = 500$, $D(q^{-1})$ changes abruptly to $D(q^{-1}) = 1 - 0.9q^{-1}$. The forgetting factor is increasing towards 0.98, from $\lambda(0) = 0.95$.

Example 4: Same as in Example 2. The innovations model is overparametrized. It is assumed that $n\hat{\beta} = 3$ and $n\hat{d} = 2$. The assumed relative degree $n\hat{\beta} - n\hat{d}$ gives $n\hat{c} = 2$. The forgetting factor is increasing towards unity.

Figure 2a,b show the true and estimated input for *Example 2*, $k = 0-200$ and $k = 800-1000$, respectively. In Figure 2c, the corresponding input estimation filter parameters are displayed. Reasonable estimates of the input are obtained already after 100 data. Beyond $k = 800$, the estimator performs very well. This can be verified by comparing the estimated parameters $\hat{Q}_1, \hat{\beta}$ at time $k = 1000$ with the ones derived in Example 1. We have

$$(\beta_1 \;\; \beta_2 \;\; Q_o \;\; Q_1) = (-0.431 \;\; 0.174 \;\; 0.432 \;\; 0.078)$$

and

$$(\hat{\beta}_1 \;\; \hat{\beta}_2 \;\; \hat{Q}_o \;\; \hat{Q}_1)_{k=1000} = (-0.345 \;\; 0.185 \;\; 0.416 \;\; 0.157) \;\; .$$

The estimation error variance $E(u(k) - \hat{u}(k))^2$ was estimated from the data series, for $k = 101 - 1000$. We compared the use of both estimated and true parameters to generate $u(k)$ above. This gave $\widehat{Ez(k)}^2_{est} = 0.96$ and $\widehat{Ez(k)}^2_{true} = 0.85$. This should be compared with the theoretical minimal variance $Ez(k)^2 = 0.82$ from (9).

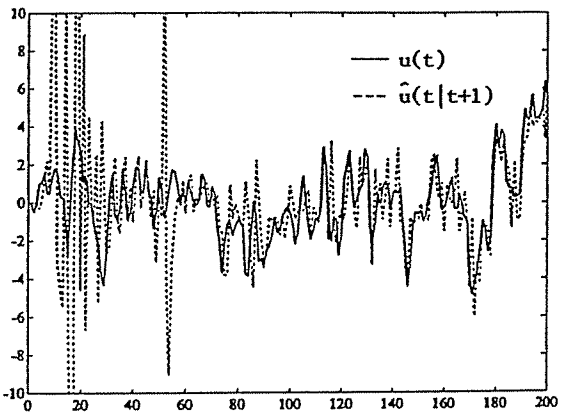

Figure 2a. Example 2. True and estimated input, $k = 0 - 200$. Correct model order, SNR=100, $\lambda(k) \to 1$.

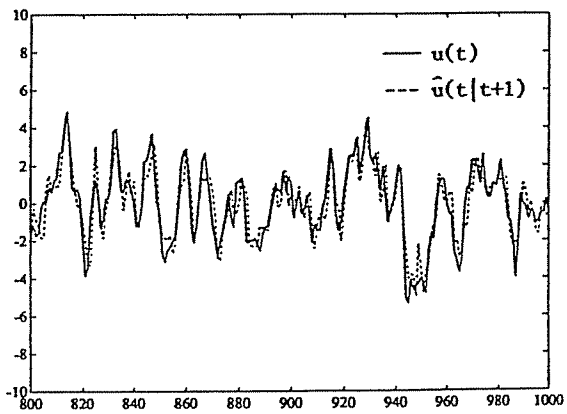

Figure 2b. $k = 800 - 1000$.

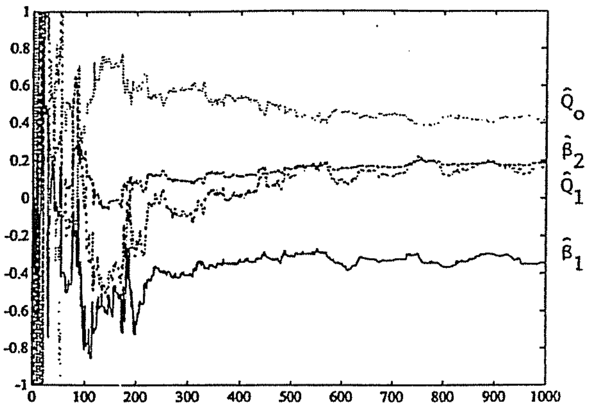

Figure 2c. Example 2. Input estimator parameters $\hat{\beta}_k$, $\hat{Q}_{1k}$, for $k = 0 - 1000$.

In Figure 3a, the true and estimated input for *Example 3* is shown, for $k = 450 - 650$. Figure 3b displays the filter coefficiens $\hat{\beta}_k$ and $\hat{Q}_{1k}$. Regarding Figures 3, we conclude that the input estimator performs well. As can be seen from Figure 3b, the parameter estimates follows the underlying parameter change in D. The $\hat{\beta}$-parameters were slightly low–pass filtered in order to obtain smooth variations in $\hat{Q}_1$. When the system changes, the signal to noise ratio improves from 100 to 590. Although there is not a dramatic improvement in the obtained input estimate, it can be detected. From $k = 550$, there is an improved fit in the peaks.

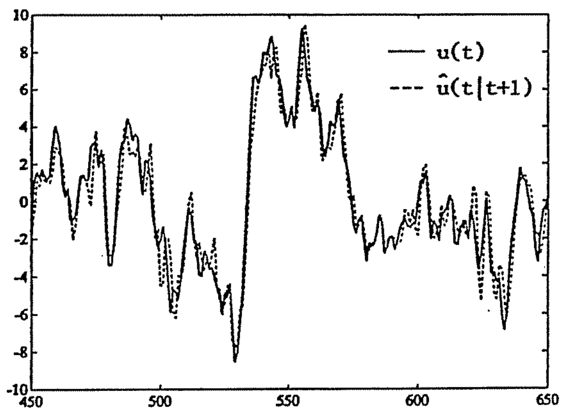

Figure 3a. Example 3. True and estimated input, $k = 450 - 650$. Correct model order assumed. At time $k = 500$, $D(q^{-1})$ changes from $1 - 0.7q^{-1}$ to $1 - 0.9q^{-1}$. $\lambda(k) \to 0.98$.

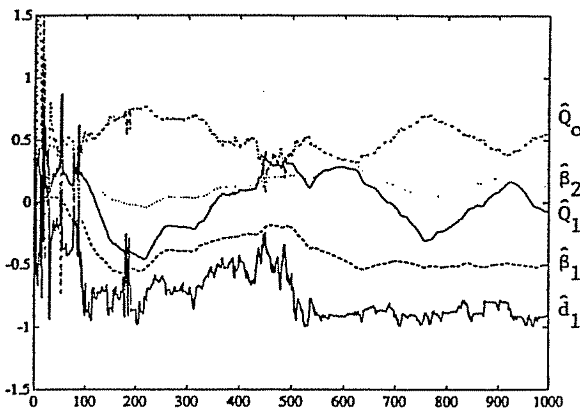

Figure 3b. The coefficients of $\hat{D}_t$ and of the input estimator polynomials $\hat{\beta}_k$, $\hat{Q}_{1k}$, for $k = 0 - 1000$.

Example 4, with Figures 4a and 4b, illustrates the difficulties with overparametrization. In Figure 4a, a typical interval of the simulated data is shown. The result is not very nice and should deter every serious user. However, this may be avoided, as pointed out in Section 3.4. Figure 4b show the trajectories of the superfluous zeros of $\hat{D}$ and $\hat{\beta}$. From $k = 500$, they follow each other closely. There is almost a common factor in the estimated transfer function $\hat{\beta}/\hat{D}$. It can be eliminated by model reduction. The model order is then reduced to that of Example 2, and we can expect a nice behaviour.

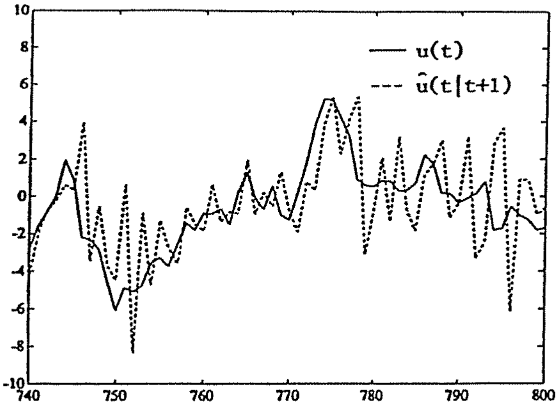

Figure 4a. Example 4. True and estimated input. Overparametrized innovations model, $n\hat{\beta} = 3$, $n\hat{d} = 2$. $k = 740 - 800$.

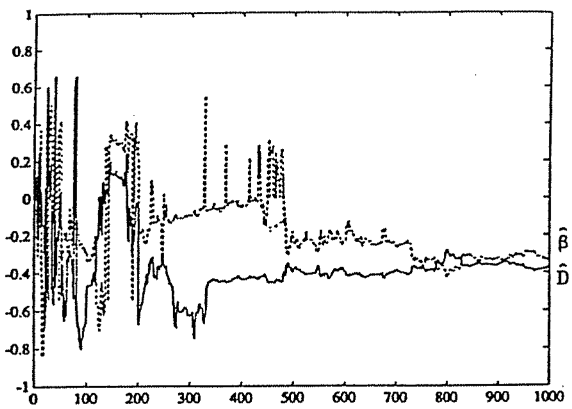

Figure 4b. Trajectories of superfluous zeros of $\hat{\beta}$ and $\hat{D}$.

6. CONCLUSIONS

The deconvolution problem has been considered in an adaptive framework. The input to a known system has been estimated, form noisy measurements of the output. The properties of the input and the measurement noise were assumed unknown. They have been estimated from output data only. We have presented an algorithm based on identification of an innovations model. Identifiability results guarantee a unique decomposition of input description and measurement noise. Two linear systems of equations are the main tools for designing a predictor, a filter or a fixed–lag smoother. The adaptive algorithm was illustrated by a numerical example and simulation experiments. Simulations with correct model order behaved well. An example illustrated difficulties with overparametrization of the innovations model. It was concluded that the difficulties may be avoided by model reduction.

An alternative application of Algorithm 1 is on–line spectral estimation of a signal which is observed through a linear system, with MA–noise. (Off–line estimation, assuming white noise, has been discussed by Tugnait[18].)

The proposed method is not without limitations. Lack of identifiability may prevent its use in some situations. Since only second order statistics is used, non–minimum phase properties cannot be estimated from output data only. This is a limitation in some applications, such as digital channel equalization.

7. REFERENCES

1. Ahlén, A., "Input estimation with application to differentiation." *Ph.D. thesis*, Uppsala University, Sweden, 1986.

2. Ahlén, A., "Identifiability of the deconvolution problem." *Automatica*, vol 26, pp177–181, January 1990.

3. Ahlén, A. and M. Sternad, "Optimal deconvolution based on polynomial methods." *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol ASSP-37, 217-225, February 1989.

4. Ahlén, A. and M. Sternad, "Adaptive input estimation", *ACASP'89*, IFAC Symposium on Adaptive systems in Control and Signal Processing, Glasgow, UK, pp 631–636, April 1989.

5. Ahlén, A. and M. Sternad, "Wiener filter design using polynomial equations". To appear in *IEEE Transactions on Signal Processing*, vol SP–29, November 1991.

6. Carlsson, B., "Digital differentiating filters and model–based fault detection." *Ph.D. thesis* Uppsala University, Sweden, 1989.

7. Carlsson, B., A. Ahlén and M. Sternad, Optimal differentiation based on stochastic signal models. *IEEE Transactions on Signal Processing*, vol SP–39, pp341–353, February 1991.

8. Carlsson, B., M. Sternad and A. Ahlén, "Digital differentiation of noisy data measured through a dynamic system" To appear in *IEEE Transactions on Signal Processing*, vol SP–40, 1992.

9. Commenges, D., "The deconvolution problem: Fast algorithms including the preconditioned conjugate gradient to compute a MAP estimator." *IEEE Transactions on Automatic Control*, vol AC–29, 229-243, March 1984.

10. Demoment, G. and R. Reynaud, "Fast minimum variance deconvolution." *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol ASSP-33, 1324-1326, October 1985.

11. Deng, Z. L., "White–noise filter and smoother with application to seismic data deconvolution." *IFAC Symposium on Identification and System Parameter Estimation*, York, UK, 621-624, July 1985.

12. Golub, G. H. and C. F. Van Loan, *Matrix computations.* North Oxford Academic, Oxford, 1983.

13. Ljung, L. and T. Söderström, *Theory and Practice of Recursive Identification.* MIT Press, Cambridge, Mass, 1983.

14. Mendel, J. M., *Optimal Seismic Deconvolution. An Estimation–based Approach.* Academic Press, New York, 1983.

15. Moir, T.J., T.G. Vishwanath and D. R. Campbell, "Real–time self–tuning deconvolution filter and smoother." *International Journal of Control*, vol 45, 969-985, March 1985.

16. Moir, T.J., D.R. Campbell and H. S. Dabis, "A polynomial approach to optimal and adaptive filtering with application to speech enhancement", *IEEE Transactions on Signal Processing*, vol SP–39, pp 1221-1224, May 1991.

17. Moore, B. C., "Principal component analysis in linear systems: controllability, observability and model reduction." *IEEE Trans. on Aut. Contr.*, vol AC-26, 17-27, 1981.

18. Tugnait, J. K.,"Spectral estimation for noisy signals observed through a linear system", *Proc. IEEE Conference on Decision and Control*, San Antonio, TX, pp 1331–1336, 1983.