Identification of FIR Wiener systems with unknown, non-invertible, polynomial non-linearities

SETH L. LACY†* and DENNIS S. BERNSTEIN‡

Wiener systems consist of a linear dynamic system whose output is measured through a static non-linearity. In this paper we study the identification of single-input single-output Wiener systems with finite impulse response dynamics and polynomial output non-linearities. Using multi-index notation, we solve a least squares problem to estimate products of the coefficients of the non-linearity and the impulse response of the linear system. We then consider four methods for extracting the coefficients of the non-linearity and impulse response: direct algebraic solution, singular value decomposition, multi-dimensional singular value decomposition and prediction error optimization.

1. Introduction


This paper is concerned with identifying Wiener systems under more general assumptions than have been previously considered. Many methods for Wiener system identification require the non-linearity to be known, invertible, differentiable, odd or require specially designed input sequences. In particular, the Wiener identification problem has been considered in Brillinger (1970), Pajunen (1985), Hasiewicz (1987), Greblicki (1992, 1994, 1997, 1998), Westwick and Kearney (1992), Wigren (1994), Westwick and Verhaegen (1996), Bai (1998), and Lovera et al. (2000) under the assumption that the non-linearity is known but one-to-one. This assumption simplifies the problem considerably since the inverse system can be viewed as a Hammerstein system wherein the input to the non-linearity is measured. If the non-linearity is known but not one-to-one, then identification is possible by first generating a candidate set of signals at the output of the linear system (Bayard and Eslami 1984, Lacy et al. 2001). These methods are applicable even if the output non-linearity is a step function in which case the output assumes at most two distinct values (Lacy et al. 2001). If the input sequence can be chosen freely, the frequency content of the input sequence can be selected such that the effect of the non-linearity can be derived from the frequency content of the output (Pintelon and Schoukens 2001).

In the present paper we consider Wiener system identification in which the output non-linearity is both unknown and not necessarily one-to-one. In this case the goal is to simultaneously identify both the linear system dynamics and the non-linearity despite the non-invertibility of the output non-linearity. To do this we assume that the non-linearity can be represented as a finite sum of polynomials. We use multi-index notation (Evans 1998, Dunkl and Xu 2001) to expand this polynomial and write the output as a linear-in-parameters sum of known terms with unknown coefficients. These coefficients consist of products of the system parameters. We then present several methods for extracting the system parameters. This approach of expanding the polynomial output non-linearity requires that the linear system output depend only on past inputs, that is, the linear dynamics are assumed to be FIR.

This paper is organized as follows. In §2 we define the problem and list the assumptions. In §3 we introduce notation used throughout the paper. In §3.1 we present an algebraic solution. In §3.2 we present a solution based on a singular value decomposition. In §3.3 we present a solution based on a multi-dimensional singular value decomposition (Andersson and Bro 2000). In §3.4 we present an optimality approach based on a prediction-error cost function. In §4 we apply all of these methods to an example to illustrate their implementation and compare their effectiveness.
**Title:** Identification of FIR Wiener systems with unknown, non-invertible, polynomial non-linearities

**Authors:** S Lacy; D Bernstein

**Abstract:** Wiener systems consist of a linear dynamic system whose output is measured through a static non-linearity. In this paper we study the identification of single-input single-output Wiener systems with finite impulse response dynamics, and polynomial output non-linearities. Using multi-index notation, we solve a least squares problem to estimate products of the coefficients of the non-linearity and the impulse response of the linear system. We then consider four methods for extracting the coefficients of the non-linearity and impulse response: direct algebraic solution, singular value decomposition, multi-dimensional singular value decomposition and prediction error optimization.
2. Problem description

Here we study the identification of a single-input, single-output, linear time-invariant system with finite impulse response whose output is measured through a static non-linearity. This system, which is represented in figure 1, is modelled by

\[ y(k) = (h_0 u(k) + h_1 u(k-1) + \cdots + h_m u(k-m)) \]

where \( u \) is the input to the system, \( y \) is the unmeasured output of the linear system, \( h_i \) are Markovian parameters, and \( z \) is the measured output of the non-linearity. We assume that the non-linear function \( N: \mathbb{R} \to \mathbb{R} \) is a polynomial of the form

\[ N(y(k)) = \sum_{i=0}^{p} c_i y(k)^i \]  

If \( N \) is not a polynomial, then (3) can be regarded as an approximation. We assume that the order \( m \) of the FIR dynamics and the degree \( p \) of the polynomial \( N \) are known. The non-linearity \( N \) is otherwise unknown and not necessarily one-to-one. In the practical situation \( m \) and \( p \) are not known. Also, it may be difficult to obtain satisfactory results if these parameters are underestimated. However, if upper bounds on these parameters are known, then the bounds can be used in (1) and (3) at the expense of increasing the computational complexity. Alternatively, the values for \( p \) and \( N \) can be incremented until satisfactory performance is achieved, at the expense of increasing the computational load at each increment.

The identification problem is to estimate the coefficients \( h_i \) and \( c_i \) using \( \ell \) measurements of \( u \) and \( z \). We adopt a two-stage approach. First, we solve a least squares problem to obtain an estimate \( \hat{\theta} \) of a vector \( \theta \) whose entries are the unknown parameters and products of the unknown parameters. Next, we present several techniques that use \( \hat{\theta} \) to estimate the individual unknown parameters. We then minimize a prediction error cost function to further refine the parameter estimates and compare to the direct approaches.

3. Wiener identification

Using (3), we rewrite equation (2) as

\[ z(k) = \sum_{i=0}^{p} c_i y(k)^i = \sum_{i=0}^{p} c_i \left( \sum_{j=0}^{m} h_j u(k-j) \right)^i \]

\[ = \sum_{i=0}^{p} c_i \sum_{|\alpha|=i} |\alpha|! h^\alpha v(k)^\alpha = \sum_{|\alpha|=0}^{p} |\alpha|! c_{|\alpha|} h^\alpha \]

\[ = \phi(k)^T \theta \]  

where

\[ h = [h_0, h_1, \ldots, h_m]^T \in \mathbb{R}^{m+1} \]

\[ \alpha = [\alpha_0, \alpha_1, \ldots, \alpha_{m+1}]^T \in \mathbb{N}_{0}^{m+1} \]

\[ v(k) = [u(k), u(k-1), \ldots, u(k-m)]^T \in \mathbb{R}^{m+1} \]

\[ \theta = [c_{|\alpha|} h^\alpha]_{|\alpha| \leq p} \in \mathbb{R}^{D_p} \]

and \( \alpha \in \mathbb{N}_0^{m+1} \) is a multi-index whose order is \( m+1 \), where \( \mathbb{N}_0 \) is the set of positive integers and zero. A multi-index is a vector whose components are non-negative integers (see Evans 1998, Dunkl and Xu 2001). We define

\[ |\alpha| = \alpha_0 + \alpha_1 + \cdots + \alpha_{m+1} = \sum_{i=0}^{m} \alpha_i \]

\[ \alpha! = (\alpha_0!) (\alpha_1!) \cdots (\alpha_{m+1}!) = \prod_{i=0}^{m} \alpha_i! \]

\[ v(k)^\alpha = v_1(k)^{\alpha_0} v_2(k)^{\alpha_1} \cdots v_{m+1}(k)^{\alpha_{m+1}} = \prod_{i=1}^{m+1} v_i(k)^{\alpha_i} \]

\[ h^\alpha = h_1^{\alpha_0} h_2^{\alpha_1} \cdots h_{m+1}^{\alpha_{m+1}} = \prod_{i=1}^{m+1} h_i^{\alpha_i} \]

The number of multi-indices of order \( m+1 \) of fixed absolute value \( i \) is given by

\[ C^{m+1}_i = \binom{m+i}{i} = \binom{m+i}{m} = \frac{(m+i)!}{m! i!} \]  

and the number of multi-indices of order \( m+1 \) of absolute value less than or equal to \( p \) is given by

\[ D^{m+1}_p = \sum_{i=0}^{p} C^{m+1}_i = \sum_{i=0}^{p} \binom{m+i}{m} \]

We need to define an order relation for multi-indices of the same order. Let \( \alpha, \beta \in \mathbb{N}_0^{m+1} \) be multi-indices. If \( |\alpha| > |\beta| \) then \( \alpha > \beta \). If \( \alpha \) and \( \beta \) have the same order,
|\alpha| = |\beta|$, then we choose the standard dictionary ordering. The notation

$$
[f(\alpha)]_{|\alpha| \leq p} = \\
\begin{bmatrix}
[f(\alpha)]_{|\alpha|=0} \\
[f(\alpha)]_{|\alpha|=1} \\
\vdots \\
[f(\alpha)]_{|\alpha|=p}
\end{bmatrix}
$$

(16)

denotes the column vector whose components are $f$ evaluated at every multi-index $\alpha$ such that $|\alpha| \leq p$. The components are ordered according to the above ordering scheme. This vector has $D_p^{m+1}$ components. Thus $\theta$ and $\phi(k)$ have $D_p^{m+1}$ components.

To estimate $\theta$ we rewrite (4) as

$$
z = \Phi^T \theta
$$

(17)

where

$$
z = [z(m + 1) \cdots z(\ell)]^T \in \mathbb{R}^{\ell-m}
$$

(18)

$$
\Phi = [\phi(m + 1) \cdots \phi(\ell)] \in \mathbb{R}^{D_p^{m+1}\times \ell-m}
$$

(19)

We assume $\Phi \Phi^T$ is non-singular, which is a persistency of excitation condition that requires $\ell \geq D_p^{m+1} + m$. Then we calculate the least squares estimate $\hat{\theta}$ of $\theta$ given by

$$
\hat{\theta} = (\Phi \Phi^T)^{-1} \Phi z
$$

(20)

Next we develop several methods for obtaining estimates $\hat{c}$ and $\hat{h}$ based on $\hat{\theta}$. Note that an arbitrary scaling and its reciprocal can be applied to the linear system and the output non-linearity. We remove this ambiguity by introducing a normalization constraint, thereby selecting a single system from a class of equivalent systems. We can normalize $\hat{c}$ and $\hat{h}$ by setting $\hat{c}_i = a$, $\hat{h}_i = a$, $||\hat{c}|| = a$, $||\hat{h}|| = a$ or various other constraints.

3.1. Direct solve

We have $m + p + 2$ unknown parameters in $h$ and $c$, one normalization constraint and $D_p^{m+1}$ equations in terms of $\theta$. We can normalize as above, then choose $m + p + 1$ independent equations. These $m + p + 1$ equations must also be independent of the constraint equation, which constitutes equation number $m + p + 2$. A symbolic manipulator such as Mathematica can be used to invert these non-linear equations and obtain estimates $\hat{h}$ and $\hat{c}$ of $h$ and $c$.

3.2. SVD

To begin, $\hat{c}_0$ can be estimated directly by

$$
\hat{c}_0 = \hat{\theta}(1)
$$

(21)

To estimate the remaining components of $\hat{c}$ and $\hat{h}$, we arrange the components of $\theta$ into the matrix $A(\theta) \in \mathbb{R}^{m+1 \times D_p^{m+1}}$, where

$$
A(\theta) = \psi h^T
$$

(22)

and

$$
\psi = [c_{|\alpha|=1} h^p]_{|\alpha| < p}
$$

(23)

Then we calculate the singular value decomposition

$$
A(\hat{\theta}) = USV^T
$$

(24)

to obtain the estimates

$$
\hat{h} = \sigma S(1, 1) V(1, :)
$$

(25)

$$
\hat{\psi} = \frac{1}{\sigma} U(1, :)
$$

(26)

where the scalar $\sigma \neq 0$ selects the normalization constraint. Finally, we extract the non-linearity coefficients $\hat{c}$ from $\hat{\psi}$. Specifically, $\hat{c}_1$ is given directly by $\hat{\psi}(1)$, while the remaining coefficients are calculated using least squares estimation and $\hat{h}$.

3.3. Multi-dimensional SVD

First, we define the tensors $A_0 \in \mathbb{R}$, $A_j \in \times_{j=1}^l \mathbb{R}^{m+1}$

$$
A_0(\theta) = c_0
$$

(27)

$$
A_1(\theta) = c_1 h = h_{A_1}
$$

(28)

$$
A_2(\theta) = c_2 h \circ h = c_2^2 h_{A_2}
$$

(29)

$$
A_3(\theta) = c_3 h \circ h \circ h = c_3^3 h_{A_3}
$$

(30)

$$
A_p(\theta) = c_p c_{p-1}^p h = c_p^p h_{A_p}
$$

(31)

where

$$
h_{A_n} = c_n^{1/n} h
$$

(32)

We use a multi-dimensional singular value decomposition (Andersson and Bro 2000) to obtain the estimate $\hat{h}_{A_n}$ of $h_{A_n}$. To do this we note that

$$
[h_{A_1} \ h_{A_2} \ \cdots \ \ h_{A_p}] = h d^T
$$

(33)

where

$$
d = [c_1 \ c_2^{1/2} \ \cdots \ c_p^{1/p}]^T
$$

(34)

Hence we compute the singular value decomposition

$$
[h_{A_1} \ \cdots \ h_{A_p}] = USV^T
$$

(35)

to obtain the estimates
\[ \hat{h} = \sigma S(1, 1) U(:, 1) \]  
\[ \hat{d} = \frac{1}{\sigma} V(:, 1) \]  
\[ \hat{\epsilon}_t = \hat{d}_{t}^{\prime} \]  

where \( \sigma \neq 0 \) selects the normalization constraint.

### 3.4. Prediction error cost function

Consider the prediction error cost function

\[ J_{pe}(\hat{c}, \hat{h}) = \| z - \hat{z} \| \]  

where \( \hat{z} \) is the response of the estimated system, that is

\[ \hat{z}(k) = \sum_{i=0}^{p} \hat{c}_i \left( \sum_{j=0}^{m} \hat{h}_j u(k-j) \right) \]

Hence

\[ J_{pe}^2(\hat{c}, \hat{h}) = \sum_{k=m+1}^{\ell} \left( z(k) - \sum_{i=0}^{p} \frac{|\alpha|!}{\alpha!} \hat{c}_i \alpha \hat{h}^{\alpha-e_i} \right)^2 \]  

The derivatives of \( \hat{z}(k) \) are

\[ \frac{\partial \hat{z}(k)}{\partial \hat{c}_i} = \sum_{|\alpha|=i} \frac{|\alpha|!}{\alpha!} v(k)^{\alpha} \hat{h}^{\alpha} = i! \sum_{|\alpha|=i} \hat{h}^{\alpha} v(k)^{\alpha} \]

\[ \frac{\partial \hat{z}(k)}{\partial h_i} = \sum_{|\alpha|=0} \frac{|\alpha|!}{\alpha!} v(k)^{\alpha} \hat{c}_i \alpha \hat{h}^{\alpha-e_i} \]

where \( e_i \) is the \( i \)th column of \( I_{m+1} \). Hence

\[ \frac{\partial J_{pe}^2}{\partial \hat{c}_i} = 2 \sum_{k=m+1}^{\ell} \left( z(k) - \hat{z}(k) \right) \left( - \sum_{|\alpha|=i} \frac{i!}{\alpha!} v(k)^{\alpha} \hat{h}^{\alpha} \right) \]

\[ = -2i! \sum_{|\alpha|=i} \hat{h}^{\alpha} \sum_{k=m+1}^{\ell} v(k)^{\alpha} (z(k) - \hat{z}(k)) \]

and

\[ \frac{\partial J_{pe}^2}{\partial h_i} = 2 \sum_{k=m+1}^{\ell} \left( z(k) - \hat{z}(k) \right) \left( - \sum_{|\alpha|=0} \frac{|\alpha|!}{\alpha!} v(k)^{\alpha} \hat{c}_i \alpha \hat{h}^{\alpha-e_i} \right) \]

\[ = -2 \sum_{|\alpha|=0} \frac{|\alpha|!}{\alpha!} \hat{c}_i \alpha \hat{h}^{\alpha-e_i} \sum_{k=m+1}^{\ell} v(k)^{\alpha} (z(k) - \hat{z}(k)) \]

The second derivatives are given by

\[ \frac{\partial J_{pe}}{\partial c_i \partial c_j} = 2 \sum_{|\alpha|=i} \frac{|\alpha|!}{\alpha!} \hat{h}^{\alpha} v(k)^{\alpha} \]

\[ \frac{\partial J_{pe}}{\partial c_i \partial h_j} = -2 \sum_{|\alpha|=0} \frac{|\alpha|!}{\alpha!} \hat{c}_i \alpha \hat{h}^{\alpha-e_i} \]

\[ \frac{\partial J_{pe}}{\partial h_i \partial h_j} = -2 \sum_{|\alpha|=0} \frac{|\alpha|!}{\alpha!} \hat{c}_i \alpha \hat{h}^{\alpha-e_i} \]

Using the above expressions, we implement a gradient-based optimization algorithm to minimize \( J_{pe} \) and obtain estimates \( \hat{c} \) and \( \hat{h} \). Assuming \( c_1 \neq 0 \), we normalize by letting \( c_1 = 1 \), and thus remove it from the optimization problem.

### 4. Example

Let \( m = 1 \), \( p = 3 \), \( \ell = 2^{13} = 8192 \), \( h = [2 \ 1 \ 1]^T \), \( c = [-20 \ 1 \ 10 \ 1]^T \), and \( N(y) = -20 + y + 10y - y^3 \), which is shown in figure 2. Thus

\[ y(k) = h_0 u(k) + h_1 u(k-1) \]

\[ z(k) = c_0 + c_1 h_1 u(k-1) + c_2 h_0 u(k) + c_2 h_2 u(k-1)^2 \]

\[ + 2c_2 h_0 h_1 u(k)u(k-1) + c_2 h_0 u(k-1)^2 \]

\[ + c_3 h_1 u(k-1)^3 + 3c_3 h_1 h_0 u(k)u(k-1)^2 \]

\[ + 3c_3 h_2 h_1 u(k)^2 u(k-1) + c_3 h_2 u(k)^3 + w(k) \]

\[ = \phi(k)^T \theta + w(k) \]
4.1. Direct solve

We have \( m + p + 2 = 6 \) unknown parameters, one normalization constraint, and \( D_p^{m+1} = 10 \) equations. We choose to normalize by setting \( \hat{c}_1 = 1 \), and solve \( m + p + 1 = 5 \) equations. Using \( \theta(1), \hat{\theta}(2), \hat{\theta}(3), \hat{\theta}(4), \hat{\theta}(7) \) and ignoring the rest of \( \theta \) we obtain

\[
\begin{align*}
c_0 &= \theta(1), \quad \hat{c}_1 = 1, \quad \hat{c}_2 = \frac{\hat{\theta}(4)}{\hat{\theta}(2)^2} \\
\hat{c}_3 &= \frac{\hat{\theta}(7)}{\hat{\theta}(2)^3}, \quad \hat{h}_0 = \theta(2), \quad \hat{h}_1 = \hat{\theta}(3)
\end{align*}
\]

Figure 3(a) shows \( \hat{h} \) for 100 simulations, each having a different realization of the input and noise sequences. Figure 3(b) shows the identified non-linearity.

4.2. SVD

Here we arrange the components of \( \theta \) into a matrix that is also an outer product of \( h \) and \( \psi \). Then we compute the singular value decomposition of this matrix to find \( \hat{h} \) and \( \psi \). Finally, we extract \( \hat{e} \) from \( \hat{\psi} \). First, \( \hat{c}_0 \) can be estimated directly as in (21). To find the remaining parameters, we arrange the components of \( \theta \) as

\[
A(\theta) = \psi h^T = [c | a + i H|^p]^{*} h^T
\]

and

\[
\begin{bmatrix}
c_1 \\
c_2 h_1 \\
c_3 h_0 \\
c_3 h_0
\end{bmatrix}
= \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix} = [0_6, 0_6]
\]

We calculate the singular value decomposition of \( A(\theta) \) as in (24). Then we obtain \( \hat{h} \) and \( \hat{\psi} \) from (25) and (26). Next, we extract the non-linearity coefficients \( c_i \) from \( \hat{\psi} \). \( \hat{c}_1 \) is given directly by \( \hat{\psi}(1) \), and we calculate the remaining \( c_i \) using least squares estimation. In this case the normalization constraint is selected by the choice of the scalar \( \sigma \). We choose to normalize by setting \( \sigma = U(1, 1) \) such that \( \hat{c}_1 = \hat{\psi}(1) = 1 \). In figure 3(c) we plot \( \hat{h} \) for 100 simulations. In figure 3(d) we plot the identified non-linearity.

4.3. Multi-dimensional SVD

We arrange the elements of \( \theta \) into several matrices as in (27)–(31), corresponding to \( c_i \) and the \( i \)th power of \( h \)

\[
\begin{align*}
A_0 &= c_0 = \theta(1) \\
A_1 &= c_1 h = c_1 \begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} \theta(3) \\ \theta(2) \end{bmatrix} = h_{A_1} \\
A_2 &= c_2 h \circ h = c_2 \begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} \theta(6) \\ \theta(5) \end{bmatrix} = h_{A_2} \circ h_{A_2}
\end{align*}
\]
Identification of FIR Wiener systems

Figure 3. Simulation results.

(a) Direct Solve Solutions: $h_1$ vs $h_0$

(b) Direct Solve Solutions: $\mathcal{N}$

(c) SVD Solutions: $h_1$ vs $h_0$

(d) SVD Solutions: $\mathcal{N}$

(e) Multi-Dimensional SVD Solutions: $h_1$ vs $h_0$

(f) Multi-Dimensional SVD Solutions: $\mathcal{N}$

(g) Optimization Solutions: $h_1$ vs $h_0$

(h) Optimization Solutions: $\mathcal{N}$
\[ A_3 = c_3 h \circ h \circ h = h_{A_3} \circ h_{A_3} \circ h_{A_3} \]

\[
= c_3 \begin{bmatrix}
    h_0^3 & h_0^2 h_1 \\
    h_0 h_1 & h_0^2 h_2 \\
    h^2_0 h_1 & h_0 h_2^2 \\
    h_0 h_1^2 & h_1^3
\end{bmatrix}
= \begin{bmatrix}
    \theta(10) & \theta(9) \\
    \theta(9) & \theta(8) \\
    \theta(8) & \theta(7)
\end{bmatrix}
\]

We use a multi-dimensional singular value decomposition (Andersson and Bro 2000) to estimate the scaled Markov vector \( h_{A_3} \). Next, we arrange these results in a matrix and compute the singular value decomposition as in (35) to estimate \( h, d, \) and \( c \). We choose \( \sigma \) to be polynomial and hence the accuracy of the initial estimate usually depends on the accuracy of the initial estimate. We used \texttt{lsqnonlin} in the MATLAB optimization toolbox to minimize the function. In Figure 3(g) we plot \( h \) for 100 simulations. In Figure 3(h) we plot the identified non-linearity. Due to noise effects, the value of the cost function evaluated at the optimal parameters is generally less than the cost function evaluated at the true parameters \( J_{pe}(\hat{c}, \hat{h}) < J_{pe}(c, h) \).

\subsection*{4.4. Prediction error cost function}

We minimize the standard prediction error cost function, \( J_{pe}(\hat{c}, h) = |z - \hat{z}| \), initialized with one of the three algorithms discussed earlier. Using the derivatives given previously we obtain \( \dot{\hat{c}} \) and \( \dot{\hat{h}} \). Although the results of the optimization are generally independent of the method used to obtain the initial estimate, the number of iterations needed by the optimization routine to converge usually depends on the accuracy of the initial estimate. We used \texttt{lsqnonlin} in the MATLAB optimization toolbox to minimize the function. In Figure 3(g) we plot \( h \) for 100 simulations. In Figure 3(h) we plot the identified non-linearity. Due to noise effects, the value of the cost function evaluated at the optimal parameters is generally less than the cost function evaluated at the true parameters \( J_{pe}(\hat{c}, \hat{h}) < J_{pe}(c, h) \).

\subsection*{4.5. Discussion}

The four methods presented, namely direct algebraic solution, singular value decomposition, multi-dimensional singular value decomposition and prediction-error minimization, produced solutions of varying accuracy. Solving some of the equations for the unknown parameters generally produced poor quality estimates, as compared to the other three. Since the direct algebraic solution method uses only a fraction of the entries of \( \theta \) to compute the estimates, it is less robust than the other three methods. In addition, the direct algebraic solution method requires user interaction to select which equations to solve.

The singular value decomposition approach produced estimates that were close to the ones obtained by the prediction error minimization, but relied only on a singular value decomposition and some subsequent least squares steps. This method is the simplest to implement both for the user and numerically.

The multi-dimensional singular value decomposition approach produced estimates comparable to the first singular value decomposition approach, but the multi-dimensional singular value decomposition is more complex to implement than the two-dimensional singular value decomposition approach.

The prediction-error minimization approach is the most complex to implement. For best results, it should be initialized using one of the previous methods. However, it produced the best estimates of both the linear dynamics and the output non-linearity.

\section{Conclusion}

This paper presented four methods for identifying FIR Wiener systems with polynomial non-linearities. We presented three methods for simultaneous direct estimation of the non-linearity and linear dynamics, and a prediction error optimization method.

Many authors have studied Wiener system identification under the assumption that the non-linearity is unknown but one-to-one (Brillinger 1970, Pajunen 1985, Hasiewicz 1987, Greblicki 1992, 1994, 1997, Westwick and Kearney 1992, Wigren 1994, Westwick and Verhaegen 1996, Bai 1998, Lovera et al. 2000). Other methods for Wiener system identification require the non-linearity to be known, invertible, monotonic, odd, even, or require the use of specially designed input sequences. In this paper we require the non-linearity to be polynomial and the linear dynamics to have finite impulse response. These two assumptions are practical in that many non-linearities can be approximated with polynomials, and that many systems with infinite impulse response can be approximated with finite impulse response dynamics.

Future work will focus on extending the method in three directions: first, the identification of sandwich non-linear systems, i.e. systems with both input and output non-linearities; second, the identification of IIR Wiener systems; and finally, identification of Wiener systems with non-polynomial output non-linearities. While the identification of sandwich non-linear systems using this approach seems tractable, overcoming the FIR and polynomial assumptions on the linear dynamics and output non-linearity appears to be more challenging.

\section*{Acknowledgements}

Supported in part by NASA under GSRP NGT4-52421 and AFOSR under grant F49620-01-1-0094 and laboratory research initiative 00VS17COR.
References


