Research Seminar

Subspace Identification of Hammerstein and Wiener Models

Speaker

Juan C. Gómez
Laboratory for System Dynamics and Signal Processing
FCEIA, Universidad Nacional de Rosario
ARGENTINA
jcgomez@fcea.unr.edu.ar

Outline

- Introduction: Motivation, New results
- A (very) brief review on Subspace State-Space System IDentification Methods
- Block-oriented Nonlinear Models
- Subspace Identification of Hammerstein Models
- Subspace Identification of Wiener Models
- Simulation Examples
- Conclusions
Motivation for Nonlinear (Subspace) Identification

- Most physical processes have a nonlinear behaviour, except in a limited range where they can be considered linear.
- The performance of controllers designed from a linear approximation is strongly influenced by a change in the operating point of the system.
- Nonlinear models are able to describe more accurately the global behaviour of the system, independently of the operating point.
- Many dynamical systems can be represented by the interconnection of static nonlinearities and LTI systems. These models are called block-oriented nonlinear models.

Subspace Methods have been very successful for the identification of LTI models in many practical applications.

Although there is a well developed theory for Subspace Identification methods for LTI systems, this is not the case for nonlinear systems. Some recent contributions in this area are: (Verhaegen & Westwick, 1996) in Subspace Identification of Hammersterin and Wiener models, and (Chen & Maciejowski, 2000) and (Favoreel et al., 1999) in Subspace Identification of bilinear systems.
The new results (Gomez & Baeyens, 2005)

- New subspace algorithms for the simultaneous identification of the linear and nonlinear parts of multivariable Hammerstein and Wiener models are presented.
- The proposed algorithms consist basically of two steps:
  **Step 1:** a standard (linear) subspace algorithm applied to an equivalent linear system whose inputs (outputs) are filtered (by the basis functions describing the static nonlinearities) versions of the original inputs (outputs).
  **Step 2:** a 2-norm minimization problem which is solved via an SVD.
- Provided the conditions for the consistency of the linear subspace algorithm used in Step 1 are satisfied, consistency of the estimates can be guaranteed.

References


They combine tools of System Theory, Numerical Linear Algebra and Geometry (projections).

They have their origin in Realization Theory as developed in the 60/70s (Ho & Kalman, 1966).

They provide reliable state-space models of multivariable LTI systems directly from input-output data.

They don’t require iterative optimization procedures → no problems with local minima, convergence and initialization.

Properties

- They don’t require a particular (canonical) state-space realization → numerical conditioning improves.
- They require a modest computational load in comparison to traditional identification methods like PEM.
- The algorithms can be (they have been) efficiently implemented in software like Matlab.
- Main computational tools are QR and SVD.
- All subspace methods compute at some stage the subspace spanned by the columns of the extended observability matrix.
- The various algorithms (e.g., N4SID, MOESP, CVA) differ in the way the extended observability matrix is estimated and also in the way it is used to compute the system matrices.
The system model

\[ x_{k+1} = Ax_k + Bu_k + Ke_k \]
\[ y_k = Cx_k + Du_k + e_k \]

State-space model in innovation form

The identification problem

To estimate the system matrices \((A, B, C, D)\) and \(K\), and the model order \(n\), from an \((N+\alpha-1)\)-point data set of input and output measurements

\[ \{u_k, y_k\}_{k=1}^{N+\alpha-1} \]

Realization-based 4SID Methods

For a LTI system, a minimal state-space realization \((A, B, C, D)\) completely defines the input-output properties of the system through

\[ y_k = \sum_{\ell=0}^{\infty} h_\ell u_{k-\ell} \]

convolution sum

where the impulse response coefficients \(h_\ell\) are related to the system matrices by

\[
\begin{cases}
D, & \ell = 0 \\
CA^{\ell-1}B, & \ell > 0
\end{cases}
\]
An estimate of the extended observability matrix can be computed by a **full rank** factorization of the impulse response Hankel matrix. This factorization is provided by the SVD of matrix $H_{ij}$.

In the absence of noise, $H_{ij}$ will be a rank $n$ matrix, and $\Sigma_i$ will contain the $n$ non-zero singular values → **model order is computed**. In the presence of noise, $H_{ij}$ will have full rank and a rank reduction stage will be required for the model order determination.

**Problems:** it is necessary to measure or to estimate (for example, via correlation analysis) the impulse response of the system → **not good**
Direct 4SID Methods

\[ Y_\alpha = \Gamma_\alpha X + H_\alpha U_\alpha + N_\alpha \]  
\text{fundamental equation} \quad (1)

\[ Y_\alpha = \begin{bmatrix}
    y_1 & y_2 & \cdots & y_N \\
    y_2 & y_3 & \cdots & y_{N+1} \\
    \vdots & \vdots & \ddots & \vdots \\
    y_{\alpha-1} & y_\alpha & \cdots & y_{N+\alpha-1}
\end{bmatrix} \]

Output block Hankel matrix

(In a similar way are defined the Input block Hankel matrix \( U_\alpha \) and the Noise block Hankel matrix \( N_\alpha \).)

\[ \Gamma_\alpha = \begin{bmatrix}
    C \\
    CA \\
    \vdots \\
    CA^{\alpha-1}
\end{bmatrix} \quad \text{Extended } (\alpha > n) \quad \text{Observability Matrix}

\[ X = [x_1, x_2, \cdots, x_N] \]

State Sequence Matrix

\[ H_\alpha = \begin{bmatrix}
    D & 0 & 0 & \cdots & 0 \\
    CB & D & 0 & \cdots & 0 \\
    CAB & CB & D & \cdots & 0 \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    CA^{\alpha-2}B & CA^{\alpha-3}B & CA^{\alpha-4}B & \cdots & D
\end{bmatrix} \quad \text{Lower triangular block Toeplitz matrix of impulse responses (unknown).}

The main idea of Direct 4SID methods

In the absence of noise \((N_\alpha = 0)\), eq. (1) becomes

\[ Y_\alpha = \Gamma_\alpha X + H_\alpha U_\alpha \] \quad (2)

and the part of the output which does not emanate from the state can be removed by multiplying (from the right) both sides of eq. (2) by the orthogonal projection onto the null space of \( U_\alpha \), \textit{i.e.} by
\[ \Pi_{\alpha}^\perp = I - U_{\alpha}^T \left( U_{\alpha} U_{\alpha}^T \right)^{-1} U_{\alpha}^\perp \quad \text{orthogonal projection} \]

such that \[ U_{\alpha} U_{\alpha}^\perp = I \]

This yields

\[ Y_\alpha U_{\alpha}^\perp = \Gamma_\alpha X U_{\alpha}^\perp \]  

(3)

**Note** that the matrix on the left depends exclusively on the input-output data. Then, a **full rank factorization** of this matrix will provide an estimate \( \Gamma_\alpha \) of the extended observability matrix. Estimates of the corresponding system matrices can be obtained by resorting to the **shift invariance property** of the extended observability matrix, and by solving a system of linear equations in the least squares sense.

The factorization is provided by the SVD of the matrix on the left side

\[ Y_\alpha U_{\alpha}^\perp = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \approx U_1 \Sigma_1 V_1^T = \left( U_1 \Sigma_1^{1/2} \right) \left( \Sigma_1^{1/2} V_1^T \right) \]

(4)

**Weighting Matrices**

Row and column weighting matrices can be introduced in (4) before performing the SVD of the matrix in the left hand side. Any choice of positive-definite weighting matrices \( W_r \) and \( W_c \) will result in consistent estimates of the extended observability matrix.
Existing algorithms employ the following choices for matrices $W_r$ and $W_c$,

- **MOESP** (Verhaegen, 1994): $W_r = I$, $W_c = \left( \frac{1}{N} \Phi \Pi_{U_{\alpha}^T}^\perp \Phi^T \right)^{-1} \Phi \Pi_{U_{\alpha}^T}^\perp$

- **CVA** (Larimore, 1990): $W_r = \left( \frac{1}{N} Y_{\alpha} \Pi_{U_{\alpha}^T}^\perp Y_{\alpha}^T \right)^{-\frac{1}{2}}$, $W_c = \left( \frac{1}{N} \Phi \Pi_{U_{\alpha}^T}^\perp \Phi^T \right)^{-\frac{1}{2}}$

- **N4SID** (Van Overschee and de Moor, 1994):

\[
W_r = I, \quad W_c = \left( \frac{1}{N} \Phi \Pi_{U_{\alpha}^T}^\perp \Phi^T \right)^{-1} \Phi
\]

---

**Computation of the system matrices**

Given an estimate $\hat{\Gamma}_{\alpha}$ of the extended observability matrix, estimates of the system matrices can be computed as:

- $\hat{C}$ : first row block of $\hat{\Gamma}_{\alpha}$
- $\hat{A}$ : solving in the least squares sense

\[
\hat{\Gamma}_{\alpha} = \Gamma_{\alpha} \hat{A}
\]

**shift-invariance property**

- $\hat{B}$ and $\hat{D}$ : solving a system of linear equations
Presence of noise

In the presence of noise

\[ Y_\alpha = \Gamma_\alpha X + H_\alpha U_\alpha + N_\alpha \]

and

\[ Y_\alpha U_\alpha^\perp = \Gamma_\alpha X + N_\alpha U_\alpha^\perp \]

The noise term can be removed by correlating it away with a suitable matrix. This can be interpreted as an oblique projection.
**Hammerstein Model Identification**

**Problem Formulation**

\[
\begin{align*}
    x_{k+1} &= Ax_k + Bv_k + \omega_k \\
    y_k &= Cx_k + Dv_k + \nu_k
\end{align*}
\]

(1) \hspace{1cm} (2)

- \(y_k \in \mathbb{R}^m\), \(x_k \in \mathbb{R}^n\), \(v_k \in \mathbb{R}^p\)
- \(\omega_k \in \mathbb{R}^n\), \(\nu_k \in \mathbb{R}^m\)

**Nonlinear subsystem**

\[v_k = N(u_k) = \sum_{i=1}^{r} \alpha_i g_i(u_k)\]

(3)

- \(g_i(\bullet) : \mathbb{R}^p \rightarrow \mathbb{R}^p, (i = 1, \ldots, r)\)
- Known basis functions
- Unknown matrix parameters \(\alpha_i \in \mathbb{R}^{p \times p}\)

**Identification problem:** to estimate the unknown parameter matrices \(\alpha_i \in \mathbb{R}^{p \times p}, (i = 1, \ldots, r)\), and \(A, B, C, \) and \(D\) characterizing the nonlinear and the linear parts, respectively, and the model order \(n\), from an \(N\)-point data set \(\{u_k, y_k\}_{k=1}^{N}\) of observed input-output measurements.

**Subspace Identification Algorithm**

\[
\begin{align*}
    x_{k+1} &= Ax_k + \sum_{i=1}^{r} B\alpha_i g_i(u_k) + \omega_k \\
    y_k &= Cx_k + \sum_{i=1}^{r} D\alpha_i g_i(u_k) + \nu_k
\end{align*}
\]

(3) \rightarrow (1), (2) \Rightarrow

**Normalization** \(\|\alpha_i\|_2 = 1\)

LSD - September 2005

J.C. Gómez
Defining $\tilde{B} = [B \alpha_1, \cdots, B \alpha_r]$, $\tilde{D} = [D \alpha_1, \cdots, D \alpha_r]$, $U_k = [g_1^T(u_k), \cdots, g_r^T(u_k)]^T$

\[
\begin{cases}
x_{k+1} = Ax_k + \tilde{B}U_k + \omega_k \\
y_k = Cx_k + \tilde{D}U_k + \nu_k
\end{cases}
\]

Fig. 6: Equivalent LTI system with input $U_k$

**Linear Subspace Algorithms** (N4SID, MOESP,CVA)

**Estimates** $\hat{A}, \hat{B}, \hat{C}, \hat{D}$, model order $n$

Defining $\alpha = [\alpha_1, \cdots, \alpha_r]^T$, then $\tilde{B} = B \alpha^T$, and $\tilde{D} = D \alpha^T$, so that

$$\Theta_{BD} = \begin{bmatrix} \tilde{B} \\ \tilde{D} \end{bmatrix} = \begin{bmatrix} B \\ D \end{bmatrix} \alpha^T$$

The **problem** then is how to compute estimates of matrices $B$, $D$, and $\alpha$ from the estimate of the matrices $\tilde{B}$, and $\tilde{D}$ (i.e., from an estimate of $\Theta_{BD}$)

It is clear that the closest, in the 2-norm sense, estimates $\hat{B}$, $\hat{D}$, and $\hat{\alpha}$ are such that

$$\left(\hat{B}, \hat{D}, \hat{\alpha}\right) = \arg\min_{B, D, \alpha} \left\{ \left\| \Theta_{BD} - \begin{bmatrix} B \\ D \end{bmatrix} \alpha^T \right\|_2^2 \right\}$$

The **solution** to this optimization problem is provided by the SVD of $\hat{\Theta}_{BD}$. 

LSD - September 2005

J.C. Gómez
Let \( \hat{\Theta}_{BD} \in \mathbb{R}^{(n+m)\times rp} \) have rank \( s>p \), and let its economy size SVD be partitioned as
\[
\hat{\Theta}_{BD} = U_1 \Sigma V_1^T = \sum_{i=1}^{s} \sigma_i u_i v_i^T = [U_1 \ U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}
\]
(4)
with \( U_1 \in \mathbb{R}^{(n+m)\times p} \), \( V_1 \in \mathbb{R}^{rp \times p} \), and \( \Sigma_1 = \text{diag} (\sigma_1, \sigma_2, \ldots, \sigma_p) \).

Then
\[
\begin{bmatrix} \hat{\Theta}_{BD} \\ \hat{\alpha} \end{bmatrix} = \arg\min_{B,D,\alpha} \left\| \begin{bmatrix} \hat{\Theta}_{BD} \\ \hat{\alpha} \end{bmatrix} - \begin{bmatrix} B \\ D \end{bmatrix} \alpha^T \right\|_2^2 = (U_1 \Sigma_1, V_1),
\]
and the approximation error is given by
\[
\left\| \hat{\Theta}_{BD} - \begin{bmatrix} \hat{B} \\ \hat{D} \end{bmatrix} \alpha^T \right\|_2^2 = \sigma_{p+1}^2.
\]

**Normalization in \( \alpha \) provided by the SVD**

---

**Identification Algorithm**

The subspace algorithm can be summarized as follows.

**Step 1:** Compute estimates of the system matrices \((A, \tilde{B}, C, \tilde{D})\), and the model order \(n\), using any available (linear) subspace algorithm, such as N4SID, MOESP, CVA.

**Step 2:** Based on the estimates \( \hat{B} \) and \( \hat{D} \) compute an estimate \( \hat{\Theta}_{BD} \) of matrix \( \Theta_{BD} \).

**Step 3:** Compute the SVD of \( \hat{\Theta}_{BD} \) and its partition as in (4).

**Step 4:** Compute the estimates of the parameter matrices \( B, D, \) and \( \alpha \) as
\[
\begin{bmatrix} \hat{B} \\ \hat{D} \end{bmatrix} = U_1 \Sigma_1 \quad \hat{\alpha} = V_1
\]
respectively.
**Result 2: Consistency Analysis**

Under some assumptions on **persistency of excitation** of the inputs, which depend on the particular subspace method used in **Step 1** of the algorithm, the estimates \( \left( \hat{A}, \hat{B}, \hat{C}, \hat{D} \right) \) are **consistent** in the sense that they converge to the true values when the number of data points \( N \to \infty \).

The consistency of \( \hat{B} \) and \( \hat{D} \), implies that of \( B, D, \) and \( \alpha \).

---

**Wiener Model Identification**

**Problem Formulation**

\[
\begin{align*}
{x_{k+1} &= Ax_k + Bu_k + \omega_k} \\
{v_k &= Cx_k + Du_k + u_k}
\end{align*}
\]

Nonlinear subsystem

\[
v_k = N^{-1}(y_k) = \sum_{i=1}^{r} \alpha_i g_i(y_k)
\]

LTI subsystem

\[
\begin{align*}
{u_k \in \mathbb{R}^p, & \quad x_k \in \mathbb{R}^n, & \quad v_k \in \mathbb{R}^m} \\
{\omega_k \in \mathbb{R}^n, & \quad \nu_k \in \mathbb{R}^m}
\end{align*}
\]

known basis functions

unknown matrix parameters

**Fig. 7: Wiener model**

\[\text{LSD - September 2005} \quad J.C. Gómez\]

\[\text{LSD - September 2005} \quad J.C. Gómez\]
Identification problem: to estimate the unknown parameter matrices \( \alpha_i \in \mathbb{R}^{m \times m}, (i = 1, \ldots, r) \), and \( A, B, C, \) and \( D \) characterizing the nonlinear and the linear parts, respectively, and the model order \( n \), from an \( N \)-point data set \( \{u_k, y_k\}_{k=1}^{N} \) of observed input-output measurements.

### Subspace Identification Algorithm

\[
\begin{align*}
\alpha & = [\alpha_1, \ldots, \alpha_r] , \ Y_k = \left[ g_1^T(y_k), \ldots, g_r^T(y_k) \right]^T \\
\alpha^+ & = \alpha, \ \tilde{C} = \alpha^+ C, \ \tilde{D} = \alpha^+ D
\end{align*}
\]

Normalization \( \|\alpha^+\|_2 = 1 \)

### Identifiability problem

Fig. 8: Equivalent LTI model with output \( Y_k \)

Linear Subspace Algorithms

(N4SID, MOESP, CVA)

Estimates \( \hat{A}, \hat{B}, \hat{C}, \hat{D}, \) model order \( n \)
The problem is how to compute estimates of matrices $C$, $D$, and $\alpha^+$ from the estimates of the matrices $\tilde{C}$, and $\tilde{D}$.

Similarly to what was done for the Hammerstein model the closest, in the 2-norm sense, estimates $\hat{C}$, $\hat{D}$, and $\hat{\alpha}^+$ are such that

$$
(\hat{C}, \hat{D}, \hat{\alpha}^+) = \arg\min_{C, D, \alpha^+} \left\{ \| \hat{C} \hat{D} - \alpha^+ [C \ D] \|^2 \right\}
$$

The solution to this optimization problem is provided by the SVD of the matrix $[\hat{C} \ \hat{D}]$.

Result 3

Let $[\hat{C} \ \hat{D}] \in \mathbb{R}^{m \times (n+p)}$ have rank $s > m$, and let its economy size SVD be partitioned as

$$
[\hat{C} \ \hat{D}] = U \Sigma V^T = \sum_{i=1}^{s} \sigma_i u_i v_i^T = [U_1 \ U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}
$$

with $U_1 \in \mathbb{R}^{m \times m}$, $V_1 \in \mathbb{R}^{(n+p) \times m}$, and $\Sigma_1 = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_m)$. Then

$$
(\hat{\alpha}^+, [\hat{C} \ \hat{D}]) = \arg\min_{C, D, \alpha^+} \left\{ \| \hat{C} \hat{D} - \alpha^+ [C \ D] \|^2 \right\} = (U_1, \Sigma_1 V_1^T),
$$

and the approximation error is given by

$$
\| \hat{C} \hat{D} - \hat{\alpha}^+ [\hat{C} \ \hat{D}] \|^2 = \sigma_{m+1}^2.
$$

Normalization in $\alpha^+$ provided by the SVD.
Identification Algorithm

The subspace algorithm can be summarized as follows.

**Step 1:** Compute estimates of the system matrices \((A, B, \tilde{C}, \tilde{D})\), and the model order \(n\), using any available (linear) subspace algorithm, such as N4SID, MOESP, CVA.

**Step 2:** Compute the SVD of \(\begin{bmatrix} \hat{C} & \hat{D} \end{bmatrix}\) and its partition as in (8).

**Step 3:** Compute the estimates of the parameter matrices \(C, D,\) and \(\alpha^+\) as

\[
\begin{bmatrix} \hat{C} & \hat{D} \end{bmatrix} = \Sigma_1 V_1^T
\]

\[
\hat{\alpha} = U_1^+
\]

respectively.

Simulation Examples

**Example 1: Hammerstein Model ID ("academic")**

- **The True System**
  \[
  G(z) = \frac{z^2 + 0.7z - 1.5}{z^3 + 0.9z^2 + 0.15z + 0.002}
  \]

  linear subsystem

  \[N(u_k) = 0.8589 u_k + 0.0149 u_k^2 - 0.5113 u_k^3 - 0.0263 u_k^4\]

  nonlinear subsystem

- **The input and noise**
  \[
  u_k = \sin(0.0005\pi k) + 0.5 \sin(0.0015\pi k) + 0.3 \sin(0.0025\pi k) + 0.1 \sin(0.0035\pi k) + \gamma_k
  \]
  \[
  \Phi_v(\omega) = \frac{0.64 \times 10^{-8}}{1.2 - 0.4 \cos(\omega)}
  \]
  input (\(\gamma_k\) white noise with variance \(10^{-6}\))

  Spectrum of the zero mean coloured noise
The Estimated Nonlinear Subsystem

\[ \hat{N}(u_k) = 0.8589 u_k + 0.0142 u_k^2 - 0.5113 u_k^3 - 0.0260 u_k^4 \]

Fig. 9: True (blue) and Estimated (green) nonlinear characteristic.

The Estimated Linear Subsystem

\[ \hat{G}(z) = \frac{0.9986 z^2 + 0.6997 z - 1.4984}{z^3 + 0.9002 z^2 + 0.1495 z + 0.0014} \]

Estimated linear subsystem

Validation results

Fig. 10: True (green) and Estimated (blue) Output.
Example 2: Hammerstein Model ID (Binary Distillation Column)

**Input:** reflux ratio ($u$)

**Outputs:**
- overhead flow rate ($y_1$)
- overhead methanol concentration ($y_2$)
- bottom flow rate ($y_3$)
- bottom methanol concentration ($y_4$)

Fig. 11: Schematic representation of the distillation column

(Weischedel & McAvoy, 1980)

Fig. 12: Left Plot: Estimation (first 1000 points), and validation (remaining 1000 points) Input Data. Right Plot: Estimation (first 1000 points) and Validation (remaining 1000 points) Output Data.
Fig. 13: True (blue) and Estimated (red) Outputs (validation data)

The Estimated Linear Subsystem

Third order model with eigenvalues at \( \{0.4916, 0.9557, 0.9726\} \)

The Estimated Nonlinear Subsystem

Third order polynomial

Fig. 14: Estimated Nonlinear Characteristic
Example 3: Wiener Model ID (pH Neutralization Process)

Fig. 15: Schematic representation of the pH Neutralization Process
(Henson & Seborg, 92, 94, 97)

- **Base:** NaOH  **Acid:** HNO₃
- **Buffer:** NaHCO₃
- **Manipulated variable:** base flow rate ($u_1$)
- **Disturbances:** buffer flow rate ($u_2$) and acid flow rate ($u_3$)
- **Output:** pH of the effluent solution ($y$)

Simulation Model based on first principles (introducing two reaction invariants for each inlet stream)

$$
\dot{x} = f(x) + g(x)u_1 + p(x)u_2 \\
h(x, y) = 0
$$

where

$$
x^\Delta = [x_1, x_2]^T = [W_a, W_b]^T
$$

$$
f(x) = \left[ \frac{u_3}{V} (W_{a3} - x_1), \frac{u_3}{V} (W_{b3} - x_2) \right]^T
$$

$$
g(x) = \left[ \frac{1}{V} (W_{a1} - x_1), \frac{1}{V} (W_{b1} - x_2) \right]^T
$$

$$
p(x) = \left[ \frac{1}{V} (W_{a2} - x_1), \frac{1}{V} (W_{b2} - x_2) \right]^T
$$

$$
h(x, y) = x_1 + 10^{y-14} - 10^{-y} + x_2 \left( \frac{1 + 2 \times 10^{y-pK_2}}{1 + 10^{pK_1-y} + 10^{y-pK_2}} \right)
$$
**Estimation and Validation data**

**Fig. 16:** Estimation (first 1000 points) and validation (remaining 600 points) input-output data.

The Estimated Linear Subsystem

Third order model

\[ \hat{G}(z) = \frac{0.0062z^2 - 0.0122z + 0.006}{z^3 - 2.9466z^2 + 2.8940z - 0.9474} \]

The Estimated Nonlinear Subsystem

Third order polynomial

\[ \hat{N}^{-1}(y_k) = 0.0319y_k^3 + 0.0358y_k^2 + 0.9989y_k \]

**Fig. 17:** Estimated Nonlinear Characteristic.
Validation results

Fig. 18: True (blue) and estimated (red) Output (Estimation/Validation data).

Conclusions

- New subspace methods for the simultaneous identification of the linear and nonlinear parts of multivariable Hammerstein and Wiener models have been presented.
- The proposed methods make use of a standard (linear) subspace method followed by a 2-norm minimization problem which is solved via an SVD.
- The proposed methods generalize all the families of linear subspace methods to this class of nonlinear models.
- The method provides consistent estimates under the same conditions on persistency of excitation required by the (linear) subspace method used as the first step of the algorithm.
- The estimated models are in a format which is suitable for their use in standard (linear) Model Predictive Control schemes.