Chapter VII

Practical Aspects of System Identification

Version 1/4.12.2005

Chapter 7. Practical Aspects of System Identification

7.1 Input/Output Data Acquisition
- **7.1.1 Acquisition Protocol**
- **7.1.2 Anti-Aliasing Filtering**
- **7.1.3 Over Sampling**

7.2 Signal Conditioning
- **7.2.1 Elimination of the DC Component**
- **7.2.2 Identification of a Plant Containing a Pure Integrator**
- **7.2.3 Identification of a Plant Containing a Pure Differentiator**
- **7.2.4 Scaling of the Inputs and Outputs**

7.3 Selection (Estimation) of the Model Complexity

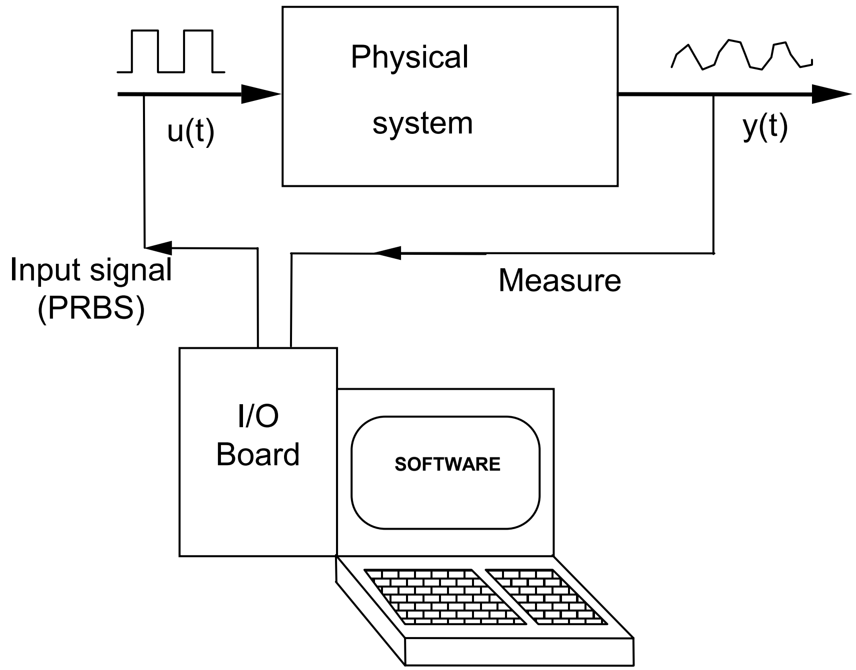
7.4 Identification of Simulated Models: Examples

7.5 Plant Identification Examples
- **7.5.1 Air Heater**
- **7.5.2 Distillation Column**
- **7.5.3 DC Motor**
- **7.5.4 Flexible Transmission**

7.6 Concluding Remarks

Input/Output Data Acquisition

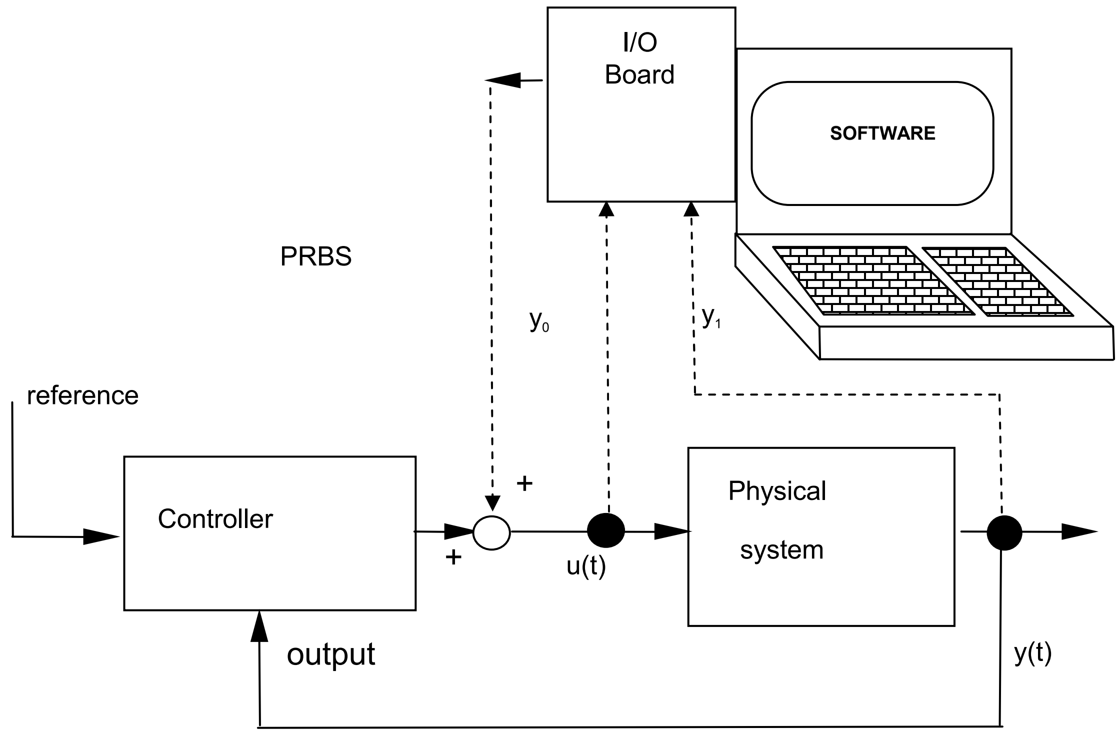
Plant operated in open loop



Plant directly excited by a PRBS signal

Input/Output Data Acquisition

Plant operated in closed loop with excitation added to the controller output

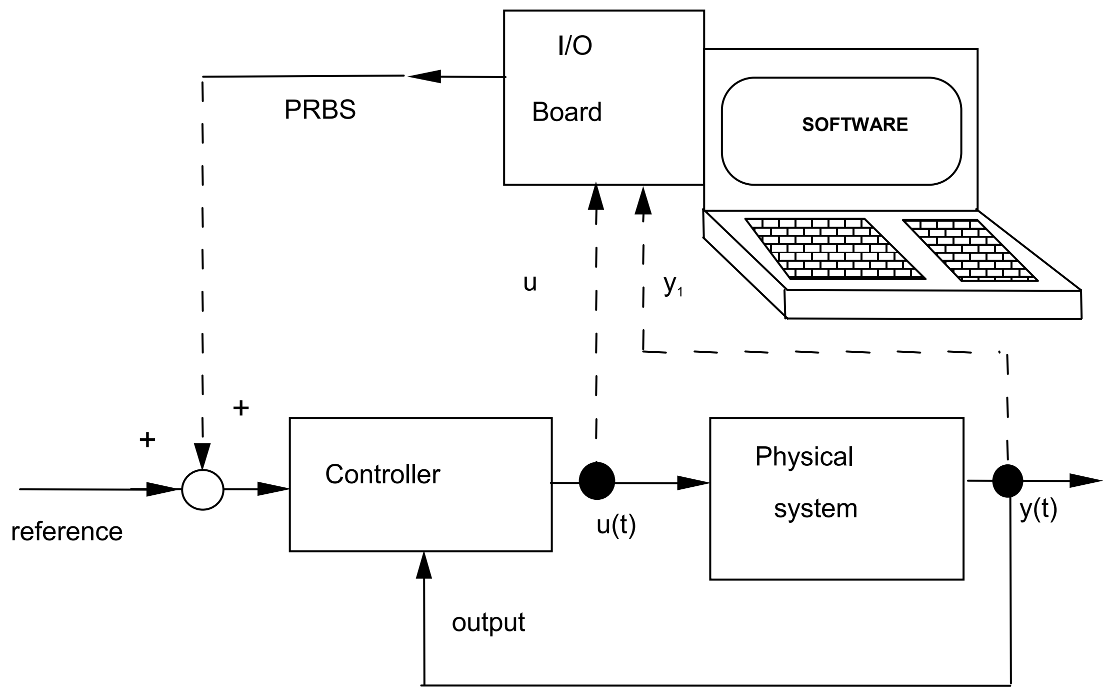


- PRBS superposed to the controller output
- Identification of the transfer function between y_0 and y

Use of a controller with integral action but limited proportional and derivative actions

Input/Output Data Acquisition

Plant operated in closed loop with excitation added to the reference



- PRBS superposed to the reference
- Identification of the transfer function between y_0 and y

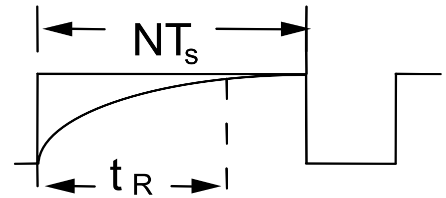
Use of a controller with integral action but limited proportional and derivative actions

Sizing of the PRBS signal

Good identification of the static gain:

Approach 1 (choice of N)

$$t_{im} = T_s.N > t_R$$



But on the other hand : $(2^N - 1)T_s = L <$ experiment length

$N \rightarrow NT_s \geq t_R$; $(2^N - 1)T_s \leq$ experiment length

Approach 2 (choice of N and of the frequency divider f_{PRBS})

$$f_{PRBS} = \frac{f_s}{p} \quad ; \quad p = 1,2,3,\ldots$$

$N, p \rightarrow pNT_s > t_R$; $p(2^N - 1)T_s \leq$ experiment length

Magnitude : **0.5%(0.15%) to 10%(5%) of the input steady state value**

Anti-aliasing filtering

- useful frequencies until 0.5 f_s
- frequencies beyond 0.5f_s modify the spectrum of the sampled signal (spectrum overlapping)

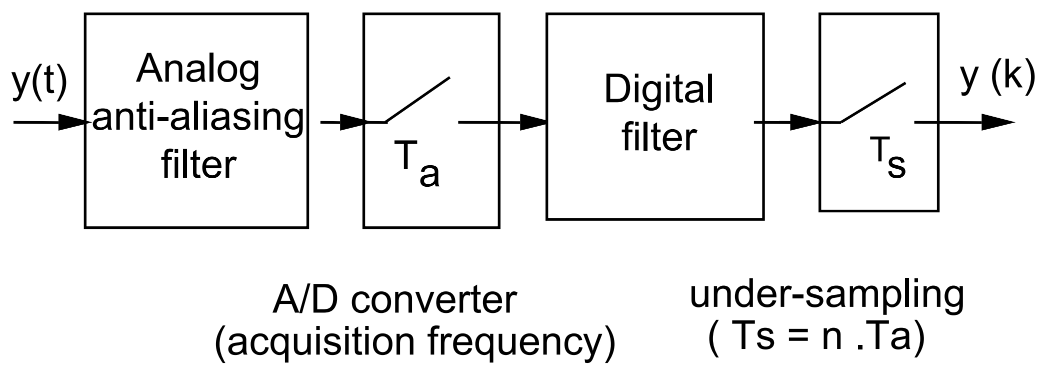
Consequence : bad identification

An anti-aliasing filter is necessary

- Solution 1 : analog filters
- Solution 2 : Over-sampling (for very slow frequency, $f < 2$Hz)

Oversampling

- In many digital control systems (DCS), the data acquisition frequency is significantly higher than that used in the control loops.
- Very small low sampling frequency (<2Hz)



$$f_a = n f_s$$

f_a = acquisition frequency
n = integer frequency multiple
f_s = sampling frequency of the control loop

Over-sampling

Digital anti-aliasing filter
For n > 3 a moving-average filter is sufficient

$$y_f(t) = \frac{y(t) + y(t-1) + \ldots + y(t-n+1)}{n}$$

Sizing of the PRBS signal

The frequency divider used for the PRBS should be a multiple of n, in order that the final sampling frequency f_s is still an integer multiple of the clock frequency of the PRBS.

PRBS clock frequency : $f_{PRBS} = \frac{1}{p} f_a$

Final sampling frequency: $f_s = \frac{1}{n} f_a = \frac{p}{n} f_{PRBS}$

Warning: (p/n) should be an integer

Signal conditioning

Elimination of the DC Component

It is necessary to work with signal variations (dynamics models)
If the DC component is not eliminated then: *bias on the identified parameters*

Stationary (or almost stationary) continuous components

a) Computation of the Mean Value (M.V.) from the I/O data set.
b) Elimination of the Mean Value (M.V.) from the I/O data set, and a new centered file is created.

Non-stationary continuous components

The I/O signals are replaced by their corresponding variations (eventually filtered).

$$y'(t) = \frac{y(t) - y(t-1)}{1 + f_1 q^{-1}} \quad ; \quad u'(t) = \frac{u(t) - u(t-1)}{1 + f_1 q^{-1}} \qquad -0,5 \le f_1 \le 0$$

Signal conditioning

Identification of a plant containing a pure integrator

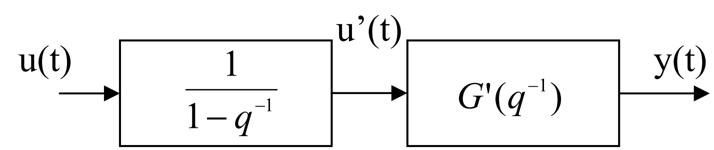
If the plant is: $G(q^{-1}) = \frac{1}{1-q^{-1}} G'(q^{-1})$

It is easier to identify *G'*

Method 1

- The input is replaced with its integral, the output is unchanged

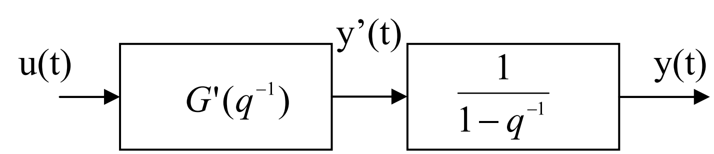
$y'(t) = y(t) \quad ; \quad u'(t) = \frac{u(t)}{1-q^{-1}}$



Method 2

- The output is replaced with its variations, the input is unchanged

$y'(t) = y(t) - y(t-1) \quad ; \quad u'(t) = u(t)$



Signal conditioning

Scaling of the I/O data set

Observations vector: $\Phi(t) = \left[-y(t), -y(t-1), \ldots, u(t), u(t-1), \ldots\right]$

Adaptation gain matrix: $F(t)^{-1} = \sum_{i=1}^{t} \Phi(i-1)\Phi^T(i-1) + \frac{1}{\delta} I \quad ; \quad \delta << 1$

If the magnitude of $u(t)$, $u(t-1)$... is very different from that of $y(t)$, $y(t-1)$, the gain matrix is not equilibrated, leading to very different convergence speed for the parameters $\hat{a}_i(t)$ and $\hat{b}_i(t)$

The I/O signals should be scaled (by multiplying, if necessary, both $u(t)$ and $y(t)$)

Once the identification is performed, the coefficients $\hat{b}_i(t)$ obtained must be re-scaled in order to obtain the good static gain

$$u_F = \alpha u \quad y_F = \beta y \qquad y = Gu\,; \quad y_F = G_F u_F \Rightarrow \beta y = G_F \alpha u$$

$$G = \frac{\alpha}{\beta} G_F$$

Selection of the model complexity

(if no algorithm for the automatic system order estimation is available)

The model:

$$A(q^{-1})y(t) = q^{-d}B(q^{-1})u(t) + C(q^{-1})e(t)$$

$$n_A = ? \quad n_B = ? \quad n_C = ? \quad d = ?$$

« a priori » choice of $\boldsymbol{n_A}$

Two cases :

1) industrial plant (temperature regulation, flow, concentration, etc.).
In general for these types of systems it holds :

$n_A \le 3$ $\qquad$ $n_A = 2$: very common value, (good initialization)

2) Electromechanical systems.
n_A results from the structural analysis of the system.

Example : flexible transmission with 2 resonant modes.
One selects $n_A=4$, since a 2nd order is needed for modeling each resonant mode.

Selection of the model complexity

Initial value for **d** and $\mathbf{n_B}$

- If no knowledge about the delay : d=0 (or d= d_{min} if a lower limit is known)
- The first coefficients of B are very small if d=0 and a delay is present
- One chooses : $n_B \geq dmax + 2$

Estimation of **d**

- Identification by RLS

$$B(q^{-1}) = b_1 q^{-1} + b_2 q^{-2} + b_3 q^{-3} + \ldots; \quad If\ |b_1| < 0.15|b_2| \rightarrow b_1 \approx 0 \rightarrow d = 1$$

$$B(q^{-1}) = q^{-1}(b_2 q^{-1} + b_3 q^{-2} + \ldots) \rightarrow d = 1$$

$$If\ |b_i| < 0.15|b_{i+1}|;\ i = 1,2,\ldots,d \rightarrow delay\ d$$

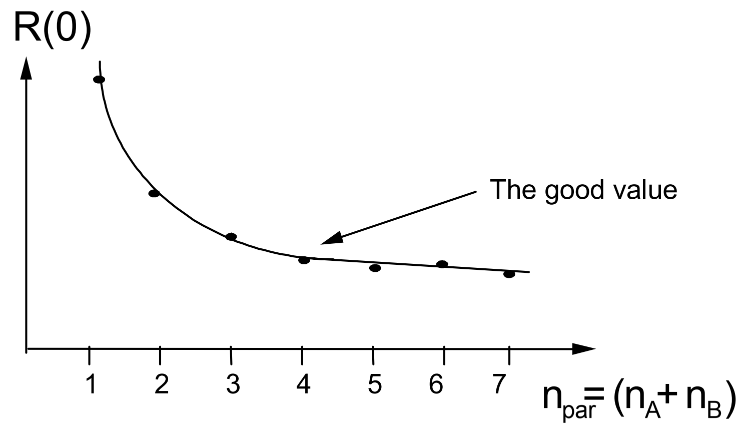
- This test can be completed by a visual validation with a step response
- The better estimation for d is done with the « validated » model

Choice for $\mathbf{n_C}$: in general $n_C = n_A$

Determination of n_A max and n_B max

Objective : to find the simplest model which validates

Test: RLS, Evolution of $R(0) = E(\varepsilon^2)$ as a function of the order



Consider $nA' = nA + 1$, nB and the corresponding variance of the residual errors $R'(0)$.

If : $R'(0) \geq 0.8 R(0)$ then it is not advisable to increase the order
(same for $n_B' = n_B + 1$)

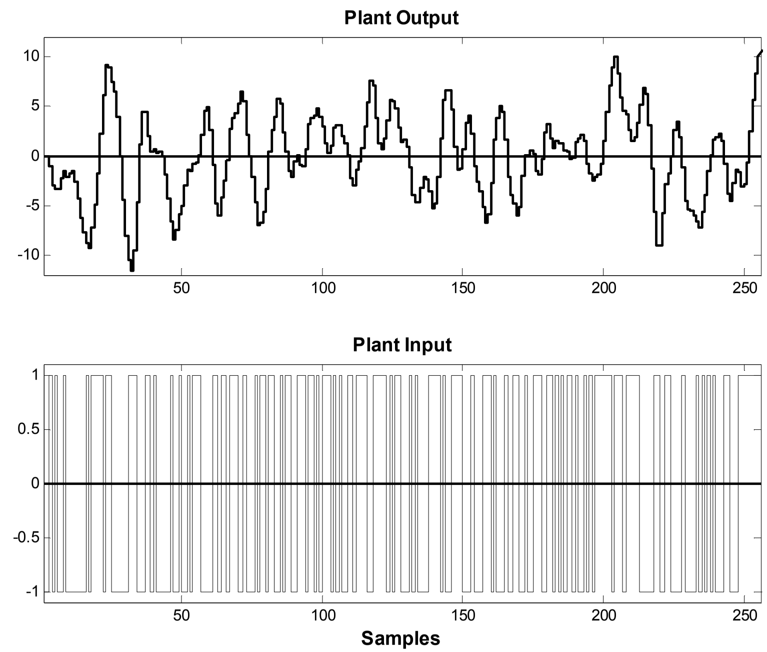
- Try different identification methods for finding a valid model without changing the orders of $\mathbf{n_A}$ and $\mathbf{n_B}$
- If no valid model is obtained then increase the orders of $\mathbf{n_A}$ and/or $\mathbf{n_B}$ (but not simultaneously)

File for testing the algorithms (simulation)

File T0 (no noise added to the measure):

$A(q^{-1})y(t) = q^{-1}B(q^{-1})u(t)$ $\qquad$ $A(q^{-1}) = 1 - 1.5q^{-1} + 0.7q^{-2}$ $\qquad$ $B(q^{-1}) = 1q^{-1} + 0.5q^{-2}$

File T0

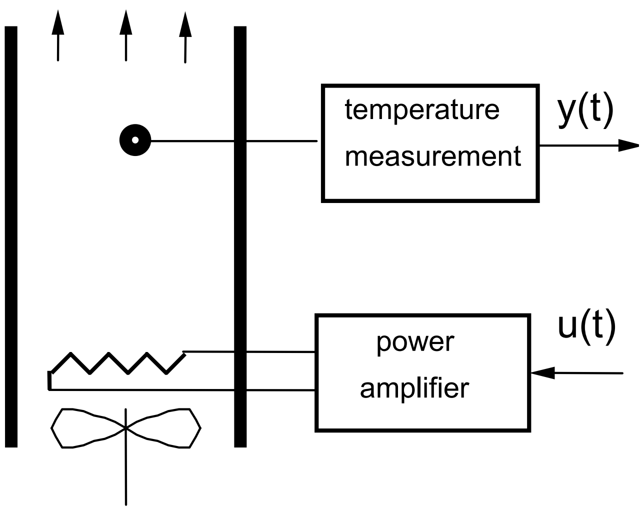


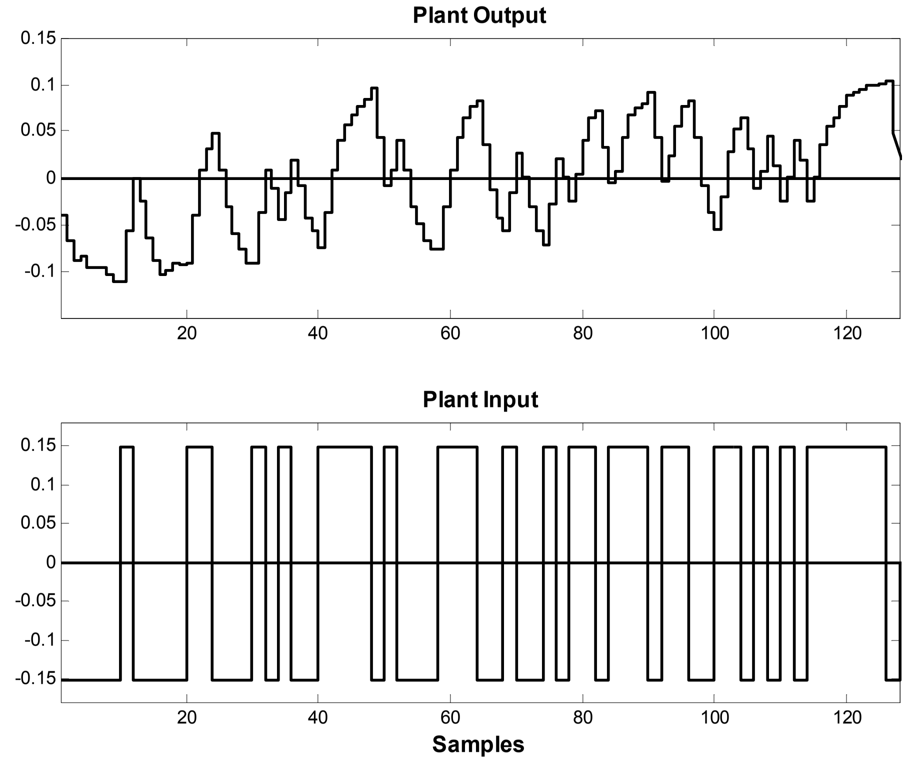
File T1 (with noise added to the measure):

$A(q^{-1})y(t) = q^{-1}B(q^{-1})u(t) + C(q^{-1})e(t)$

Available on the book website

Air heater

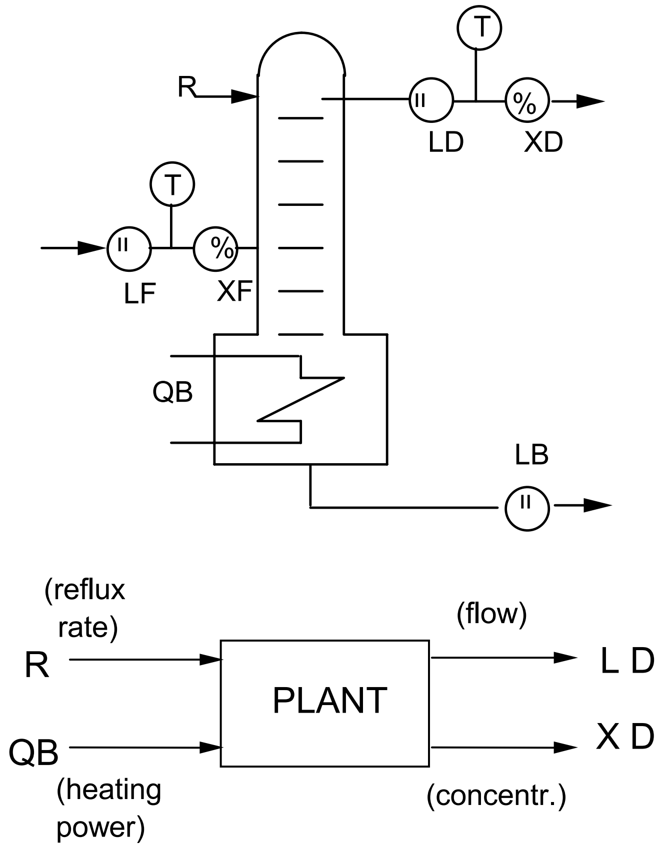


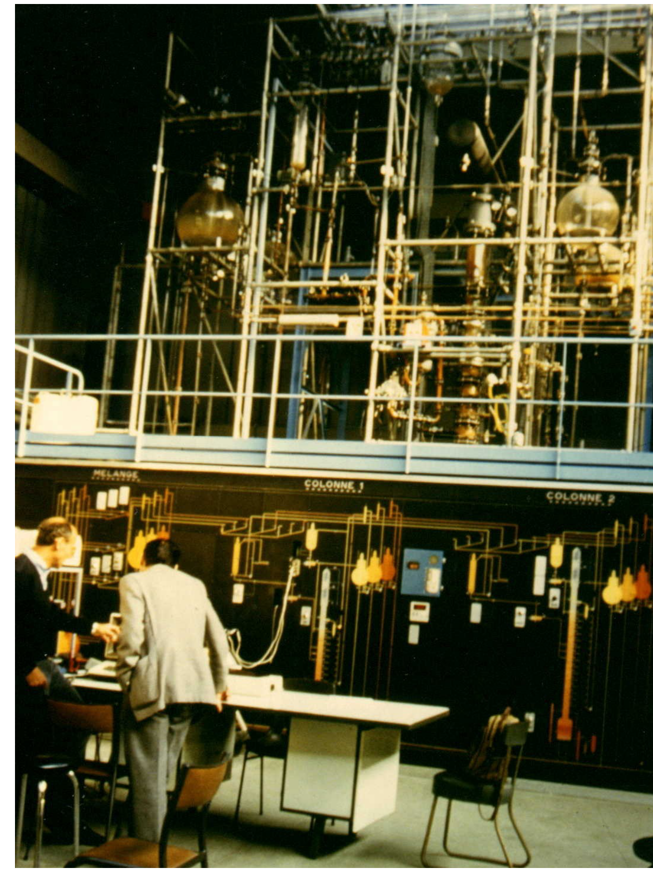


File : Aero.c (see book website)

See details of the results in the book (pg. 296 - 300)

Distillation column

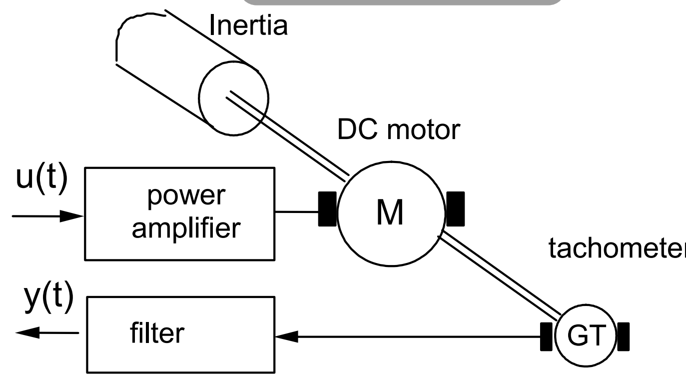


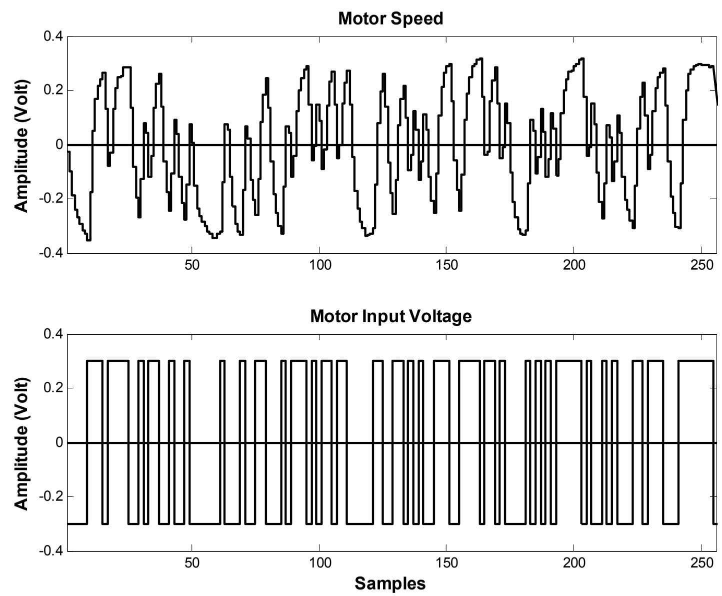


File : QXD (see book website) Input : QB Output: XD

See details of the results in the book (pg. 300 - 305)

DC Motor

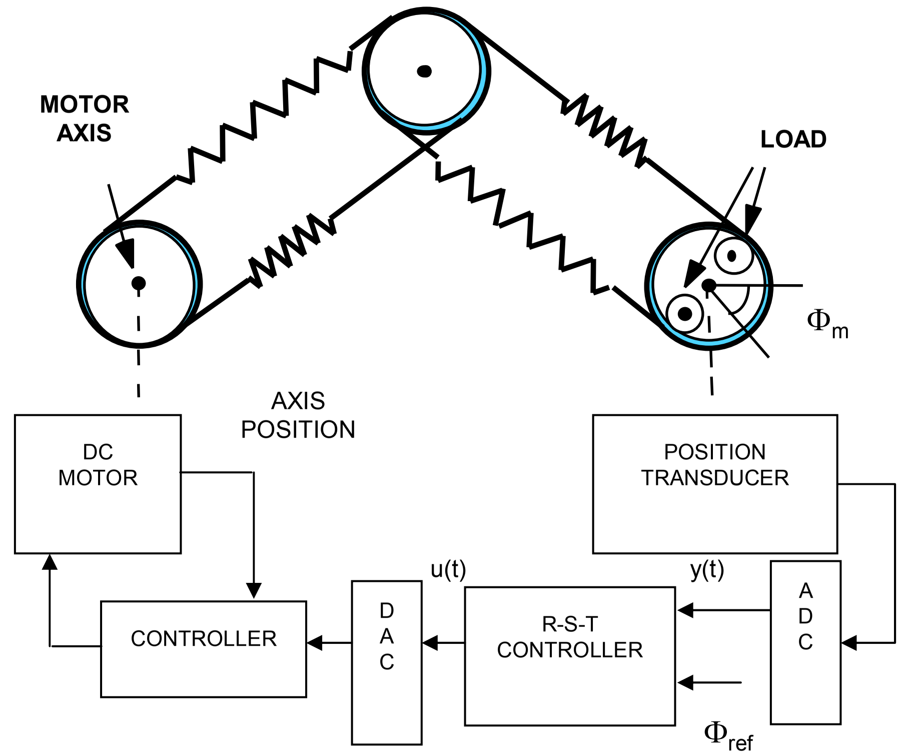




File : Mot3.c (see book website)

See details of the results in the book (pg. 305 - 309)

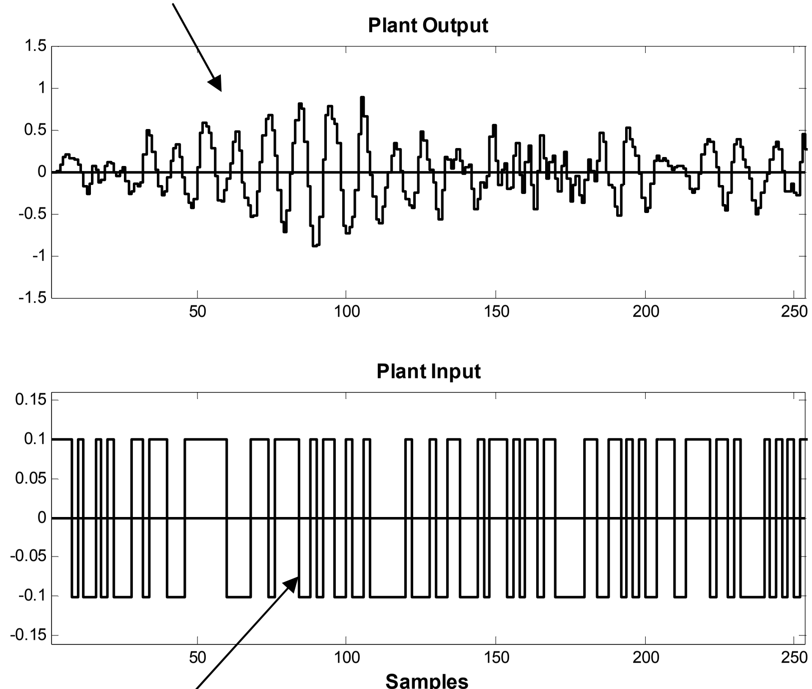
Flexible transmission



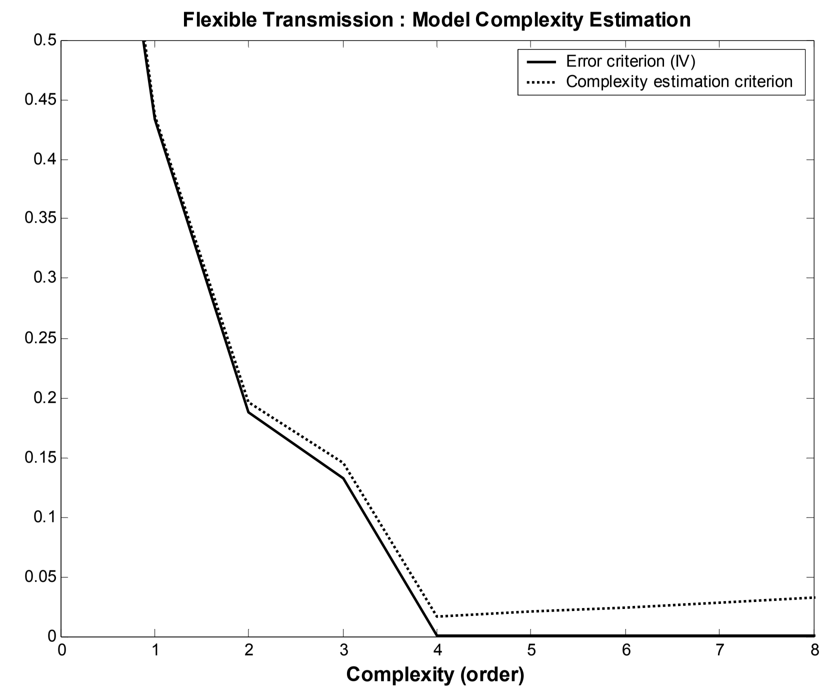
File : Poulbo1.C (see book website)

Flexible transmission identification

Measured position (y)



Excitation signal (u)
(PRBS)



Complexity estimation
Order $n = \max(n_A, n_B + d)$

Remark : The system is characterized by two resonant modes

Flexible transmission identification

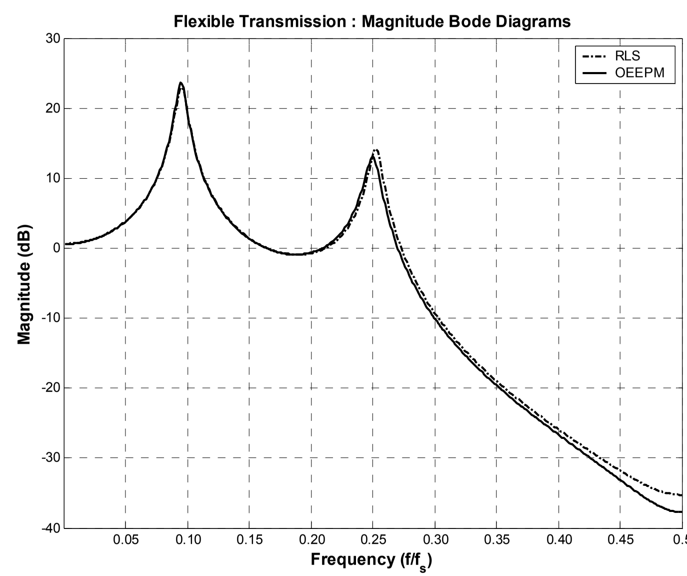
S=1 M=1 (RLS) A=1 FILE: POULBO1.C NS=254 DELAY D=2
COEFFICIENTS OF POLYNOMIAL A: A(1) = -1.5748
A(2) = 1.8329
A(3) = -1.4784
A(4) = 0.8895
COEFFICIENTS OF POLYNOMIAL B: B(1) = 0.3010
B(2) = 0.4181
VALIDATION TEST: Whiteness of the residual error
System variance: 0.3317 Model variance: 0.1053 Error variance R(0): 0.0007
NORMALIZED AUTOCORRELATION FUNCTIONS
Validation Criterion: Theor. Val.: |RN(i)| 0.136, Pract. Val.: |RN(i)| 0.15
RN(0) = 1.0000 → RN(1) = -0.5727 ←
→ RN(2) = 0.2360 ← RN(3) = -0.0475
RN(4) = -0.0158

R.L.S.

S=3 M=3 (OEEPM) A=1 FILE:POULBO1.C NS=254 DELAY D=2
COEFFICIENTS OF POLYNOMIAL A: A(1) = -1.60955
A(2) = 1.87644
A(3) = -1.49879
A(4) = 0.88574
COEFFICIENTS OF POLYNOMIAL B: B(1) = 0.30530
B(2) = 0.39430
COEFFICIENTS OF POLYNOMIAL C: C(1) = -0.67530 C(2) = 0.2283
C(3) = -0.0653 C(4) = -0.0585
VALIDATION TEST: Whiteness of the residual error
System variance: 0.1061 Model variance: 0.1055 Error variance R(0): 0.0004
NORMALIZED AUTOCORRELATION FUNCTIONS
Validation Criterion: Theor. Val.: |RN(i)| 0.136, Pract. Val.: |RN(i)| 0.15
RN(0) = 1.0000
RN(1) = -0.0425
RN(2) = 0.0959 RN(3) = -0.0563
RN(4) = -0.0407

O.E.E.P.M.

Good Validation



Frequency responses of the models identified

For more details see book (pg. 309 – 314)

A few concluding remarks

- The acquisition should be done by applying a signal rich in frequencies in the bandwidth of the system
- The PRBS is a basic signal that must be sized according to the application considered (as a function of the test time length and the time response)
- Attention to the filtering before the acquisition (anti-aliasing filtering)
- The I/O signals must be centered !
- The models identified must be validated!

This can help:

Matlab (Scilab) functions for generation of PRBS, complexity estimation,parameter identification and model validation are available on the book website