System Identification
Lecture 10
Experiment Design and MIMO models

Paul Van den Hof

Control Systems Group
Department of Electrical Engineering
Eindhoven University of Technology

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Contents

Experiment Design

- Preparatory experiments
- Design of input signals
- Sample frequency
- Data (pre)processing

Preparatory experiments

Objective

Obtain global information on the dynamics of the system, that can be used as prior information for identification and validation of an accurate parametric model.

Preparatory experiments

- Noise measurements on output
- Transient analysis (step responses)
  - Area of linearity
    Which input amplitudes can be applied to stay within linear behaviour?
  - time constants
    Fastest and slowest time constants determine relevant frequency-area
  - static gain
    Steady state response; behavior for frequency 0
  - choice of input signals
    Which of multiple inputs has the strongest effect on the output?

- Correlation-/frequency analysis (periodic signal, white noise)
  - Frequency range of interest
  - Length of experiment
    (At least 10 times slowest time constant, but factor dependent on SNR)
  - Time delays

Possibilities for preparatory experiments are strongly dependent on practical circumstances

Input design for identification

Criteria for designing an input signal:

- Present in relevant frequency area
- High power to achieve large signal-to-noise ratio (SNR)
- Amplitude small enough to warrant linear process behaviour

Input signals - (P)RBS

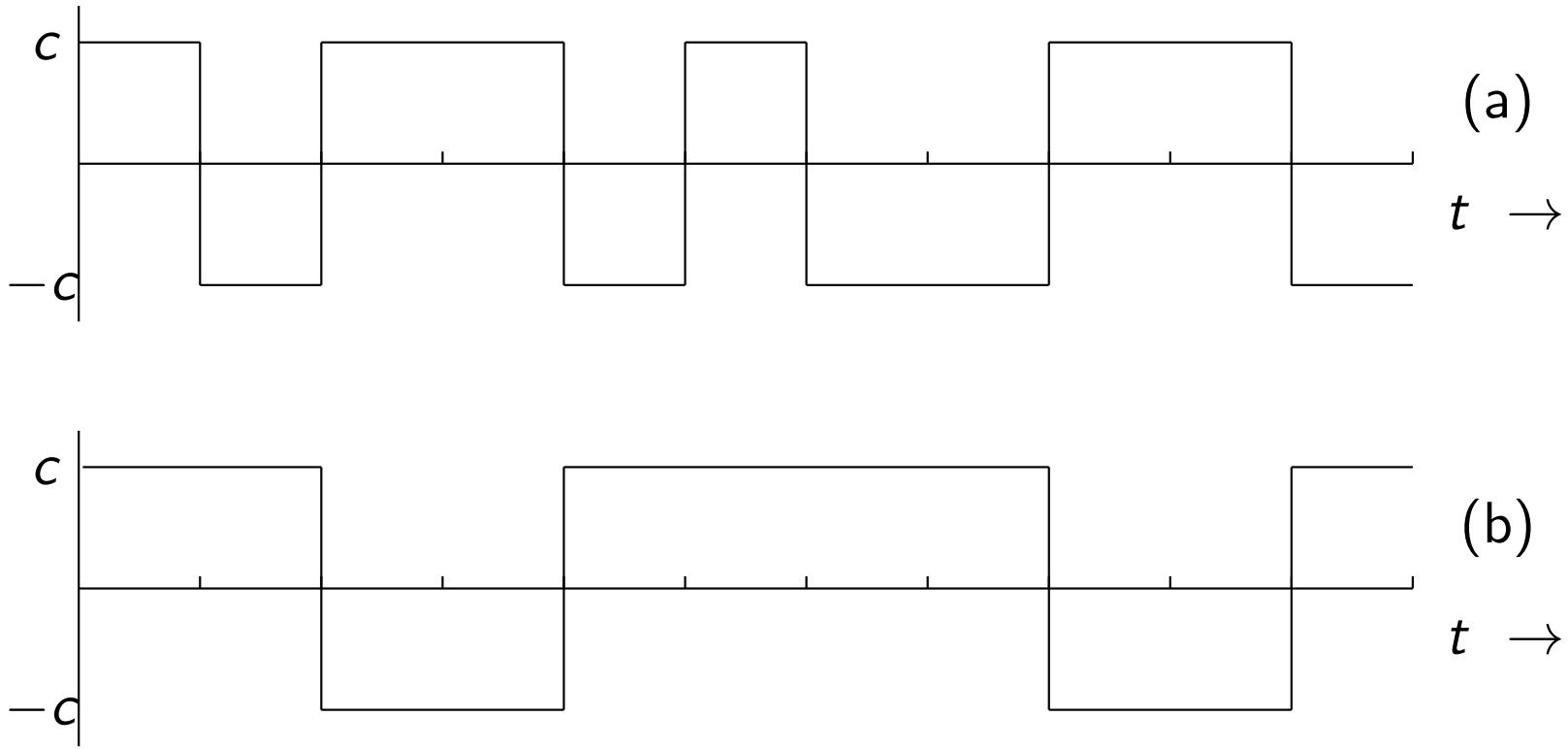
(Pseudo-)random binary sequence

- Computer-generated binary signal
- Properties of white noise (flat spectrum)
- Maximal signal power under amplitude constraints

$$\text{RBS:} \qquad u(t) = c \cdot \text{sign}[w(t)]$$

with $w(t)$ a white stochastic process.
Consequence:

$$\begin{aligned} R_u(\tau) &= c^2 \delta(\tau) \\ \Phi_u(\omega) &= c^2 \end{aligned}$$



(a) Typical RBS with clock period equal to sampling interval ($N_c = 1$);
(b) RBS with increased clock period $N_c = 2$.

Influencing frequency content of signal

- Filter $\rightarrow$ amplitude bound is lost
- Change of "clock period":

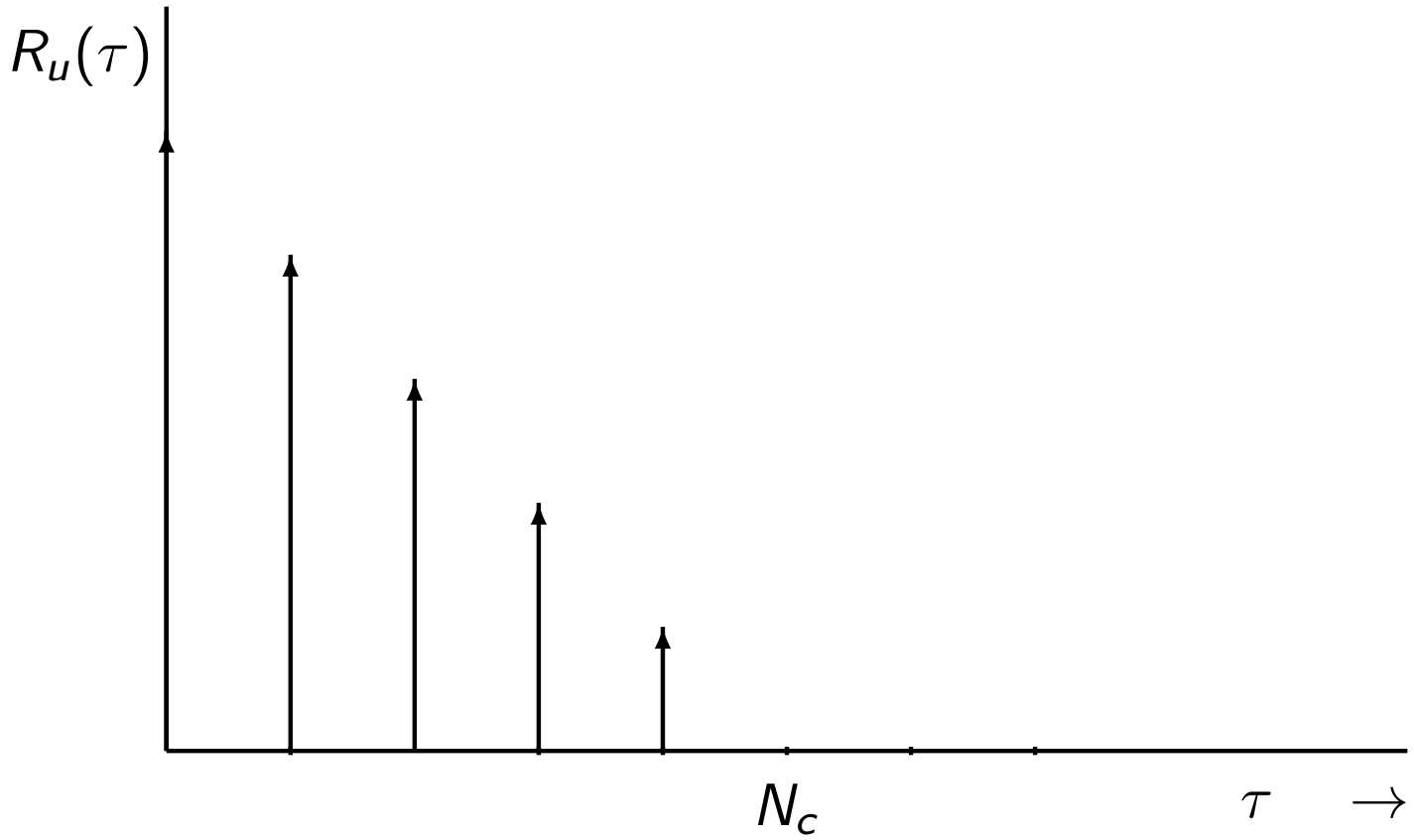
$$u(t) := c \cdot sign(w(int(t/N_c))$$

corresponds with N_c-times repeating of every value of the original signal (stretching)

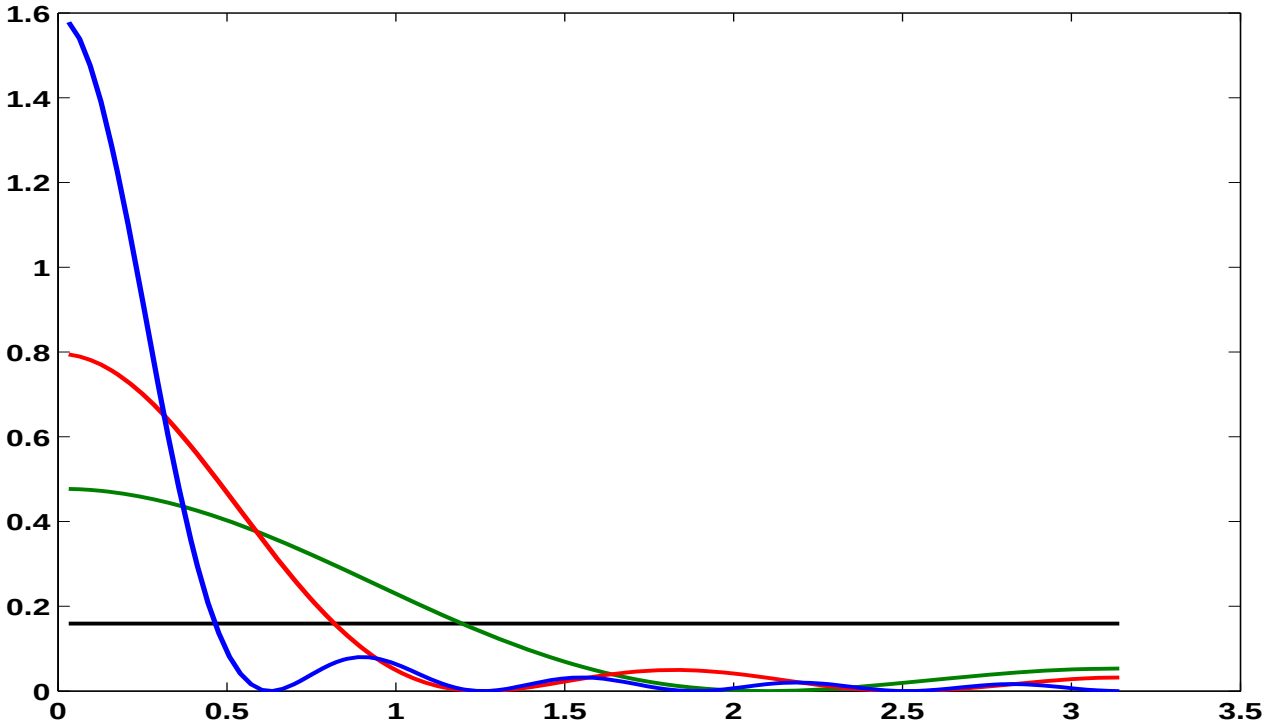
Consequence:

$$\begin{aligned} R_u(\tau) &= c^2 \cdot \frac{N_c - \tau}{N_c} & 0 \leq \tau \leq N_c \\ &= 0 & |\tau| \geq N_c \end{aligned}$$

Consequence for autocorrelation function:



Power spectral density $\Phi_u(\omega)$ for different values of N_c:



Spectrum $\frac{1}{2\pi}\Phi_u(\omega)$ of (P)RBS with basic clock period $N_c = 1$ (black), $N_c = 3$ (green), $N_c = 5$ (red), and $N_c = 10$ (blue).

Alternatives

RBS

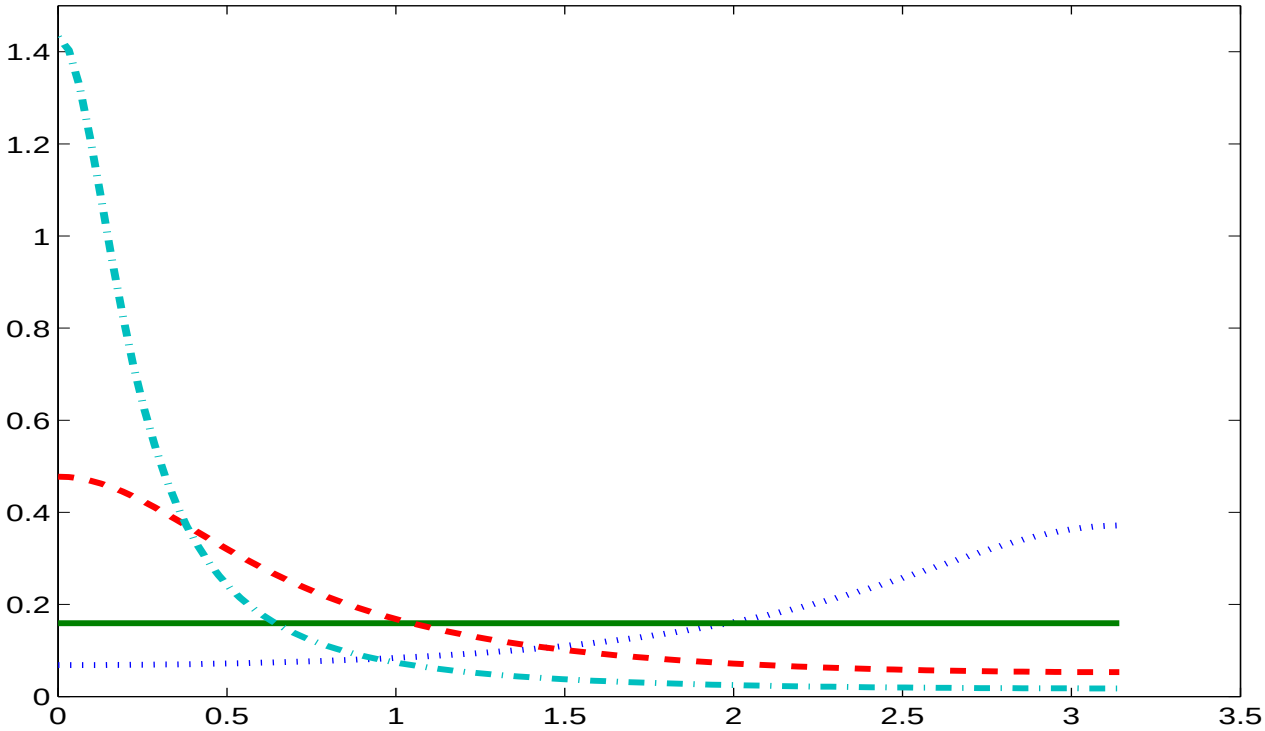
$$u(t) = \begin{cases} u(t-1) & \text{with probability } p \\ -u(t-1) & \text{with probability } 1-p \end{cases}$$

For $p = 0.5$, spectrum is flat.

PRBS

binary signal u constructed from a deterministic shift-register

Power spectral density $\Phi_u(\omega)$ for different values of p:



Spectrum $\frac{1}{2\pi}\Phi_u(\omega)$ for RBS with non-switching probabilities $p = 0.5$ (solid),$p = 0.75$ (dashed), $p = 0.90$ (dash-dotted) and $p = 0.3$ (dotted).

Periodic signals

$$u = \sum_{k=1}^{r} \alpha_k sin(\omega_k t + \phi_k)$$

- Choose the desirable frequency grid/frequency points
- Only amplitudes α_k determine the frequency spectrum
- Phases are typically chosen randomly (to avoid amplitude peaks)

Question:

When estimating a second order model (with 4 parameters):

- How many sinusoids do we need for consistency of $G(q, \hat{\theta}_N)$?
- Is it important where we choose the excited frequencies?

Optimal experiment design

More advanced problems:

- Find $\Phi_u(\omega)$ such that a particular scalar measure of the asymptotic parameter covariance matrix P is minimized (e.g. det P); [1]
- Find the least costly experiment (power and length of input signal) such that a certain confidence bound on the parameters / frequency response is achieved; [2]

[1] G.C. Goodwin and R.L. Payne (1977). *Dynamic System Identification: Experiment Design and Data Analysis*. Academic Press, New York.

[2] X. Bombois, G. Scorletti, M. Gevers, P.M.J. Van den Hof and R. Hildebrand. Least costly identification experiment for control. *Automatica, Vol. 42* (2006), pp. 1651-1662.

Sampling frequency

$$\omega_s = \frac{2\pi}{T_s}$$

Distinguish between situations:

- Data-acquisition and nonparametric identification
- Identification / Model application

Data-acquisition and nonparametric identification

- The higher ω_s the larger the frequency range that is captured, and the more data points (for averaging out noise)
- Lower bound: $\omega_s/2 >$ frequency range of interest.
  Rule of thumb for 1st order process: $\omega_s \geq 10 \cdot \omega_b$ with ω_b bandwidth.
- Upper bound determined by storage capacity.

Parametric identification / Model use

An upper bound appears to hold:

- Numerical aspects.
  Discrete-time system matrix: $A_d = e^{A_c T_s} \to I$ for $T_s \to 0$.
  All poles cluster around $z = 1$. 'Physical' length of range of difference equation becomes smaller and smaller.
- Stress on high-frequency range.
  Prediction horizon (1 step forwards) becomes smaller.

$$\theta^* = \arg\min_{\theta \in \Theta} \int_{-\pi/T_s}^{\pi/T_s} \Phi_\varepsilon(\omega, \theta) d\omega$$

When the integrand has no roll-off, errors in the high-frequency range become dominating.
E.g. ARX: then $\Phi_\varepsilon(\omega, \theta)$ contains

$$|G_0(e^{i\omega T_s}) - G(e^{i\omega T_s}, \theta)|^2 \cdot |Ae^{i\omega T_s}, \theta)|^2 \Phi_u(\omega)$$

Upper bound:

$$\omega_s \leq 30 \cdot \omega_b$$

Combined bounds (for 1st order system):

$$10\omega_b \leq \omega_s \leq 30\omega_b$$

With $\omega_b \sim 1/\tau$ and $\tau_{set,95} \sim 3\tau$ this leads to

$$\frac{1}{15}\tau_{set,95} \leq T_s \leq \frac{1}{5}\tau_{set,95}$$

Sampling strategy:

- Data-acquisition and preparatory experiments (including non-parametric identification) with high value of ω_s
- Reduction of ω_s before parametric identification (apply anti-aliasing filter)

Summary Experiment Design

- Experimental setup: gradually build-up from non-parametric information to parametric
- Optimal experiment design requires information on the system (!)
- Required length of experiment is determined by system dynamics and SNR
- Signals can be designed to have amplitude constraints (to avoid nonlinearities)
- Sampling should be done fast enough but not too fast

Multivariable (MIMO) models

What changes if we have multiple inputs and multiple outputs?

$$y(t) = G(q)u(t) + H(q)e(t)$$

$y(t), e(t) \in \mathbb{R}^p$, $u(t) \in \mathbb{R}^m$, $cov(e) = \Lambda_0 \in \mathbb{R}^{p \times p}$

m: number of inputs

p: number of outputs

$$u(t) = \begin{bmatrix} u_1(t) \\ \vdots \\ u_m(t) \end{bmatrix}; \quad y(t) = \begin{bmatrix} y_1(t) \\ \vdots \\ y_p(t) \end{bmatrix}$$

- Scalar transfer function $\rightarrow$ transfer function matrix:

$$G(q) = \begin{bmatrix} G_{11}(q) & \cdots & G_{1m}(q) \\ \vdots & \vdots & \vdots \\ G_{p1}(q) & \cdots & G_{pm}(q) \end{bmatrix}$$

$$H(q) = \begin{bmatrix} H_{11}(q) & \cdots & H_{1p}(q) \\ \vdots & \vdots & \vdots \\ H_{p1}(q) & \cdots & H_{pp}(q) \end{bmatrix}$$

H stable, inversely stable, and monic, i.e. $\lim_{z\to\infty} H(z) = I$, or

$$H(z) = I + h_1 z^{-1} + h_2 z^{-2} + \cdots$$

- Predictor / prediction error remains the same:

$$\varepsilon(t,\theta) = H(q,\theta)^{-1}[y(t) - G(q,\theta)u(t)]$$

$\varepsilon(t,\theta)$ is a p-dimensional vector;

- Identification criterion (scalar) is slightly generalized:

$$V_N(\theta) = \frac{1}{N}\sum_{t=0}^{N-1} \varepsilon^T(t,\theta)\Lambda^{-1}\varepsilon(t,\theta)$$

Λ^{-1} weighs the different components of ε with respect to eachother,
necessary e.g. in situations where the different components of $\varepsilon(t,\theta)$ have different scales (nm and km.....), or different noise levels.

- Model structures:
  This is the major change: different options for representing e.g. $G(q,\theta)$:
  - FIR:

$$G(q,\theta) = B_0 + B_1 q^{-1} + \cdots B_{n_b} q^{-n_b}, \qquad B_i \in \mathbb{R}^{p\times m}$$

  - Polynomial fractions, (ARX):

$$\begin{aligned} G(q,\theta) &= A(q,\theta)^{-1} B(q,\theta) \\ H(q,\theta) &= A(q,\theta)^{-1} \end{aligned}$$

    with

$$A(q,\theta) = I + A_1 q^{-1} + \cdots A_{n_a} q^{-n_a}$$

    possibly with $\{A_j\}_{j=1,n_a}$ restricted to being diagonal (i.e. same poles in each row of G)

Adaptations of the (SISO) theory:

- Maximum likelihood results for prediction error method remains valid provided that

$$\Lambda = \Lambda_0$$

  i.e. the identification criterion is weighted with the inverse of the noise covariance matrix.
  When this is unknown it needs to be estimated $\rightarrow \Lambda(\theta)$, and the ML criterion gets more complicated.

- Persistence of excitation condition for the (multivariable) input signal u, becomes:

$$\Phi_u(\omega) > 0 \quad \text{(matrix inequality)}$$

  in a sufficient number of points ω.
  Interpretation: each input signal should have a component that is not linearly dynamically related to the other inputs.

For MIMO models, the structural issues of polynomial models, has been a stimulus to work more often with state space models: A, B, C, D:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) + v(t) \end{aligned}$$

$$G(z) = D + C(zI - A)^{-1}B$$

$A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$, $D \in \mathbb{R}^{p \times m}$.

State space model is a Box-Jenkins structure (independent parametrizations for G and H):

$$\begin{aligned} x(t+1) &= \begin{bmatrix} A_G & 0 \\ 0 & A_H \end{bmatrix} x(t) + \begin{bmatrix} B_G \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ B_H \end{bmatrix} e(t) \\ y(t) &= \begin{bmatrix} C_G & C_H \end{bmatrix} x(t) + D_G u(t) + e(t) \end{aligned}$$

leading to

$$\begin{aligned} G(z) &= D_G + C_G(zI - A_G)^{-1}B_G \\ H(z) &= I + C_H(zI - A_H)^{-1}B_H \end{aligned}$$

States of G and H are decoupled:

$$x(t) = \begin{bmatrix} x_G(t) \\ x_H(t) \end{bmatrix}$$

Summary MIMO models

- All SISO principles remain in some form
- Technicalities occur (choice of orders/structures)
- Higher complexity