Nonlinear System Identification
A User-Oriented Roadmap

Johan Schoukens

18th IFAC Symposium on System Identification, SYSID 2018

VRIJE UNIVERSITEIT BRUSSEL

Technische Universiteit Eindhoven University of Technology


TC57
A380
A380

Linear

Nonlinear

Time-Varying

Outline

Why is nonlinear SI so involved?

Linear or nonlinear SI? A users decision

The lead actors in SI

Extensive case study

Conclusions

Outline

- Why is nonlinear SI so involved?
  - From hyperplane to manifold
  - Model errors
  - Process noise
- Linear or nonlinear SI? A users decision
- The lead actors in SI
- Extensive case study
- Conclusions

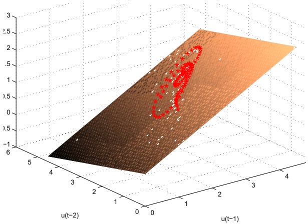
From hyperplane to manifold [1]

Linear models: a hyperplane

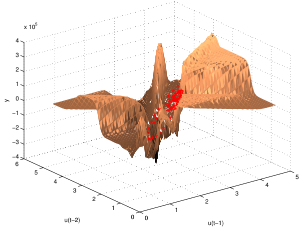
Nonlinear models: a manifold

- only known where domain is sampled
- extrapolation should be avoided

Linear



Nonlinear



[1] Acknowledgement Ljung, Bode Lecture 2003

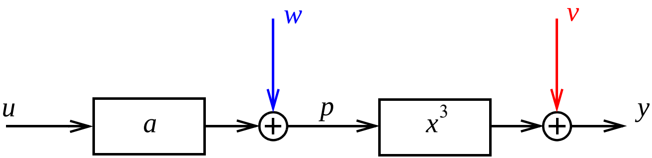
Model errors

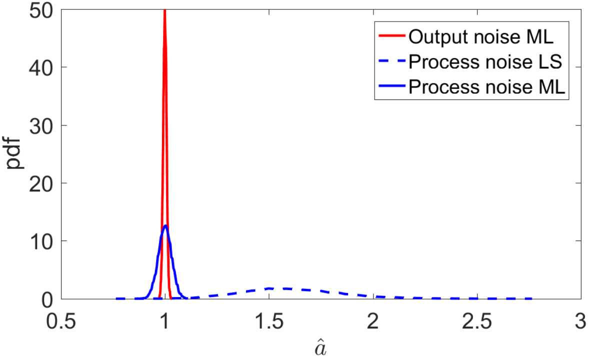
Impossible to avoid in many cases

Affect experiment design and choice cost function

Residuals no longer independent of input

Process noise





Outline

Linear or nonlinear SI? A users decision

Nonlinear SI much more ‘expensive’ than linear SI

Make a well informed decision

- Do we face a nonlinear identification problem?
- Safe to use a linear system identification approach?
- How much to gain with a nonlinear model?

Linear or nonlinear SI? A users decision

Nonlinear SI much more 'expensive' than linear SI

Make a well informed decision

- Do we face a nonlinear identification problem?
- Safe to use a linear system identification approach?
- How much to gain with a nonlinear model?

Detection, qualification, quantification NL distortions

- Characterize nonlinear behavior
- No increase of the measurement time
- Little user interaction

Linear or nonlinear SI? A users decision

Nonlinear SI much more 'expensive' than linear SI

Make a well informed decision

- Do we face a nonlinear identification problem?
- Safe to use a linear system identification approach?
- How much to gain with a nonlinear model?

Detection, qualification, quantification NL distortions

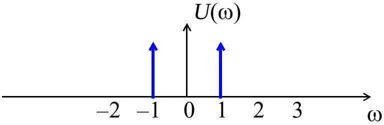
- Characterize nonlinear behavior
- No increase of the measurement time
- Little user interaction

Tool: well-designed periodic excitations

$$u_0(t) = \frac{2}{\sqrt{N}} \sum_{k=1}^{N/2-1} U_k cos\left(2\pi k f_0 t + \varphi_k\right)$$

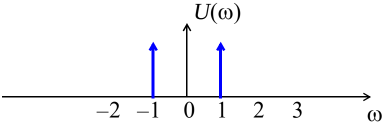
Understanding nonlinear systems: $y = u^3$

$$u(t) = 2\cos \omega t$$
$$= e^{j\omega t} - e^{-j\omega t} \text{ with } \omega = 1$$



Understanding nonlinear systems: $y = u^3$

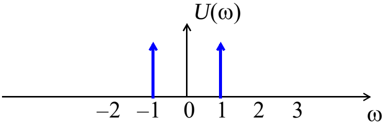
$$u(t) = 2\cos\omega t$$
$$= e^{j\omega t} - e^{-j\omega t} \text{ with } \omega = 1$$



$$y = u^3$$
$$= (e^{j\omega t} - e^{-j\omega t})(e^{j\omega t} - e^{-j\omega t})(e^{j\omega t} - e^{-j\omega t})$$

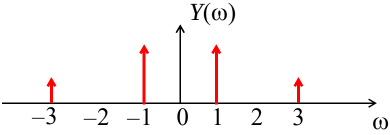
Understanding nonlinear systems: $y = u^3$

$$
\begin{aligned}
u(t) &= 2\cos\omega t \\
&= e^{j\omega t} - e^{-j\omega t} \text{ with } \omega = 1
\end{aligned}
$$

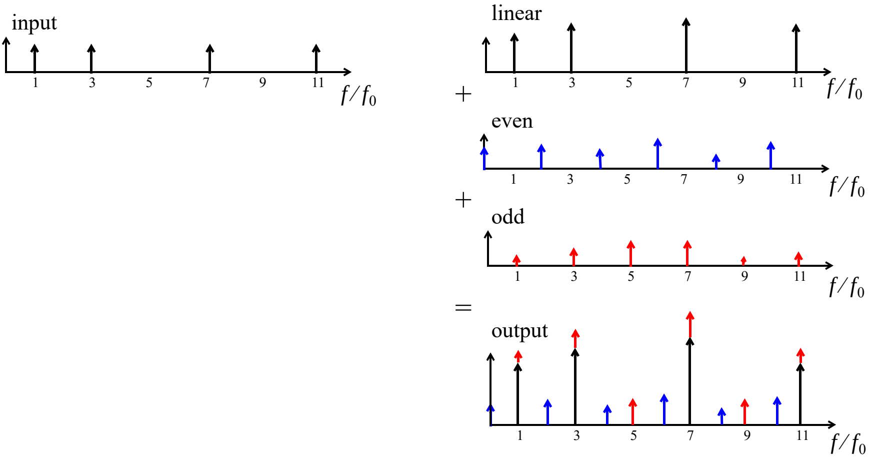


$$
\begin{aligned}
y &= u^3 \\
&= (e^{j\omega t} - e^{-j\omega t})(e^{j\omega t} - e^{-j\omega t})(e^{j\omega t} - e^{-j\omega t})
\end{aligned}
$$

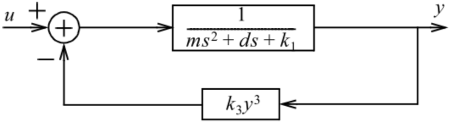
1	1	1	3
1	1	-1	1
1	-1	1	1
1	-1	-1	-1
-1	1	1	1
-1	1	-1	-1
-1	-1	1	-1
-1	-1	-1	-3



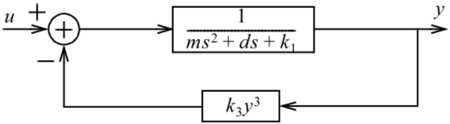
Detection and qualification of nonlinear distortions

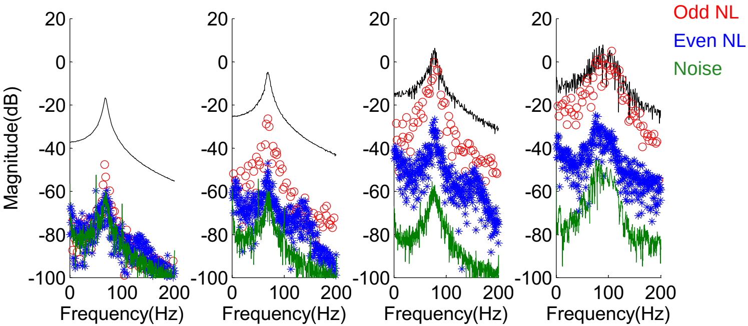


Example: the forced Duffing oscillator



Example: the forced Duffing oscillator

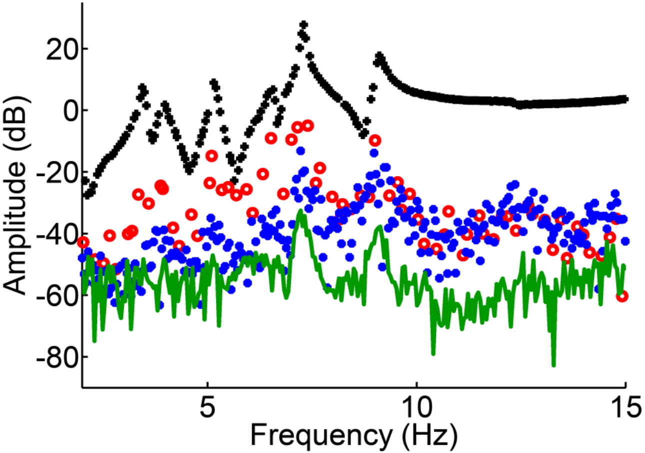




Example: an air fighter[2]

[2] Acknowledgement M. Vaes (VUB), B. Peeters, J. Debille (Siemens Industry Sofware), T. Dossogne, J.P. Noël, C. Grappasaonni, G. Kerschen (ULg)

Example: an air fighter[2]

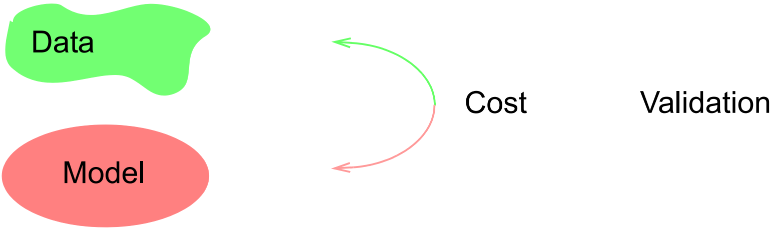


Output Odd NL Even NL Noise

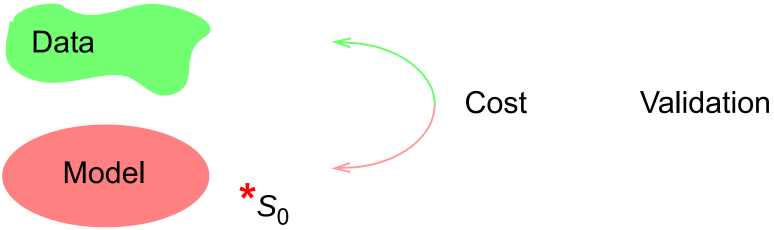
[2] Acknowledgement M. Vaes (VUB), B. Peeters, J. Debille (Siemens Industry Sofware), T. Dossogne, J.P. Noël, C. Grappasaonni, G. Kerschen (ULg)

Outline

Lead actors in SI from a nonlinear perspective

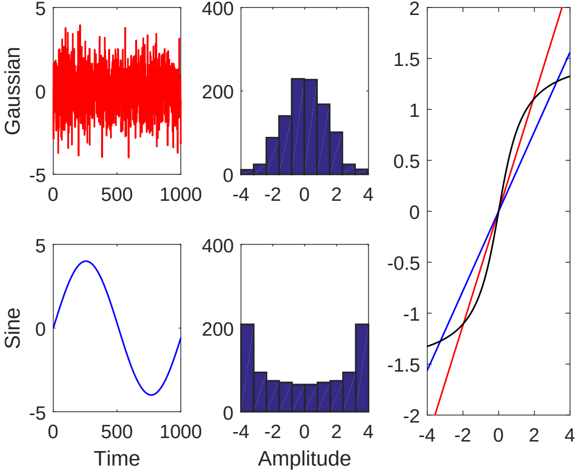


Lead actors in SI from a nonlinear perspective

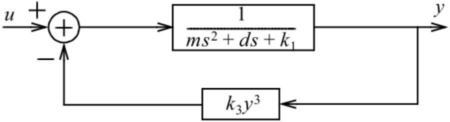


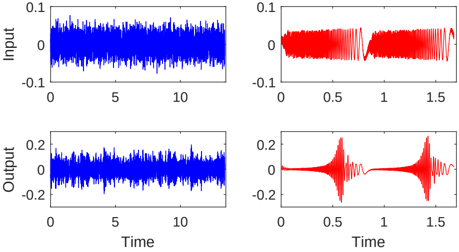
Impact model errors?

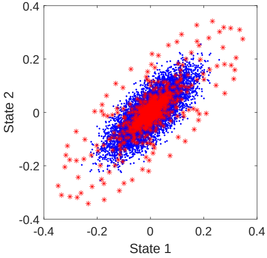
Data: amplitude distribution



Data: cover domain of interest







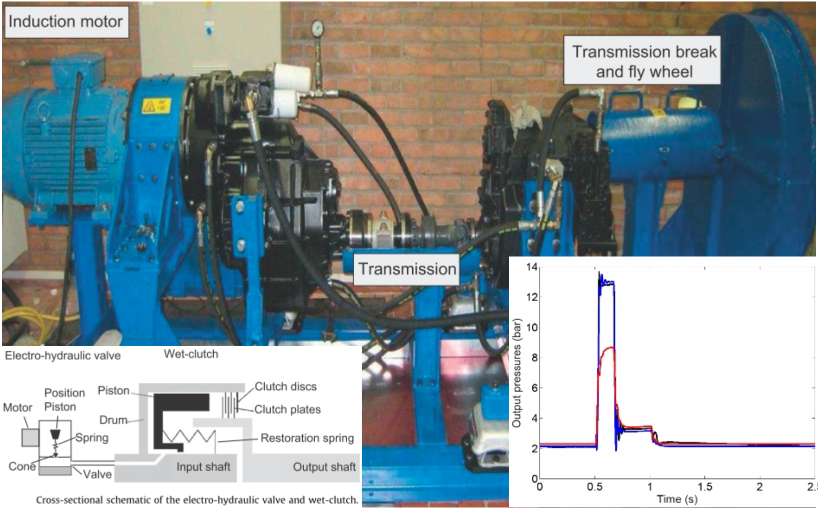
Data: experiment design

Optimal experiment design is still an open problem

User guidelines

- Use periodic excitations
  - Nonparametric distortion
  - No user interaction
  - Separation of plant and noise model
  - No inference with model errors
- Cover amplitude and frequency range of interest
  - Necessary but not sufficient condition
  - Pay attention to the state domain
  - Strongly linked to application: use excitation with same nature
- Use linear SI insights to excite the dynamics

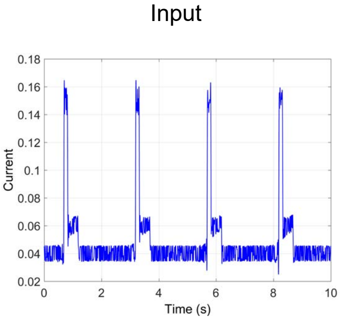
Experiment design: example[3]

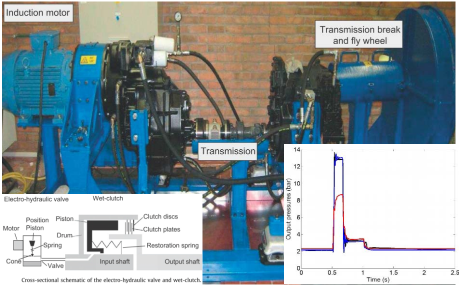


Cross-sectional schematic of the electro-hydraulic valve and wet-clutch.

3 Acknowledgement W.D. Widanage, A. Van Mulders (VUB), J. Stoev, G. Pinte (FMTC)

Experiment design: example[3]





[3] Acknowledgement W.D. Widanage, A. Van Mulders (VUB), J. Stoev, G. Pinte (FMTC)

Cost

Minimize the distance between the data and the model

Model errors dominate $\rightarrow$ move away from ML paradigm

Should reflect user's need how to shape model errors

Least squares cost functions in TD and FD

$$\hat{\theta}_N = \text{argmin}_\theta \sum_{t=1}^{N} \|y(t) - \hat{y}(t|\theta)\|^2$$

$$\hat{\theta}_N = \text{argmin}_\theta \sum_{f \in F} \|Y(f) - \hat{Y}(f|\theta)\|^2$$

Can be combined with regularization

Model

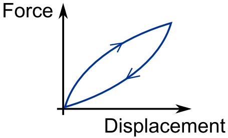
Estimates future outputs

$$\hat{y}(t|\theta, Z^{t-})$$

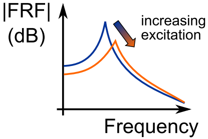
- Prediction model
  - $Z^t = u(s), y(s); s \leq t$
- Simulation model
  - $Z^t = u(s); s \leq t$
- Nonlinear function
  - $q(t) = f(p(t))$ with p, q signals in the model
  - f a static multivariate function
- Choice f, p, q
  - Users choice: white box - black box models
  - System behavior: open loop - dynamic NL closed loop

System behavior: open loop - dynamic NL closed loop

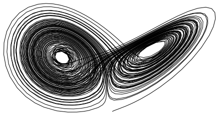
Hysteresis



Varying dynamics



Chaos



System behavior: open loop - dynamic NL closed loop

NL Open loop

NL not captured in a dynamic feedback loop
Fading memory

Examples

- NFIR
- Volterra
- Block-oriented models: Wiener, Hammerstein, Wiener-Hammerstein, Hammerstein-Wiener
- Nonlinear state space with lower triangular structure

Dynamic NL closed loop

Covers complex non-fading memory behavior
Can become unstable

Examples

- NIIR and NARX
- Closed loop block-oriented systems
- Nonlinear state space

Users choice: white box - black box models

- White models
  - Dedicated: new model for new problem
  - Expensive
  - Compact
  - Provide physical insight
- Black box models
  - Generic methodology
  - Behavior modeling
  - Explosion of parameters
  - Can provide intuitive insight

Users choice: white box - black box models

White models: physical models

- Estimate value physical parameters

Smoke-grey models: semi-physical modeling

- Natural selection of the variables

Steel-grey models: linearization based models

- Models depend on working point and nature of excitation

Slate-grey models: block oriented models

- Structural insight can be injected

Black models: universal approximators

- Volterra, NARX, Nonlinear state space

Pit-black models: nonparametric smoothing

Black box models complexity
Keep the exploding number of parameters under control

Control the model complexity

- Reduce number of parameters using regularization
- Data driven structure retrieval

Regularization

- Smooth solutions
  - Impose smoothness
  - number of parameters not changed
- Sparse solutions
  - Force parameters to zero
  - Brute force

Data driven structure retrieval

- Decouple multivariate nonlinear function $p = f(q)$
- Search for a natural basis
- From combinatorial to linear grow of the complexity

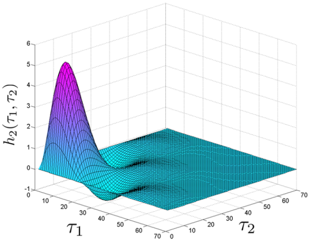
Regularization: impose smoothness on Volterra kernel

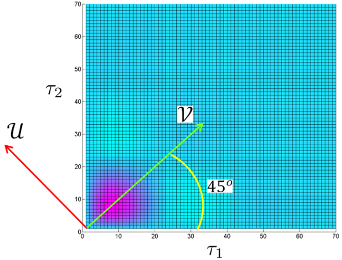
Model

$$y_0(t) = \sum_{0}^{n_1} g_1(\tau)u(t-\tau) + \sum_{0}^{n_2}\sum_{0}^{n_2} g_\alpha(\tau_1,\tau_2)u(t-\tau_1)u(t-\tau_2) + \ldots$$

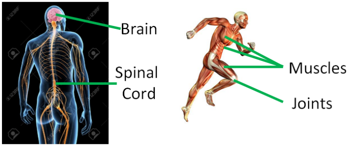
Cost

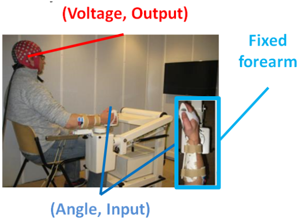
$$V = \frac{1}{N}\sum_{t=1}^{N}(y(t)-y_{mod}(t,\theta))^2 + [\theta_1^T \theta_2^T]\begin{bmatrix} P_1^{-1} & 0 \\ 0 & P_2^{-1}\end{bmatrix}\begin{bmatrix}\theta_1 \\ \theta_2\end{bmatrix}$$





Example: Volterra model wrist-brain sensorimotor system[4]





Experiment Design

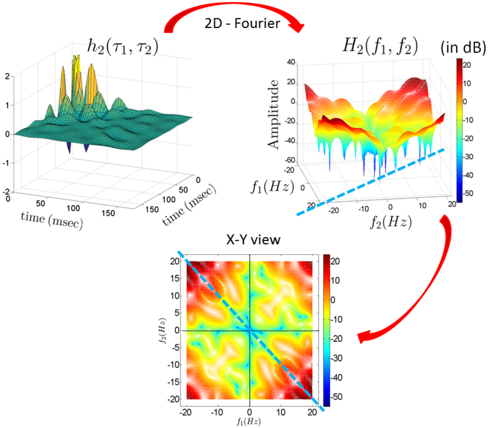
Random odd multisine $[1, 3, 5, 7, 9, 11, 13, 15, 19, 23]$Hz
Averaged over 210 periods
7 Realizations

Model

2^{nd} degree Volterra kernel
Fixed delay 20 ms
Memory length 130 ms (33 samples)

[4] Acknowledgement G. Birpoutsoukis (VUB), M. Vlaar, F. Van der Helm and A. Schouten (TU Delft)

Example: Volterra model wrist-brain sensorimotor system



Results

Linear kernel 10% VAF (Variance Acounted For)

2^{nd} degree Volterra model 45% VAF

2^{nd} degree Volterra model 60% VAF on improved experiment

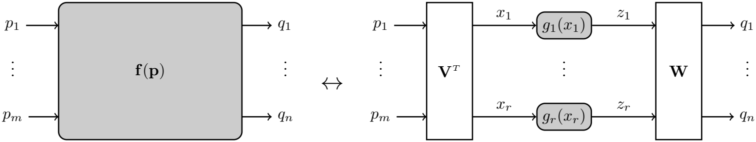
high-pass system

Data driven structure retrieval: decoupling[5]

Multivariate nonlinear function $q = f(p)$

Decouple

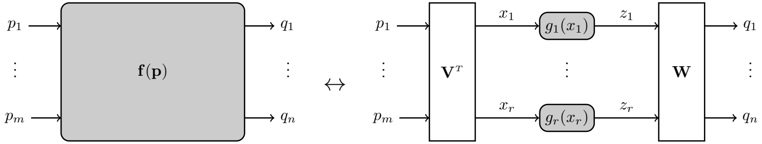
$$q = W\mathbf{g}(V^\top p)$$
$$\mathbf{g}_i = g_i(x_i),\, i = 1, \ldots, r \text{ with } x = V^\top p$$



[5] Acknowledgement D. Philippe, M. Ishteva, K. Tiels (VUB)

Decoupling: example

$$
\begin{aligned}
q_1 &= f_1(p_1, p_2) \\
&= 54p_1^3 - 54p_1^2 u_2 + 8p_1^2 + 18p_1p_2^2 + 16p_1p_2 - 2p_2^3 + 8p_2^2 + 8p_2 + 1 \\
q_2 &= f_2(p_1, p_2) \\
&= -27p_1^3 + 27p_1^2p_2 - 24p_1^2 - 9p_1p_2^2 - 48p_1p_2 - 15p_1 + p_2^3 - 24p_2^2 - 19p_2 - 3
\end{aligned}
$$



$$
\begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -3 & -1 \end{bmatrix} \begin{bmatrix} 2x_1^2 - 3x_1 + 1 \\ x_2^3 - x_2 \end{bmatrix}, \quad \text{with} \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 & -2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}
$$

Decoupling: from complexity control to model reduction

Complexity control

Number of parameters

Full model $O(n^d)$

Decoupled model $O(d)$

Force all $g_i(x) = g(x), \forall i$

$$q = W\mathbf{g}(V^\top p)$$

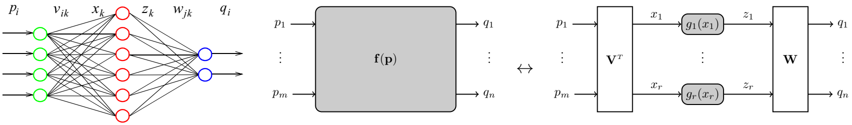
$$\mathbf{g}_i = g(x_i), i = 1, \ldots, r \text{ with } x = V^\top p$$

Model reduction

Reduce number of branches

Balance complexity versus model errors

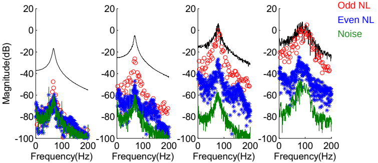
Decoupling: Link with Neural Networks



Outline

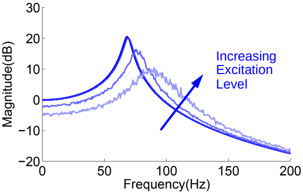
Forced Duffing Oscilator: Initial nonparametric analysis

Nonlinear distortion analysis



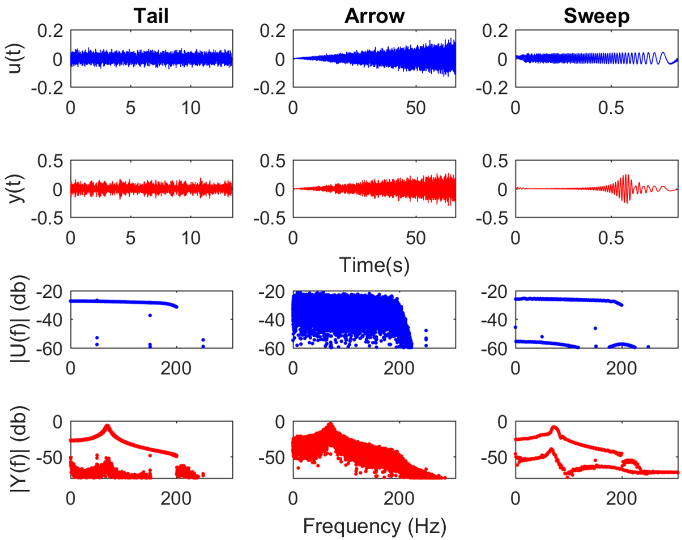
Conclusion: nonlinear distortions dominate noise

FRF measurement



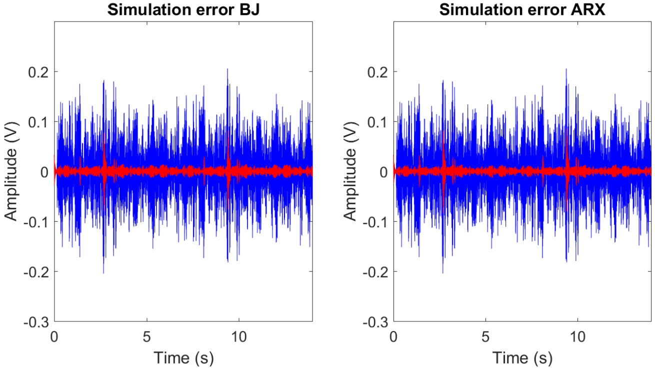
Conclusion: Nonlinear feedback model needed

Forced Duffing Oscilator: the data[6]

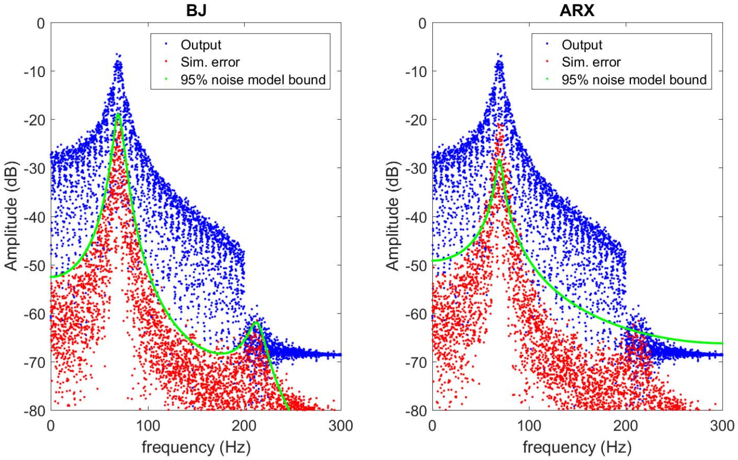


[6] Data available on http://www.nonlinearbenchmark.org (Silverbox Benchmark)

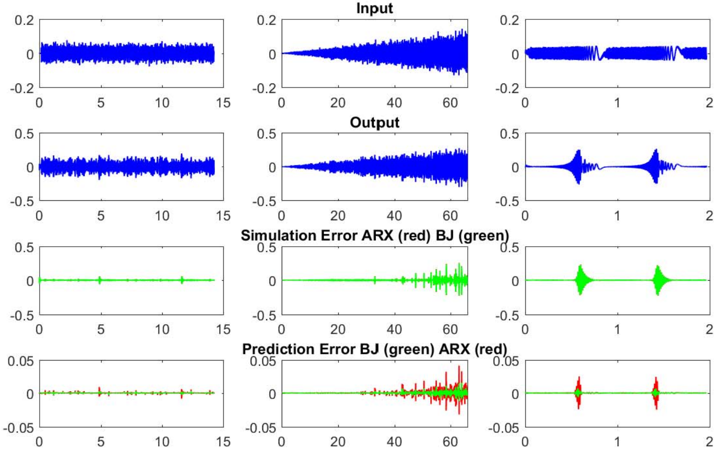
Forced Duffing Oscilator: linear models



Forced Duffing Oscilator: linear models



Forced Duffing Oscilator: linear models

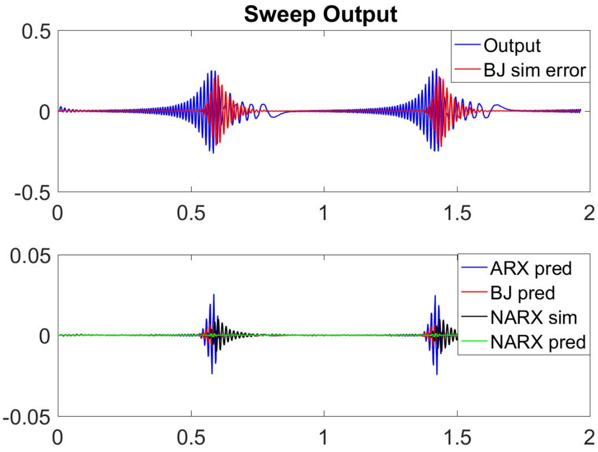


Forced Duffing Oscilator: from linear to nonlinear models

NARX

$$y(t) = P(u(t), u(t-1), u(t-2), y(t-1), y(t-2))$$

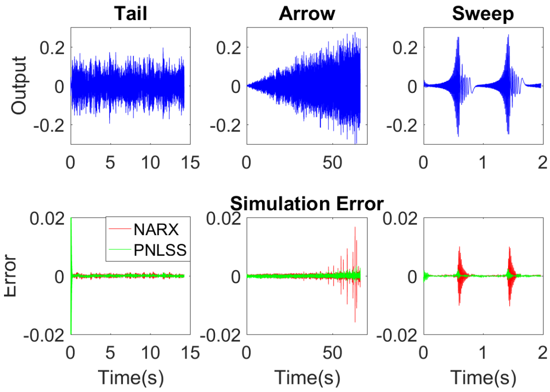
P Polynomial degree 3



Forced Duffing Oscilator: black box nonlinear state space model[7]

Nonlinear State space

- 2 states
- Polynomial degree 3



[7] Acknowledgement K. Tiels (University of Uppsala)

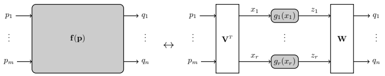
Forced Duffing Oscilator: from black box to highly structured models[8]

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = A \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + Bu(k) + \begin{bmatrix} f_1(x_1(k), x_2(k), u(k)) \\ f_2(x_1(k), x_2(k), u(k)) \end{bmatrix}$$

$$y(k) = C \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + Du(k)$$

f polynomial degree 3

Decouple $f(x_1(k), x_2(k), u(k))$



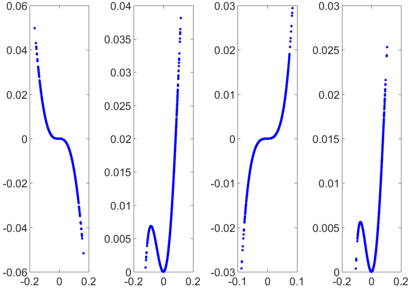
[8] Acknowledgement J. Decuyper (VUB)

Forced Duffing Oscilator: Decoupled Model

4 branches

Polynomial degree 3 $\rightarrow$ 5

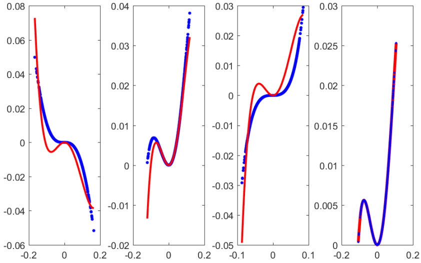
	BLA	NLSS	Decoupled
RMS error	12%	0.49%	0.40%
n_{θ_L}	5	5	5
$n_{\theta_{NL}}$	0	30	12



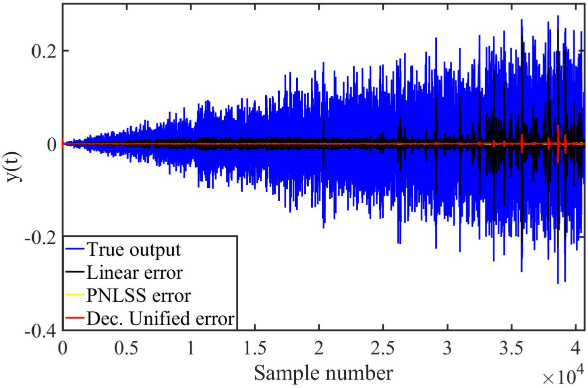
Forced Duffing Oscilator: Decoupled + Equal Branches

Impose all branches are equal

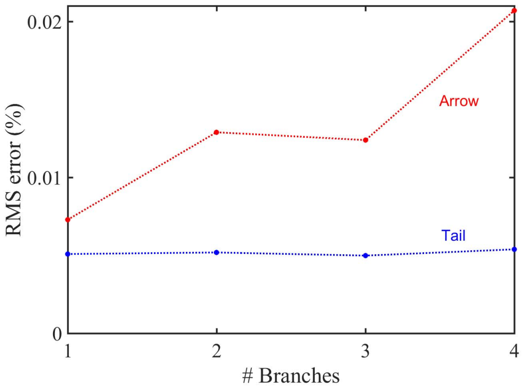
	BLA	NLSS	Decoupled	Equal Branches
RMS error	12%	0.49%	0.40%	0.40%
n_{θ_L}	5	5	5	5
$n_{\theta_{NL}}$	0	30	12	6



Forced Duffing Oscilator: Decoupled + Equal Branches



Forced Duffing Oscilator: Decoupled, Equal + Single Branch

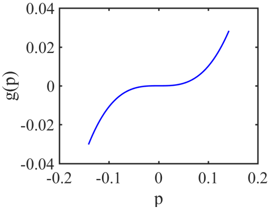


Forced Duffing Oscilator: Decoupled, Equal, Single Branch

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = A \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + Bu(k) + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} g(p)$$

$$y(k) = C \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + Du(k)$$

$$p = v_1 x_1(k) + v_2 x_2(k) + v_3 u(k)$$



Forced Duffing Oscilator: Final model

Black box

Data driven structure retrieval

Single branch

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = A \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + Bu(k) + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} g(p)$$

$$y(k) = C \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + Du(k)$$

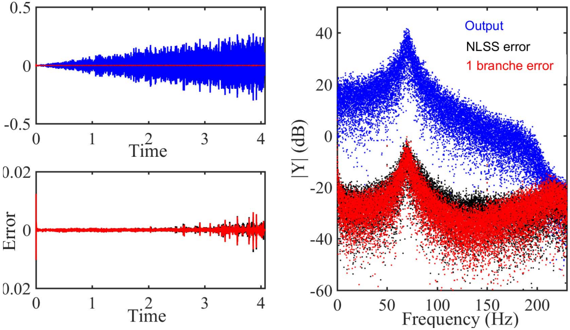
$$p = v_1 x_1(k) + v_2 x_2(k) + v_3 u(k)$$

Forced Duffing Oscillator: Final model

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = A \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + Bu(k) + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} g(p)$$

$$y(k) = C \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + Du(k)$$

$$p = v_1 x_1(k) + v_2 x_2(k) + v_3 u(k)$$



Conclusions

Why is nonlinear SI so involved?

- From hyperplane to manifold
- Model errors
- Process noise

Linear or nonlinear SI? A users decision

- Nonparametric distortion analysis
- Guidance model structure selection

The lead actors in SI

- Impact model errors on Experiment, Model, Cost Function

Data driven insight

- Structure retrieval
- Model reduction

Thanks to ...

My coauthors on nonlinear system identification

My discussion partners (during my sabbatical leaves)

Koen en Jan for their help on the presentation

Annick

This presentation is inspired on the tutorial article
Nonlinear System Identification. A User-Oriented Roadmap.
J. Schoukens and L. Ljung
Work in progress