TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Data-driven analytics and model learning in interconnected systems

Paul Van den Hof

Biomedical imaging and modeling cluster
TU/e Biomedical Engineering Dept
23 October 2023

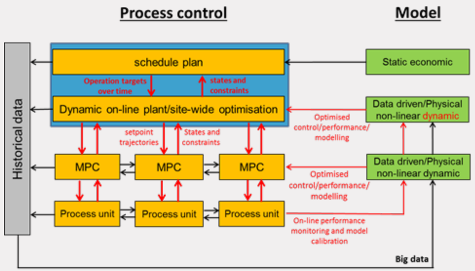
www.sysdynet.eu
www.pvandenhof.nl
p.m.j.vandenhof@tue.nl

erc European Research Council

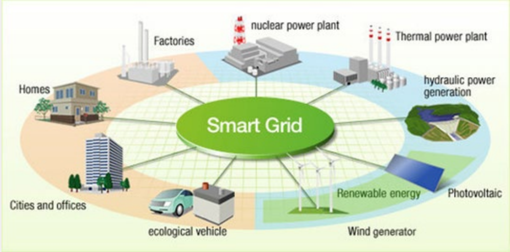
MATLAB Toolbox: www.sysdynet.net

Introduction – dynamic networks

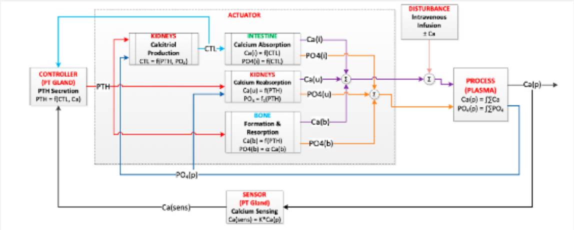
Decentralized process control



Smart power grid

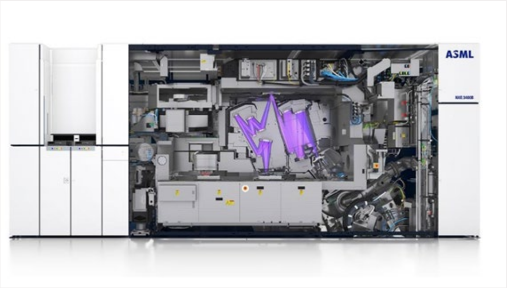


Physiological models

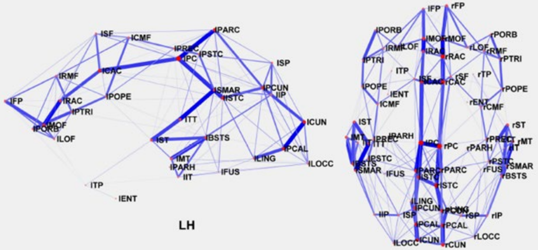


Christie, Achenie and Ogunnaike (2014)

Complex machines

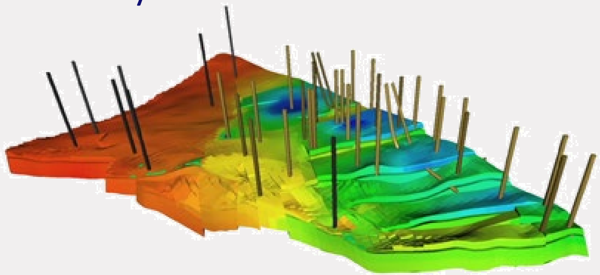


Brain network



P. Hagmann et al. (2008)

Hydrocarbon reservoirs



Mansoori (2014)

Introduction

Overall trend:

- (Large-scale) interconnected dynamic systems
- Many opportunities for sensing / collecting data
- The scope of control/optimization/diagnostics enlarges (larger systems)
- The infrastructure for control/optimization/diagnostics becomes distributed: local actions (multi agents)
- Data is “everywhere”, big data era, AI/machine learning tools

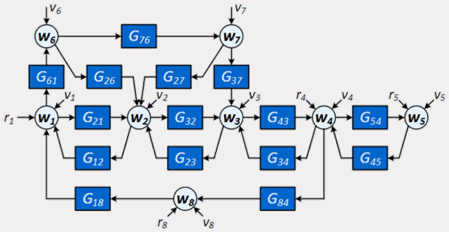
- Model-based operations require accurate/relevant models
- → **Learning and validating models/actions from data**
- → **Exploiting the available physical information (topology, dynamics, disturbances)**

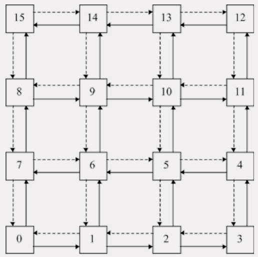
TU/e

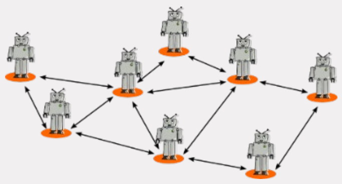
Network models

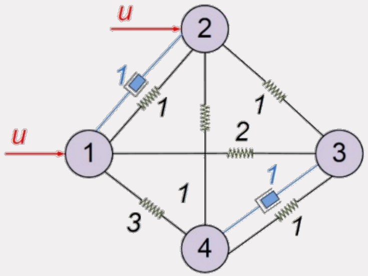
Network models

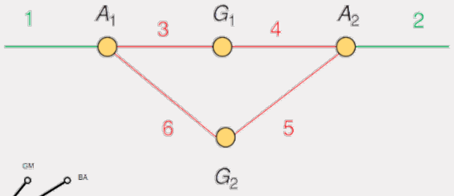
- scalable, describing the physics
- dynamic elements with cause-effect
- handling feedback loops (cycles)
- combining physical and cyber components
- centered around measured signals
- allow disturbances and probing signals

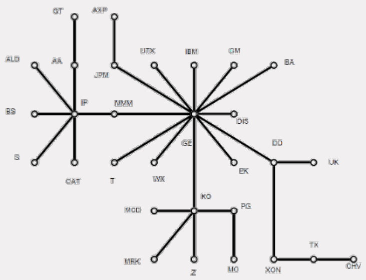


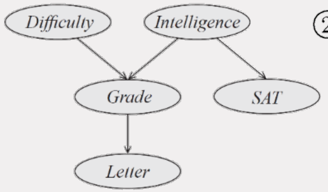


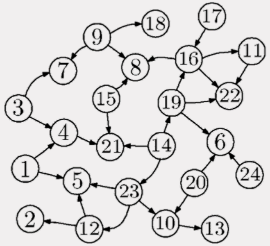


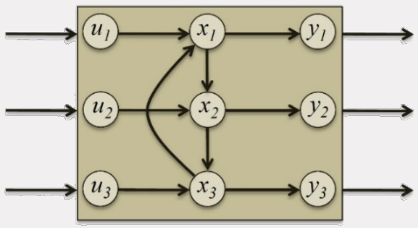


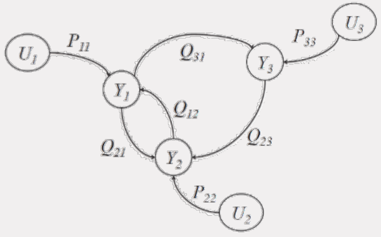






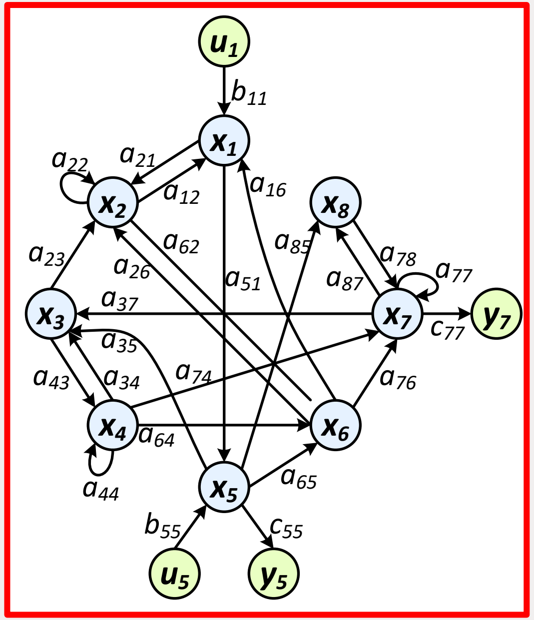






D. Materassi and M.V. Salapaka (2012)
R.N. Mantegna (1999)
www.momo.cs.okayama-u.ac.jp
J.C. Willems (2007)
D. Koller and N. Friedman (2009)
E.A. Carara and F.G. Moraes (2008)
P.E. Paré et al (2013)
P.M.J. Van den Hof et al (2013)
X.Cheng (2019)
E. Yeung et al (2010)

Network models



State space representation

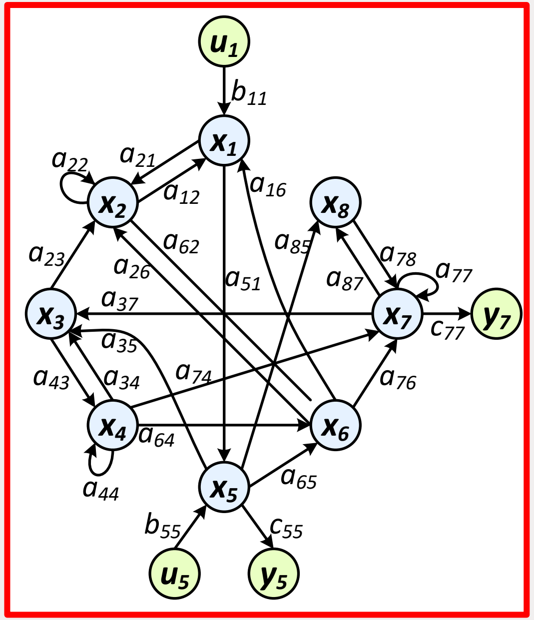
$$x(k+1) = Ax(k) + Bu(k)$$
$$y(k) = Cx(k) + Du(k)$$

- States as **nodes** in a (directed) graph
- State transitions (1 step in time) reflected by a_{ij}
- Transitions are encoded in links
- Ultimate break-down of system structure
- Actuation (u) and sensing (y) reflected by separate links

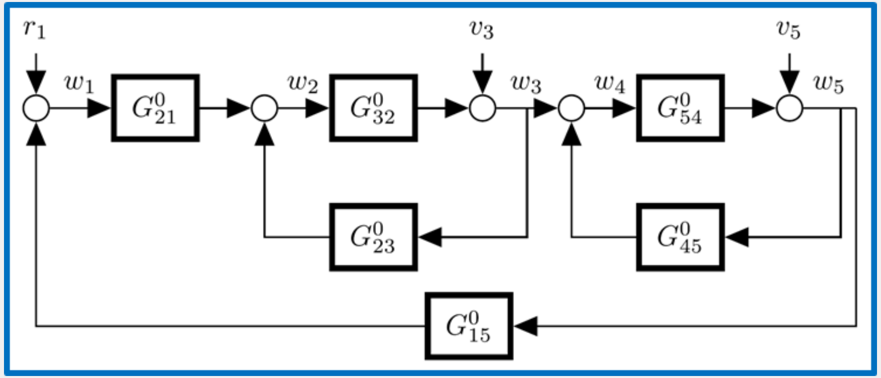
For data analytics problems:

- Lump unmeasured states in dynamic **modules**

Network models



State space representation [1]

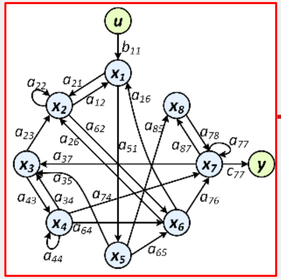


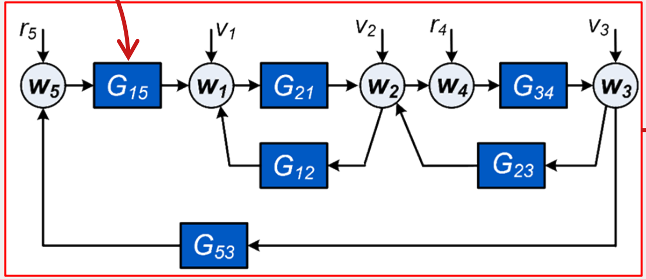
Module representation [2]

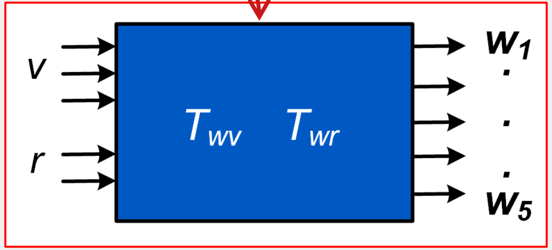
[1] Goncalves, Warnick, Sandberg, Yeung, Yuan, Scherpen,…

[2] VdH, Dankers, Goncalves, Warnick, Gevers, Bazanella, Hendrickx, Materassi, Weerts,…

Dynamic network models - zooming



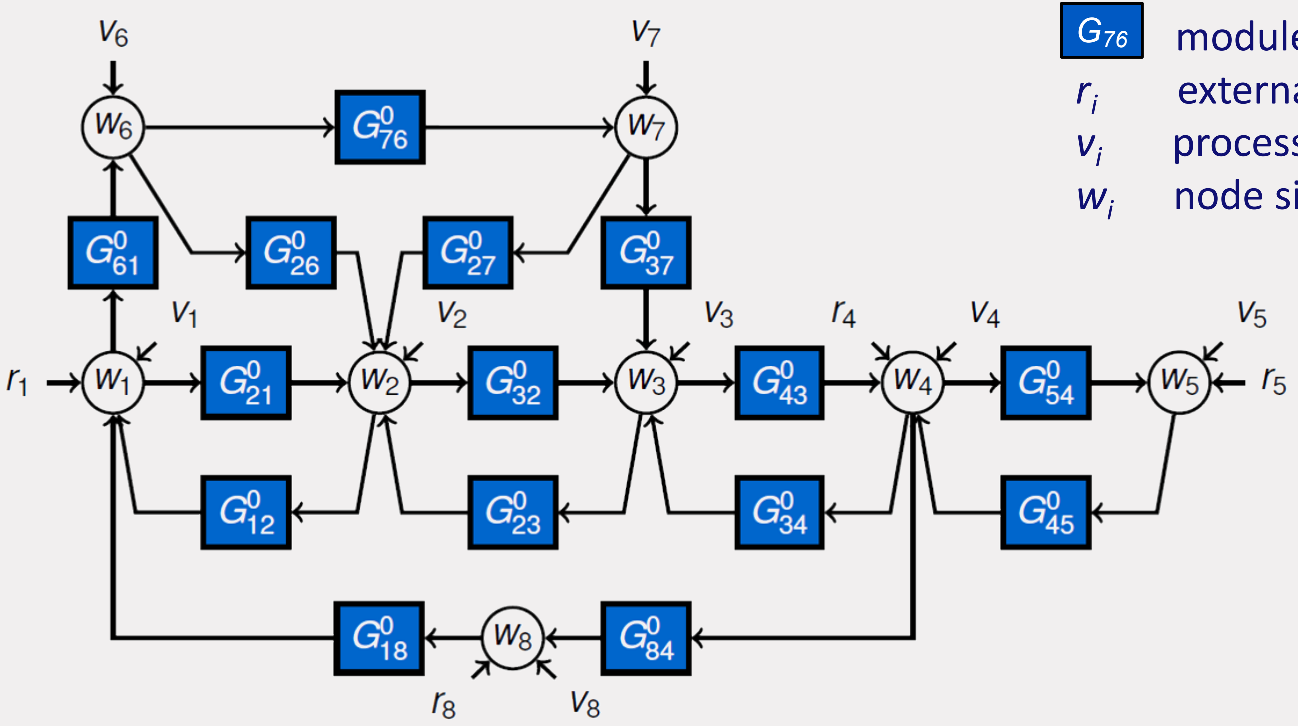




Increasing level of detail

Decreasing structural information

Dynamic network setup – Module framework



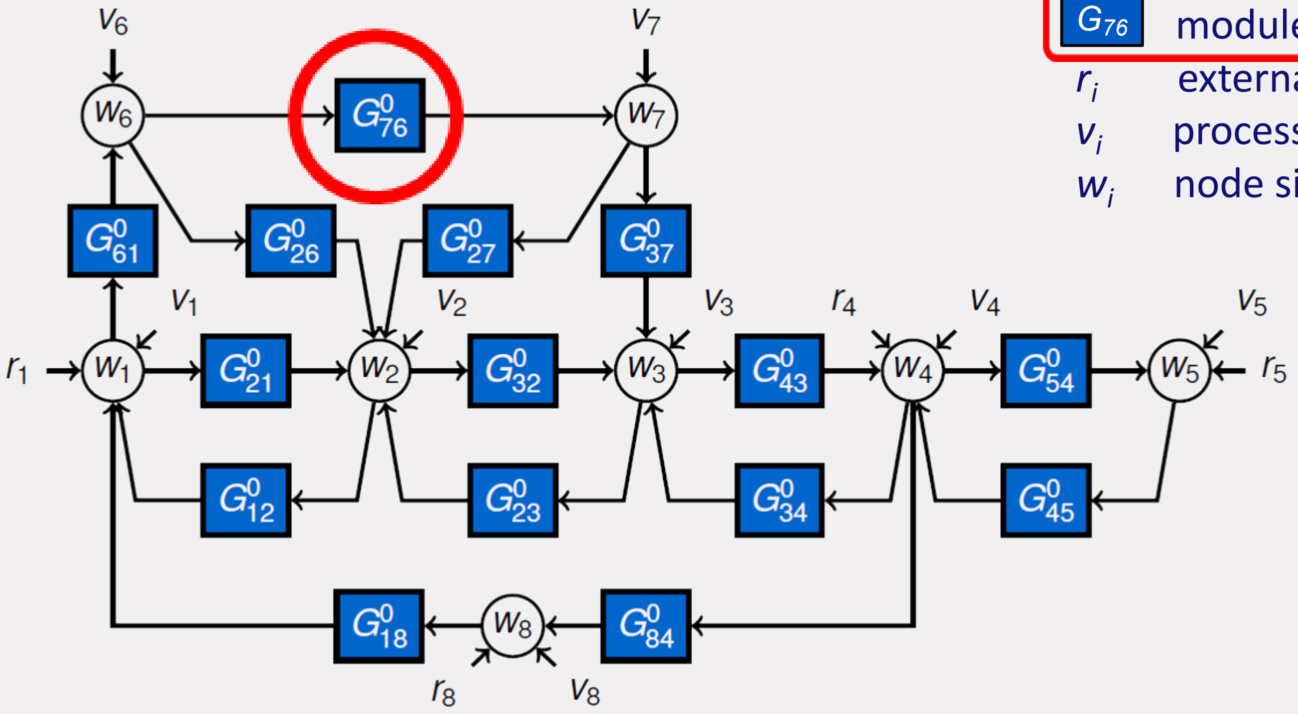
G_{76} module

r_i external excitation

v_i process noise

w_i node signal

Dynamic network setup – Module framework



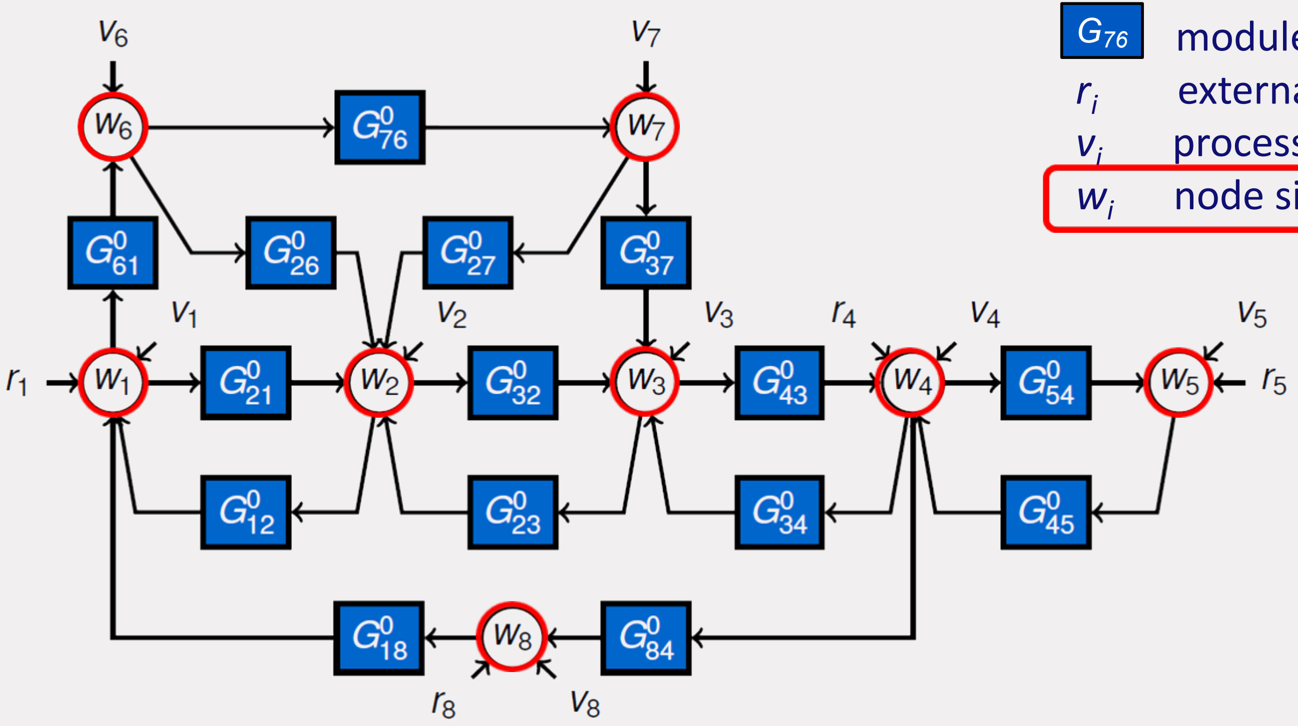
G_{76} module

r_i external excitation

v_i process noise

w_i node signal

Dynamic network setup – Module framework



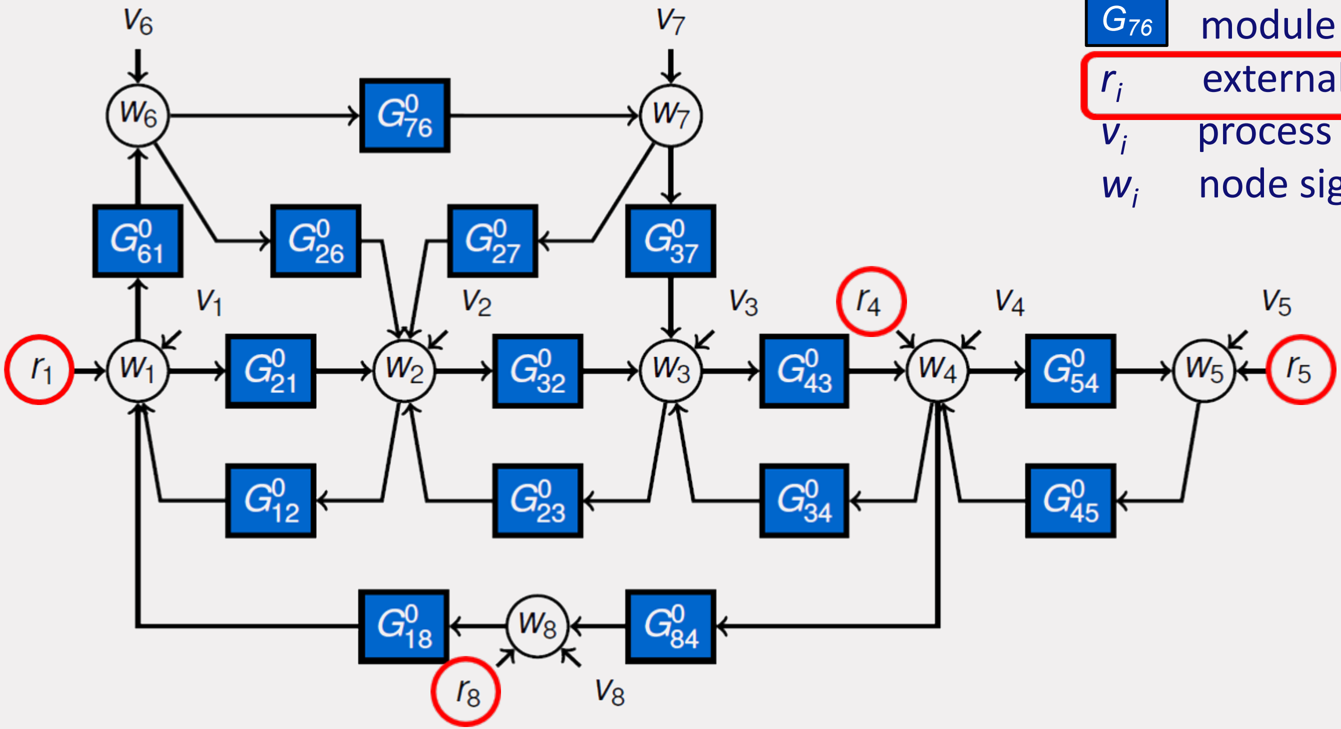
G_{76} module

r_i external excitation

v_i process noise

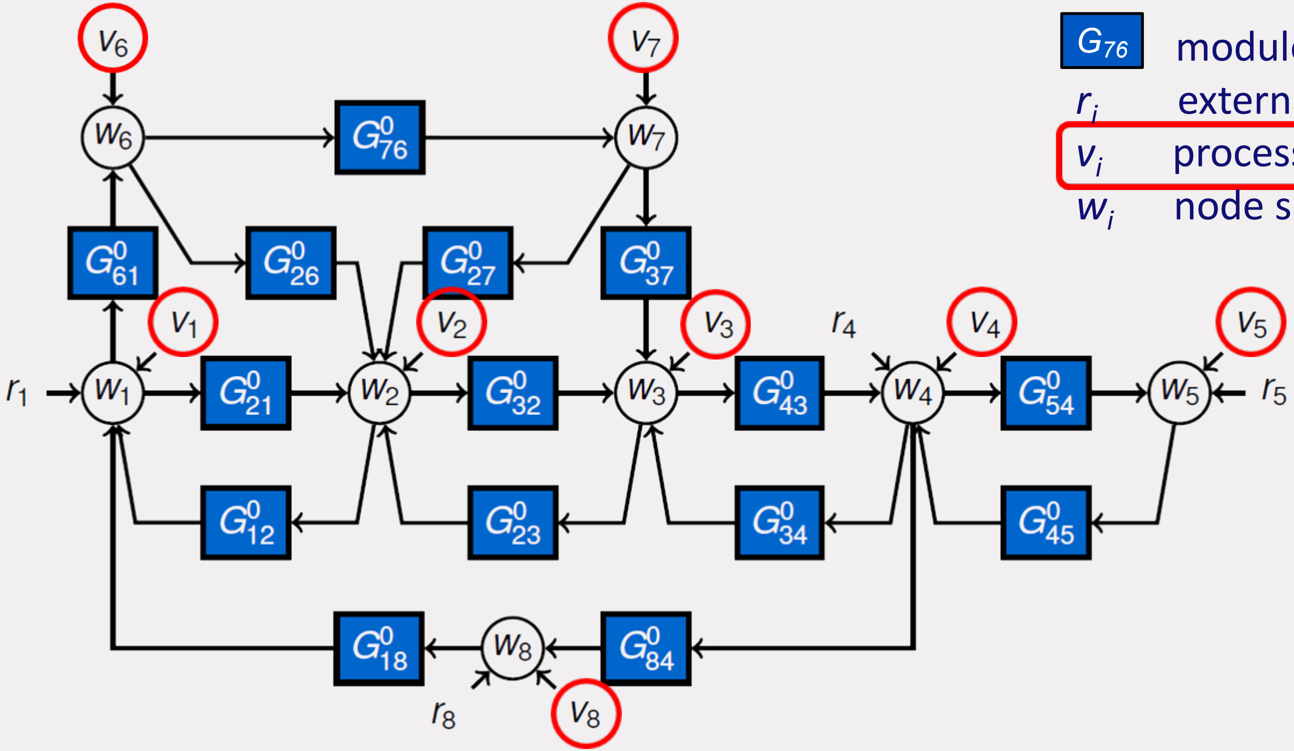
w_i node signal

Dynamic network setup – Module framework



G_{76}	module
r_i	external excitation
v_i	process noise
w_i	node signal

Dynamic network setup – Module framework



G_{76}	module
r_i	external excitation
v_i	process noise
w_i	node signal

Dynamic network setup – Module framework

Collecting all equations:

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & G^0_{12} & \cdots & G^0_{1L} \\ G^0_{21} & 0 & \ddots & G^0_{2L} \\ \vdots & \ddots & \ddots & \vdots \\ G^0_{L1} & G^0_{L2} & \cdots & 0 \end{bmatrix}}_{\text{Network matrix } G^0(q)} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} + R^0 \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_K \end{bmatrix} + H^0 \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{bmatrix}$$

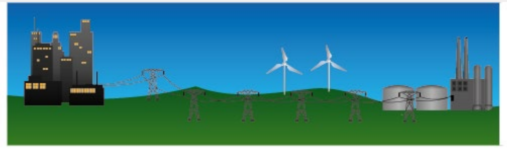
$$w(t) = G^0(q)w(t) + \underbrace{R^0(q)r(t)}_{u(t)} + v(t); \qquad v(t) = H^0(q)e(t); \quad cov(e) = \Lambda$$

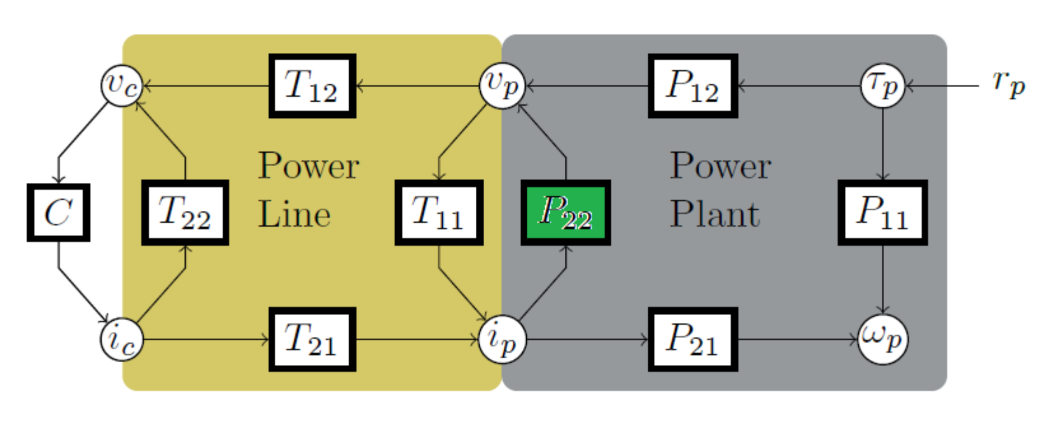
- The topology of the network is encoded in the structure (non-zero entries) of G^0.
- Measured node signals might be subject to sensor noise: $\tilde{w}_k = w_k + s_k$

J. Gonçalves and S. Warnick, IEEE TAC, 2008.
PVdH et al., Automatica, 2013.

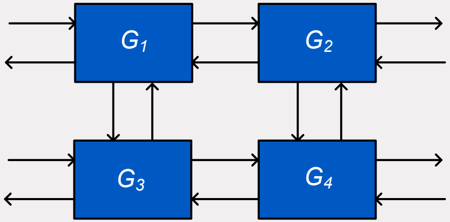
Alternative models

Bilaterally coupled (two-port) system:





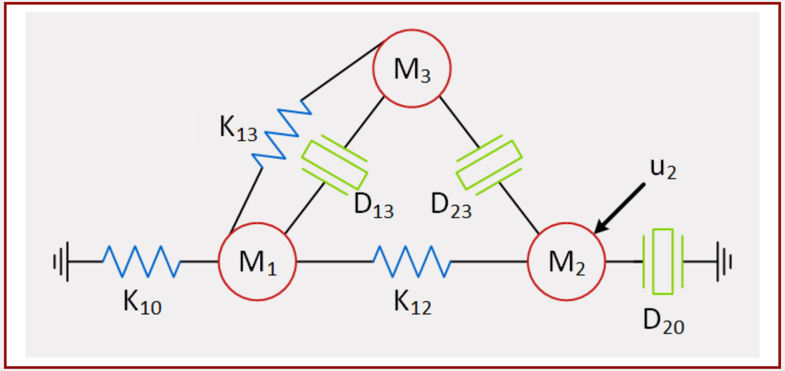
Fully captured in the module-framework



A. Dankers, 2014

Alternative models

Diffusively coupled networks:



$$\begin{bmatrix} M_1 & & \\ & M_2 & \\ & & M_3 \end{bmatrix}\begin{bmatrix} \ddot{w}_1 \\ \ddot{w}_2 \\ \ddot{w}_3 \end{bmatrix} + \begin{bmatrix} 0 & & \\ & D_{20} & \\ & & 0 \end{bmatrix}\begin{bmatrix} \dot{w}_1 \\ \dot{w}_2 \\ \dot{w}_3 \end{bmatrix} + \begin{bmatrix} K_{10} & & \\ & 0 & \\ & & 0 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

$$+ \begin{bmatrix} D_{13} & 0 & -D_{13} \\ 0 & D_{23} & -D_{23} \\ -D_{13} & -D_{23} & D_{13}+D_{23} \end{bmatrix}\begin{bmatrix} \dot{w}_1 \\ \dot{w}_2 \\ \dot{w}_3 \end{bmatrix} + \begin{bmatrix} K_{12}+K_{13} & -K_{12} & -K_{13} \\ -K_{12} & K_{12} & 0 \\ -K_{13} & 0 & K_{13} \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 0 \\ u_2 \\ 0 \end{bmatrix}$$

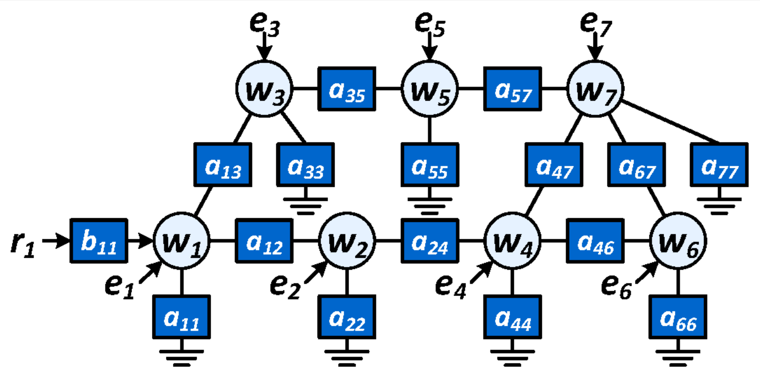
$$M_j \ddot{w}_j(t) + D_{j0}\dot{w}_j(t) + \sum_{k \neq j} D_{jk}(\dot{w}_j(t) - \dot{w}_k(t)) + K_{j0} w_j(t) + \sum_{k \neq j} K_{jk}(w_j(t) - w_k(t)) = u_j(t),$$

$$[\ \underbrace{A(p)}_{diagonal} + \underbrace{B(p)}_{Laplacian}\ ]\, w(t) = u(t)$$ $A(p), B(p)$ polynomial $\qquad p = \frac{d}{dt}$

Alternative models

Diffusively coupled networks

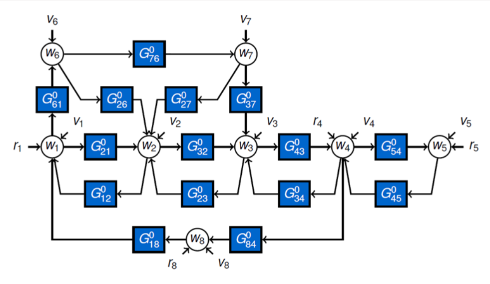
The related graph is bi-directional:



$$[\ \underbrace{Q(p)}_{diagonal} - \underbrace{P(p)}_{hollow\&symmetric}\ ]\ w(t) = u(t)$$

Q, P polynomial

Dynamic network setup – Module framework



Many data-analytics and data-driven modeling challenges appear

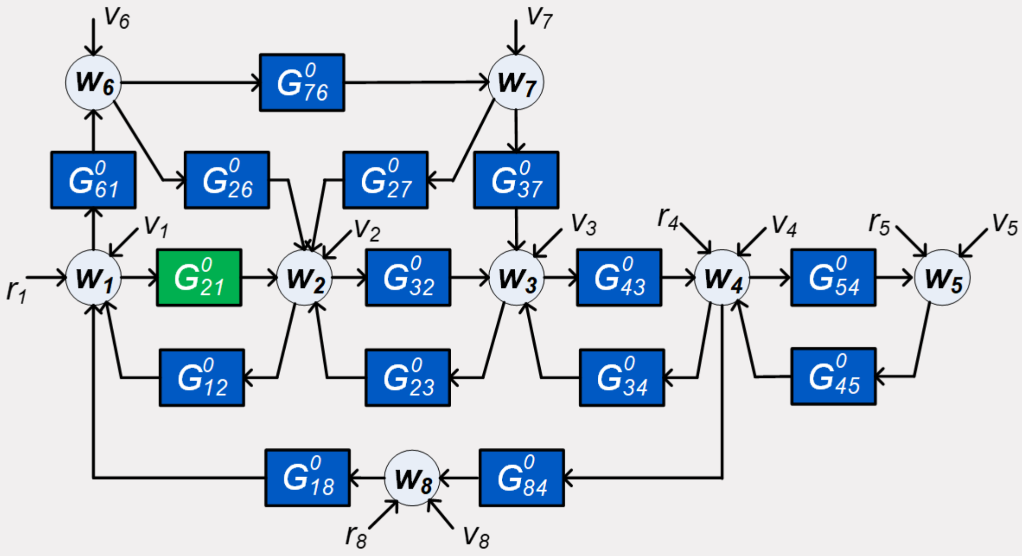
Sensor locations: $\{w_k(t)\}_{k=\cdots;}$

Actuator locations: $\{r_j(t)\}_{j=\cdots;}$

- Estimate or validate a single module/subnetwork (known topology)
- Estimate or validate the full network
- Estimate or validate the topology
- Identifiability
- Detect a fault and diagnose its location
- Exploit active probing (experiment design)
- User prior knowledge of modules/topology
- Scalable algorithms

Single module identification

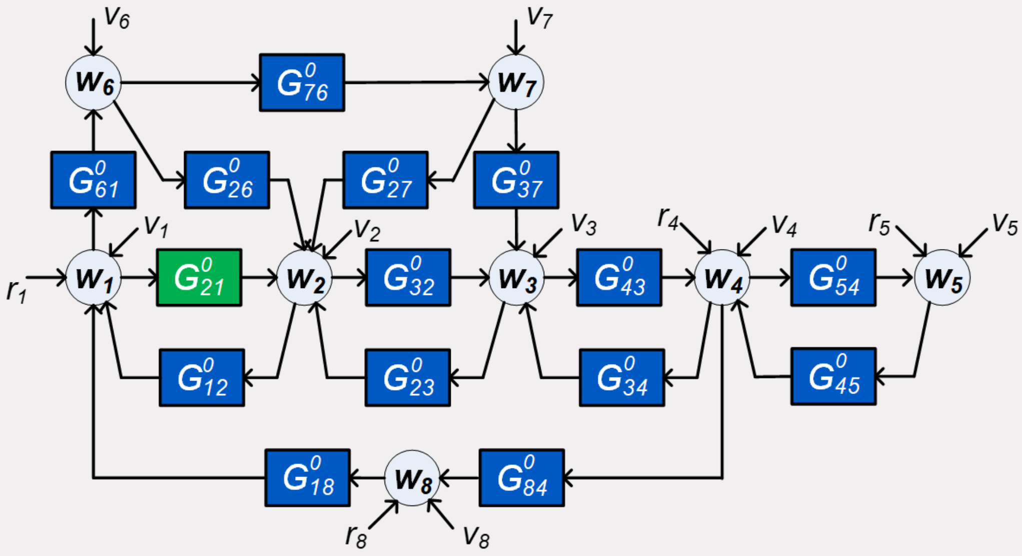
Single module identification



For a network with
known topology:

- Identify G^0_{21} on the basis of measured signals
- Which signals to measure? Preference for local measurements
- When is there enough excitation / data informativity?

Single module identification



Different types of methods:

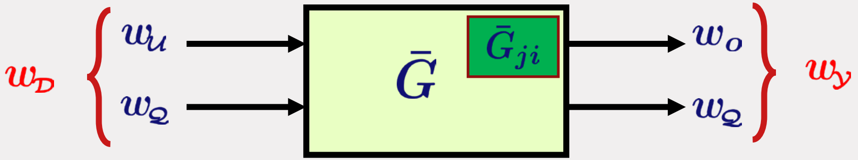
Indirect methods:

- Rely on mappings $\boldsymbol{r} \to \boldsymbol{w}$ and on sufficient excitation signals $\boldsymbol{r}$

Direct methods:

- Rely on mappings $\boldsymbol{w} \to \boldsymbol{w}$ and use excitation from both $\boldsymbol{r}$ and $\boldsymbol{v}$ signals

Single module identification

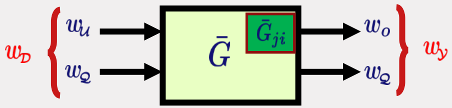


Conditions for arriving at a consistent model estimate:

1. Module invariance: $\bar{G}_{ji} = G^0_{ji}$
2. Handling of confounding variables
3. Data-informativity
4. *Technical conditions on presence of delays*

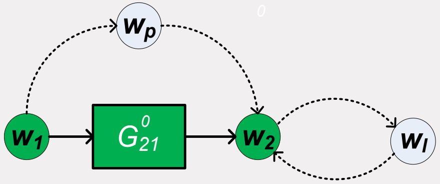
Path-based conditions on the selected signals and the network graph

Single module identification



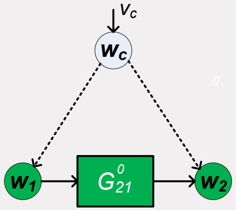
Conditions for arriving at a consistent model estimate:

1. Module invariance: $\bar{G}_{ji} = G^0_{ji}$



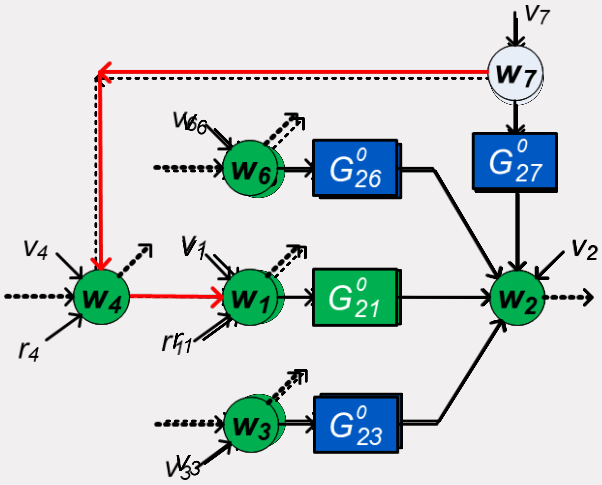
PPL condition: all parallel paths and loops around the output should be blocked by a measured node that is present in $w_{\mathcal{D}}$

2. Handling of confounding variables



No correlated disturbances between $w_{\mathcal{Y}}$ and signals in $w_{\mathcal{U}}$ that are in-neighbors of $w_{\mathcal{Y}}$

Confounding variables – solutions



Non-measurable w_7 is a confounding variable

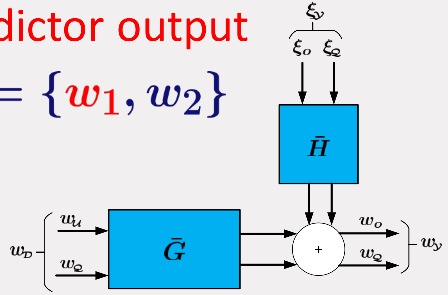
Two possible solutions:

1. Include w_4 → add predictor input

$$w_{\mathcal{D}} = \{w_1, w_3, w_4, w_6\} \quad w_{\mathcal{Y}} = \{w_2\}$$

2. Predict w_1 too → add predictor output

$$w_{\mathcal{D}} = \{w_1, w_3, w_6\} \quad w_{\mathcal{Y}} = \{w_1, w_2\}$$

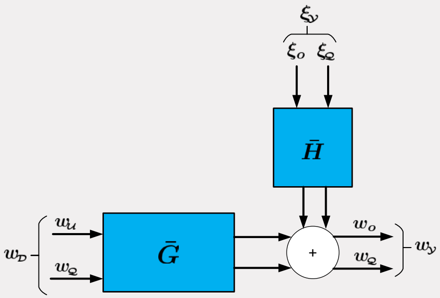


- There are degrees of freedom in choosing the predictor model

[1] A.G. Dankers et al., *Proc. IFAC World Congress*, 2017.
[2] K.R. Ramaswamy, *IEEE TAC,* 2021.

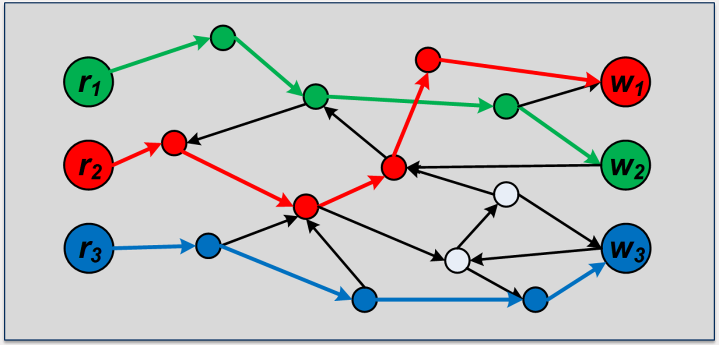
TU/e

Data informativity (path-based condition)



Data-informativity for estimating $(\bar{G}, \bar{H})$ is obtained if $\Phi_\kappa(\omega) > 0$ for almost all ω, with $\kappa = \begin{bmatrix} w_{\mathcal{D}} \\ \xi_{\mathcal{Y}} \end{bmatrix}$.

This is satisfied **generically** if there are $dim(\kappa)$ **vertex disjoint paths** from all external network signals to $\boldsymbol{\kappa}$



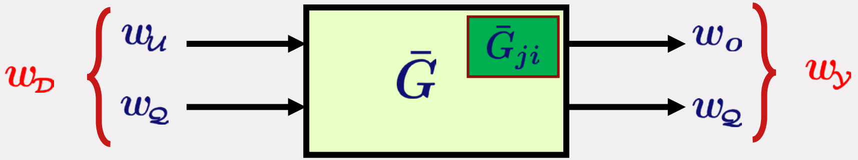
$b_{\mathcal{R}\to\mathcal{W}} = 3$

[1] Van der Woude, 1991

[2] Hendrickx, Gevers & Bazanella, CDC 2017, TAC 2019.

[3] VdH et al., CDC 2020.

Single module identification



Conditions for arriving at a consistent model estimate:

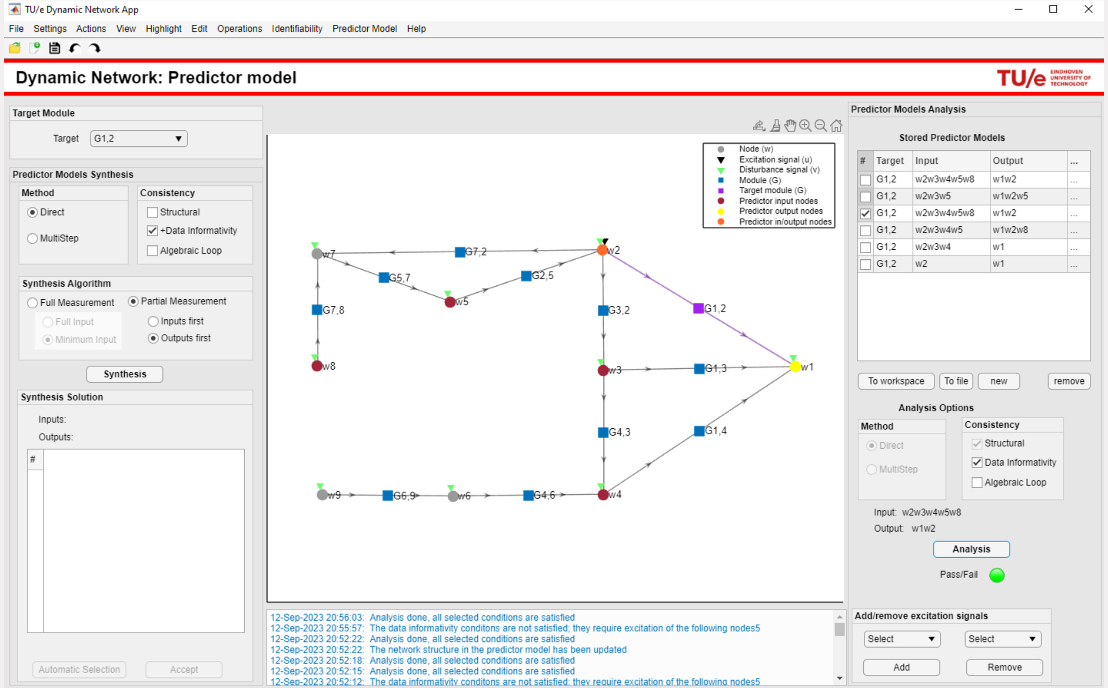
1. Module invariance: $\bar{G}_{ji} = G^0_{ji}$
2. Handling of confounding variables
3. Data-informativity
4. *Technical conditions on presence of delays*

Path-based conditions on the selected signals and the network graph

Single module identification

Different synthesis algorithms can provide predictor models that satisfy the conditions

Multiple solutions
for either full/partial
measurement



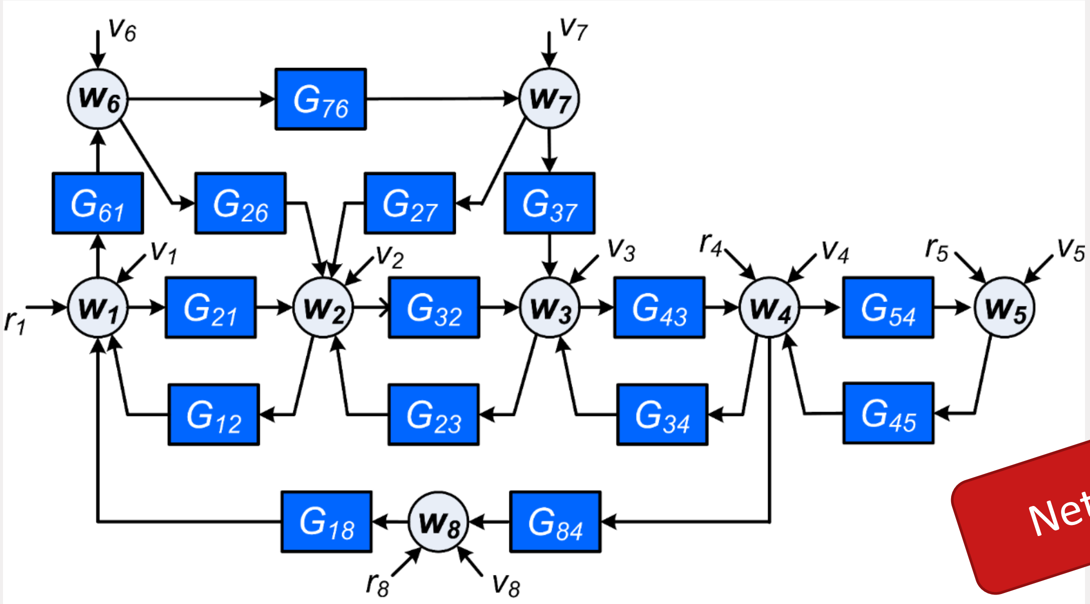
[1] K.R. Ramaswamy et al., IEEE-TAC, 2021.
[2] S.J.M. Fonken et al., CDC, 2023.
[3] Control Systems Group TU/e, SYSDYNET Toolbox for MATLAB, 2023, www.sysynet.net.

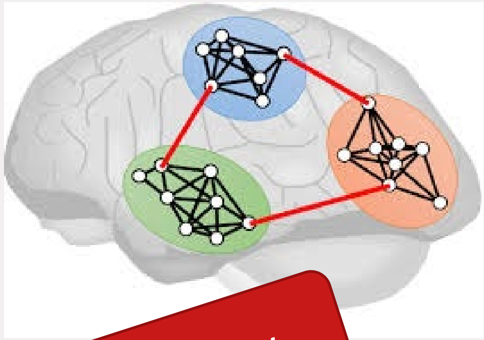
Summary single module identification

- Path-based conditions that the predictor model should satisfy
- Different algorithms for synthesizing predictor model
- Degrees of freedom in sensor / actuator placement
- Onec a predictor model is constructed, estimation comes down to a “classical’’ MISO/MIMO estimation problem

Generic network identifiability

Network identifiability

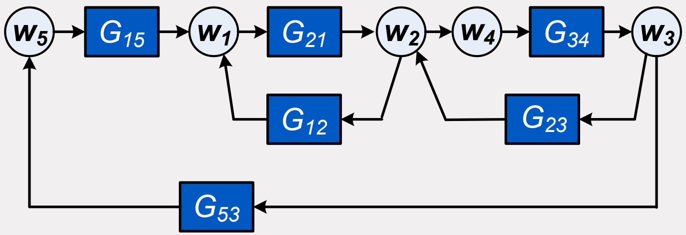


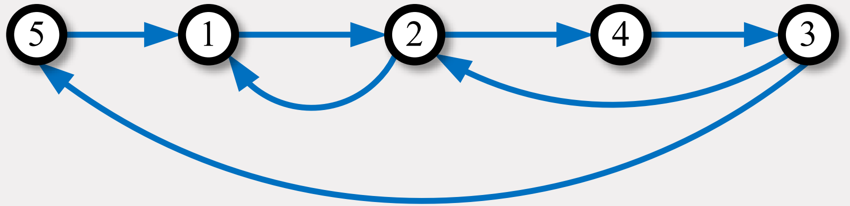


Network identifiability

Under which conditions can we estimate the topology and/or dynamics of the full network?

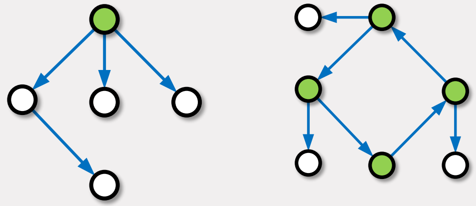
Where to allocate external excitations for network identifiability?



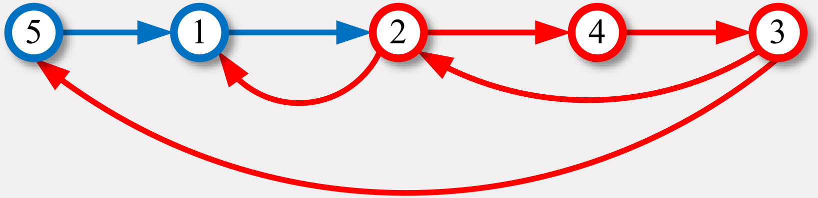


All indicated modules are parametrized

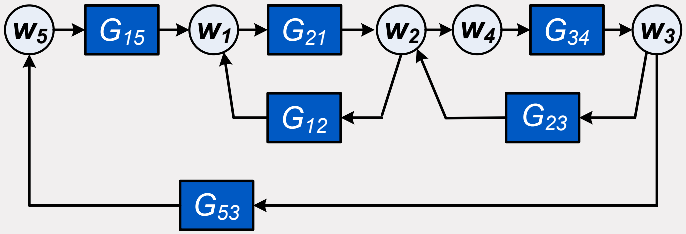
Decompose the network graph in pseudo-trees:

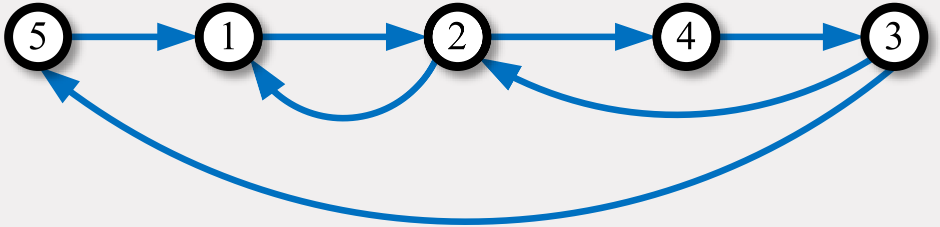


Two disjoint pseudo-trees

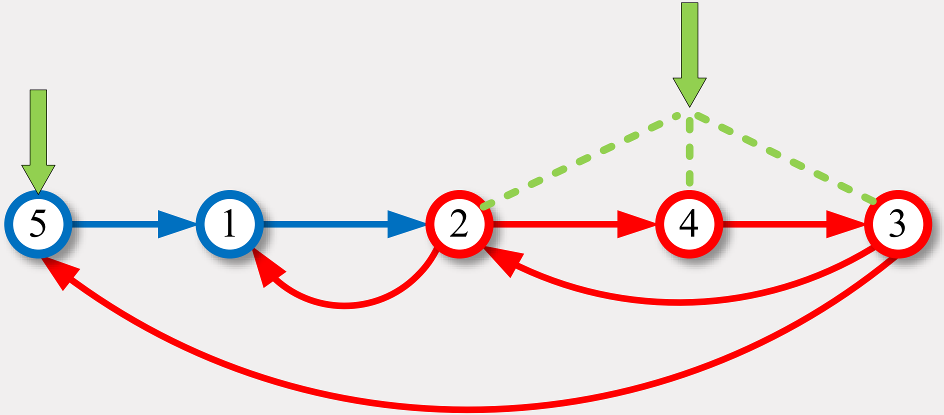


Where to allocate external excitations for network identifiability?



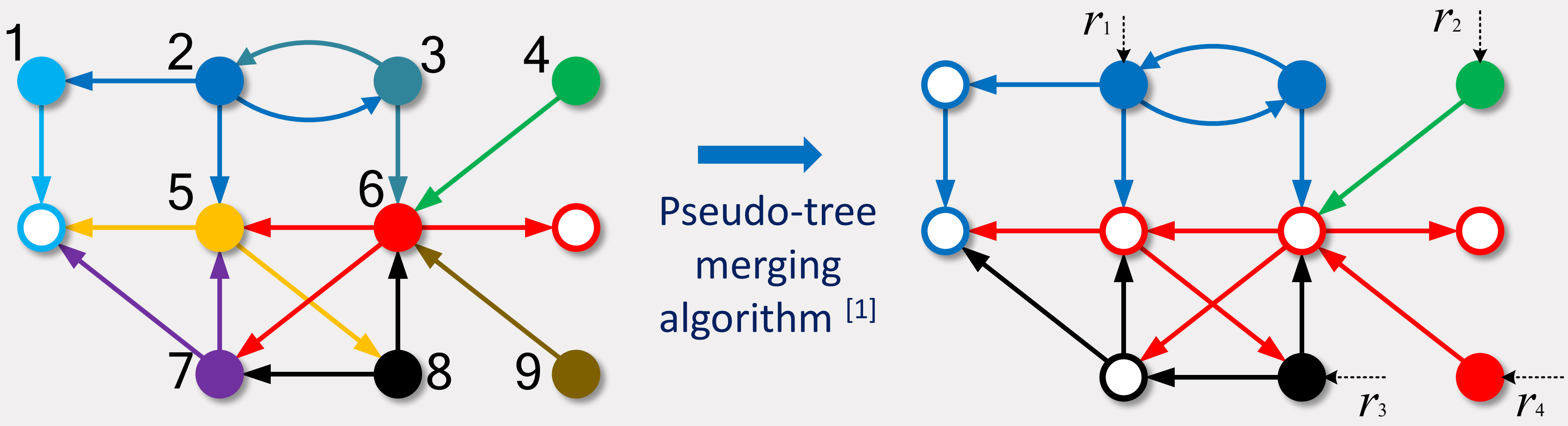


Two independent excitations
guarantee
generic network identifiability



[1] X. Cheng, S. Shi and PVdH, TAC 2022.

Where to allocate external excitations for network identifiability?



- Nodes are signals w and external signals (r, e) that are input to parametrized link
- Known (nonparametrized) links do not need to be covered

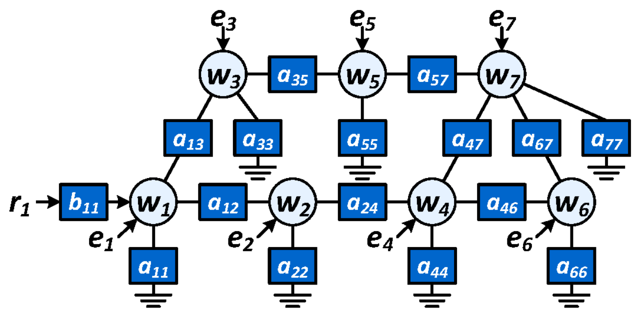
[1] X. Cheng, S. Shi and PVdH, TAC, February 2022.

Modeling considerations

Modeling considerations

- Module framework for dynamic networks (with directed graphs) most extensively developed.

- Diffusively coupled modelling framework is highly attractive for physical systems[1]



- Algorithms available for full network and single connection identification

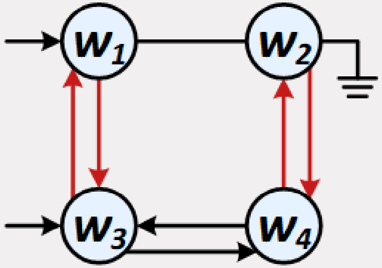
- PPL condition simplifies to: measure all neighbours

[1] E.M.M. Kivits and PVdH, TAC, June 2023; CDC 2022, IFAC 2023

Modeling considerations

Ultimate challenge

- Combine both frameworks for cyber (directed) – physical (non-directed) systems



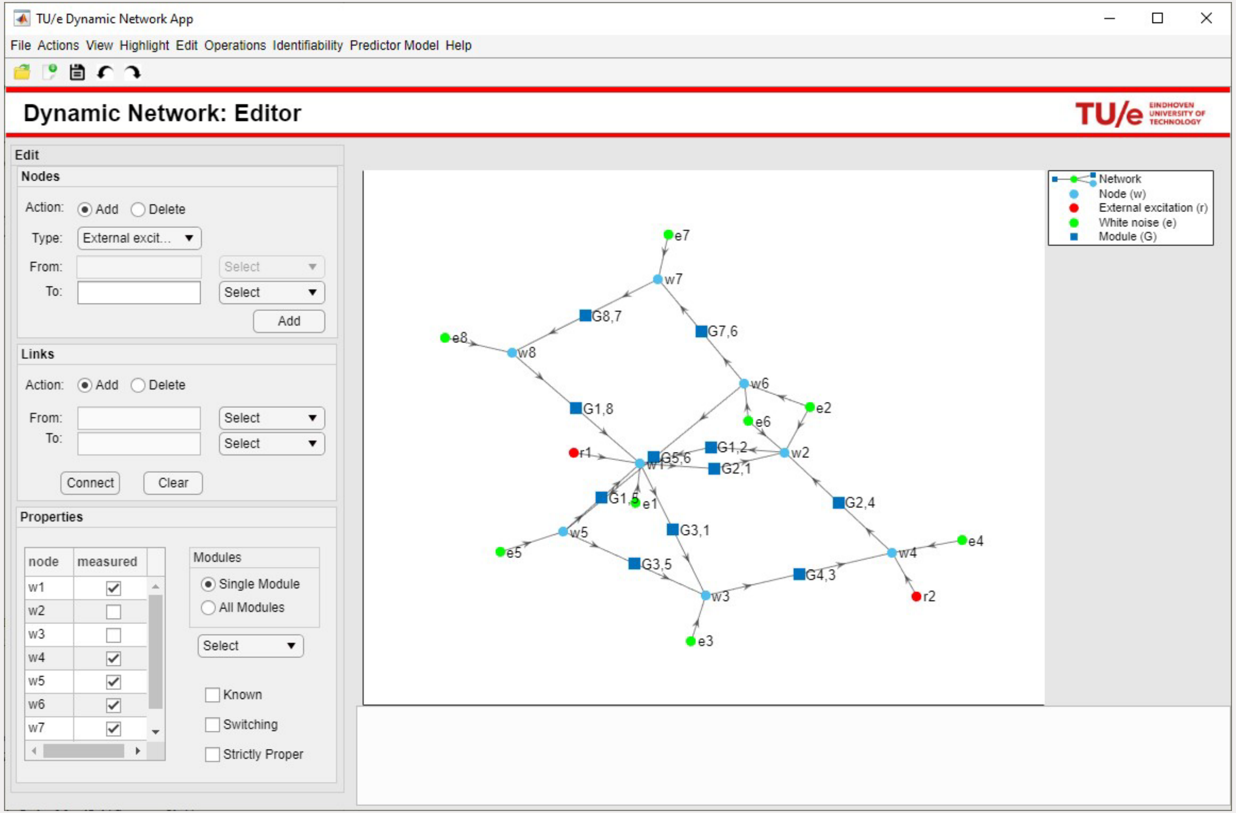
- And build the theory for related data-processing steps

Conclusions

Conclusions

- Rich framework
- Exploiting structure/topology
- Effective use of prior/physical information on model structures/ parameters
- So far, mainly developed for linear dynamics
- Free choice for actual estimation / machine learning algorithms
- Looking for attractive application opportunities in the biomed eng domain

Algorithms implemented in SYSDYNET MATLAB Toolbox



Structural analysis and operations on dynamic networks

- Edit and manipulate
- Assign properties to nodes and modules
- Immersion of nodes, PPL test
- Generic identifiability analysis and synthesis
- Predictor model selection for single module ID

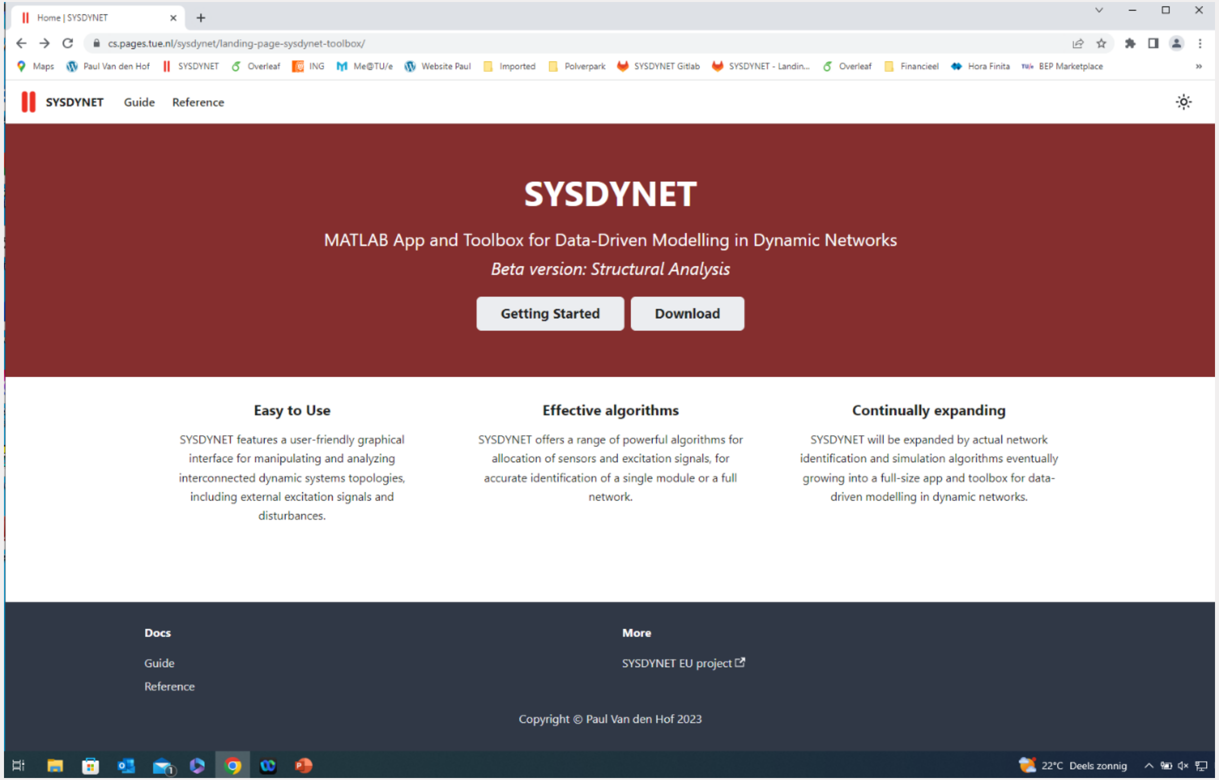
to be complemented with

- estimation algorithms for single module and network ID;
- topology estimation
- Network simulation

Beta-version to be downloaded from www.sysdynet.net

In cooperation with VersionBay embracing change TU/e

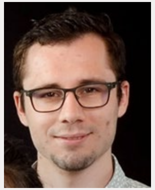
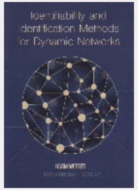
MATLAB Toolbox



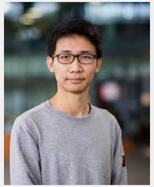
ERC SYSDYNET Team: data-driven modeling in dynamic networks

Research team:

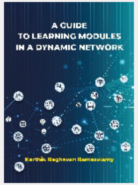

Arne Dankers

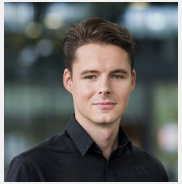
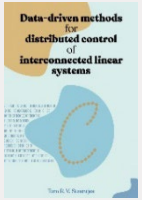
Harm Weerts




Shengling Shi

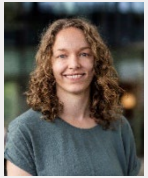



Karthik Ramaswamy

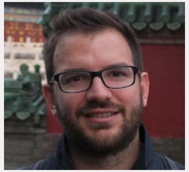
Tom Steentjes

Lizan Kivits

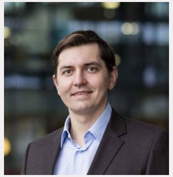


Stefanie Fonken

Xiaodong Cheng



Giulio Bottegal



Mircea Lazar

Tijs Donkers

Jobert Ludlage

Wim Liebregts

Ilja v. Oort

Mannes Dreef

+ MSc students

Co-authors, contributors and discussion partners:

Donatello Materassi, Manfred Deistler, Michel Gevers, Jonas Linder, Sean Warnick, Alessandro Chiuso, Håkan Hjalmarsson, Miguel Galrinho, Martin Enqvist, Johan Schoukens, Xavier Bombois, Peter Heuberger, Péter Csurcsia

Minneapolis, Vienna, Louvain-la-Neuve, Linköping, KTH Stockholm, Padova, Brussels, Salt Lake City, Lyon.

erc
European Research Council

TU/e

Further reading

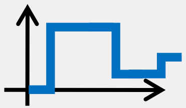
- P.M.J. Van den Hof, A. Dankers, P. Heuberger and X. Bombois (2013). Identification of dynamic models in complex networks with prediction error methods - basic methods for consistent module estimates. *Automatica*, Vol. 49, no. 10, pp. 2994-3006.
- A. Dankers, P.M.J. Van den Hof, X. Bombois and P.S.C. Heuberger (2015). Errors-in-variables identification in dynamic networks - consistency results for an instrumental variable approach. *Automatica*, Vol. 62, pp. 39-50, 2015.
- A. Dankers, P.M.J. Van den Hof, P.S.C. Heuberger and X. Bombois (2016). Identification of dynamic models in complex networks with predictior error methods - predictor input selection. *IEEE Trans. Autom. Contr.*, *61 (4),* pp. 937-952, 2016.
- H.H.M. Weerts, P.M.J. Van den Hof and A.G. Dankers (2018). Identifiability of linear dynamic networks. *Automatica*, 89, pp. 247-258, March 2018.
- H.H.M. Weerts, P.M.J. Van den Hof and A.G. Dankers (2018). Prediction error identification of linear dynamic networks with rank-reduced noise. *Automatica, 98,* pp. 256-268, December 2018.
- H.H.M. Weerts, J. Linder, M. Enqvist and P.M.J. Van den Hof (2019). Abstractions of linear dynamic networks for input selection in local module identification. *Automatica, Vol. 117*, July 2020.
- R.J.C. van Esch, S. Shi, A. Bernas, S. Zinger, A.P. Aldenkamp and P.M.J. Van den Hof (2020). A Bayesian method for inference of effective connectivity in brain networks for detecting the Mozart effect. *Computers in Biology and Medicine*, *Vol. 127*, paper 104055, December 2020.
- K.R. Ramaswamy, G. Bottegal and P.M.J. Van den Hof (2020). Learning linear models in a dynamic network using regularized kernel-based methods. *Automatica, Vol. 129,* Article 109591, July 2021.
- P.M.J. Van den Hof and K.R. Ramasmwamy (2021). Learning local modules in dynamic networks. *Proc. of Machine Learning Res., Vol. 144,* pp. 176-188.
- K.R. Ramaswamy and P.M.J. Van den Hof (2021). A local direct method for module identification in dynamic networks with correlated noise. *IEEE Trans. Automatic Control*, *Vol. 66*, no. 11, pp. 3237-3252, November 2021.
- X. Cheng, S. Shi and P.M.J. Van den Hof (2022). Allocation of excitation signals for generic identifiability of linear dynamic networks. *IEEE Trans. Automatic Control*, *Vol. 67*, no. 2, pp. 692-705, February 2022.
- S. Shi, X. Cheng and P.M.J. Van den Hof (2022). Generic identifiability of subnetworks in a linear dynamic network: the full measurement case. *Automatica, Vol. 117* (110093), March 2022.
- S.J.M. Fonken, K.R. Ramaswamy and P.M.J. Van den Hof (2022). A scalable multi-step least squares method for network identification with unknown disturbance topology. *Automatica, Vol. 141* (110295), July 2022.
- K.R. Ramaswamy, P.Z. Csurcsia, J. Schoukens and P.M.J. Van den Hof (2022). A frequency domain approach for local module identification in dynamic networks. *Automatica, Vol. 142* (110370), August 2022.
- S. Shi, X. Cheng and P.M.J. Van den Hof (2023). Single module identifiability in linear dynamic networks with partial excitation and measurement. *IEEE Trans. Automatic Control, Vol. 68(1),* pp. 285-300, January 2023.
- X.Bombois, K. Colin, P.M.J. Van den Hof and H. Hjalmarsson (2023). On the informativity of direct identification experiments in dynamical networks. *Automatica, Vol. 148* (110742), February 2023.
- E.M.M. Kivits and P.M.J. Van den Hof (2023). Identification of diffusively coupled linear networks through structured polynomial models. *IEEE Trans. Automatic Control, Vol. 68(6),* pp. 3513-3528, June 2023.
- X. Cheng, S. Shi, I. Lestas and P.M.J. Van den Hof. A necessary condition for network identifiability with partial excitation and measurement. To appear as Technical Note in *IEEE Trans. Automatic Control, Vol. 68 (11)*, November 2023.
- T.R.V. Steentjes, M. Lazar and P.M.J. Van den Hof. On a canonical distributed controller in the behavioral framework. To appear in Syst. & Control Lett., 2023.

Papers available at www.sysdynet.eu

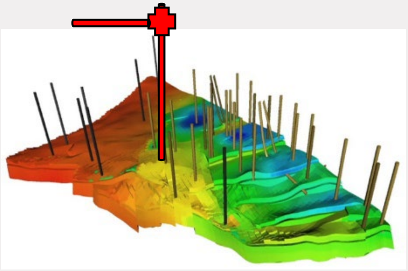
The end

Use cases

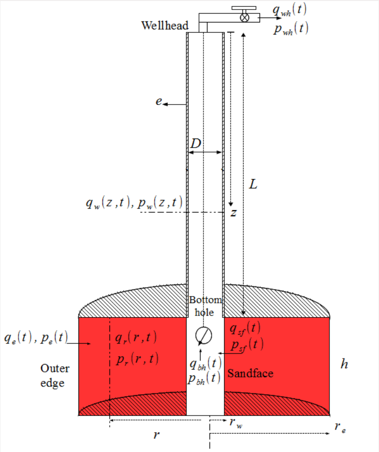
Oil reservoir modeling through well testing

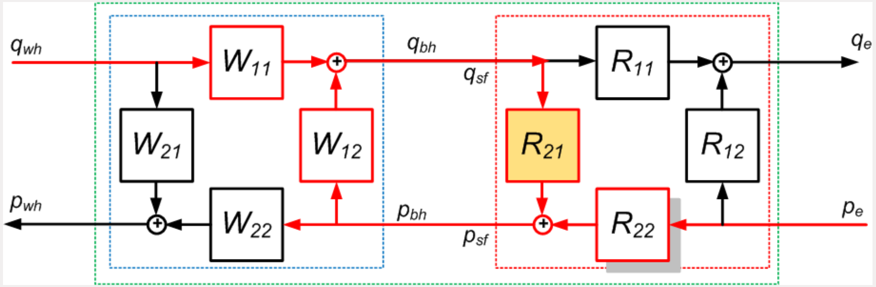


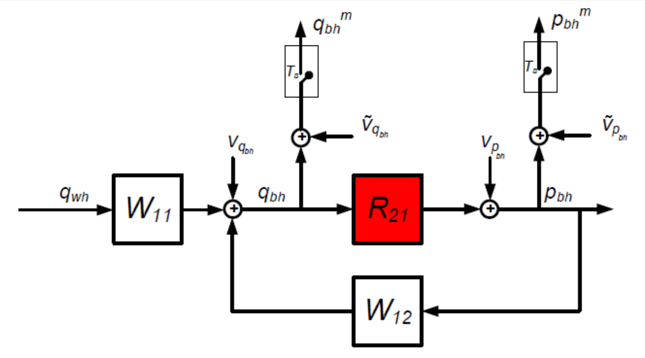
flow rate is varied



Measuring pressure



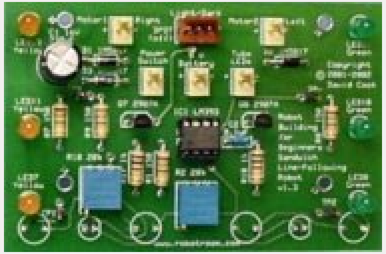


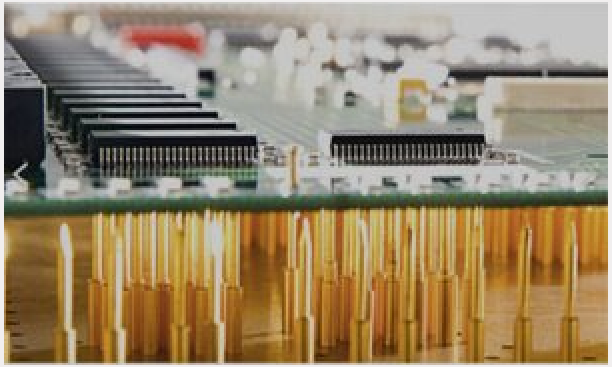


Errors-in-variables problem in closed-loop setting

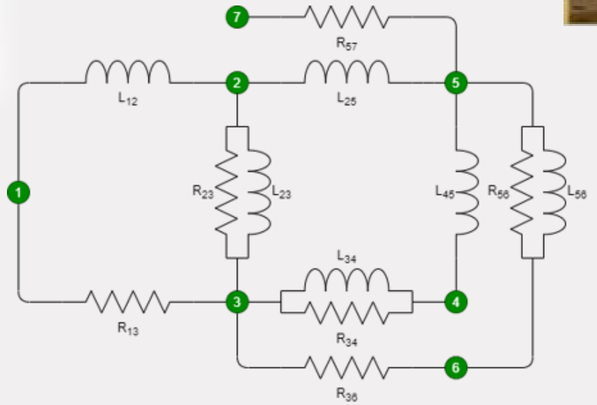
M. Mansoori, P.M.J. Van den Hof, J.D. Jansen and D. Rashtchian (2015). *SPE Journal*, Vol. 20, no. 5, pp. 1005-1027.

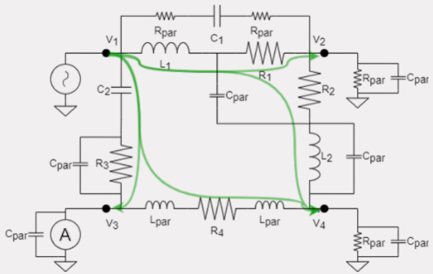
Application: Printed Circuit Board (PCB) Testing





Source: Altium

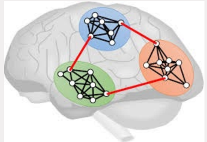




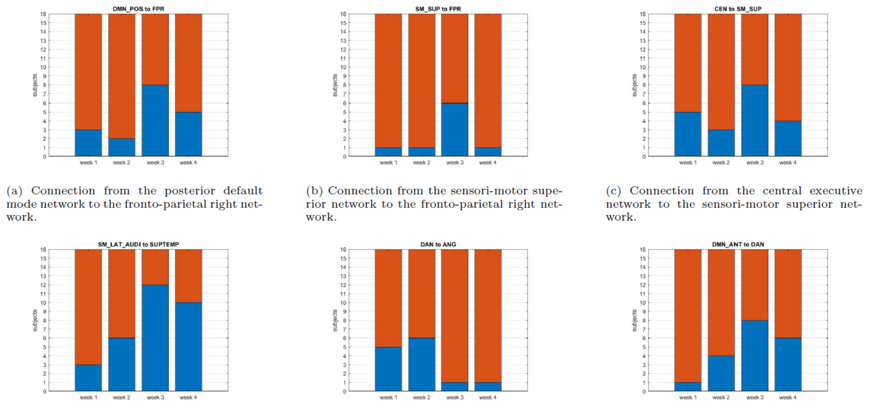
Detection of

- component failures
- parasitic effects

Neurodynamic effect of listening to Mozart music



Identifying changes in network connections in the brain, after intensely listening for one week (Sonate K448), based on fMRI data



(a) Connection from the posterior default mode network to the fronto-parietal right network.

(b) Connection from the sensori-motor superior network to the fronto-parietal right network.

(c) Connection from the central executive network to the sensori-motor superior network.

(d) Connection from the lateral sensori-motor network to the superior temporal gyrus.

(e) Connection from the dorsal attention network to the angular gyrus.

(f) Connection from the anterior default mode network to the dorsal attention network.

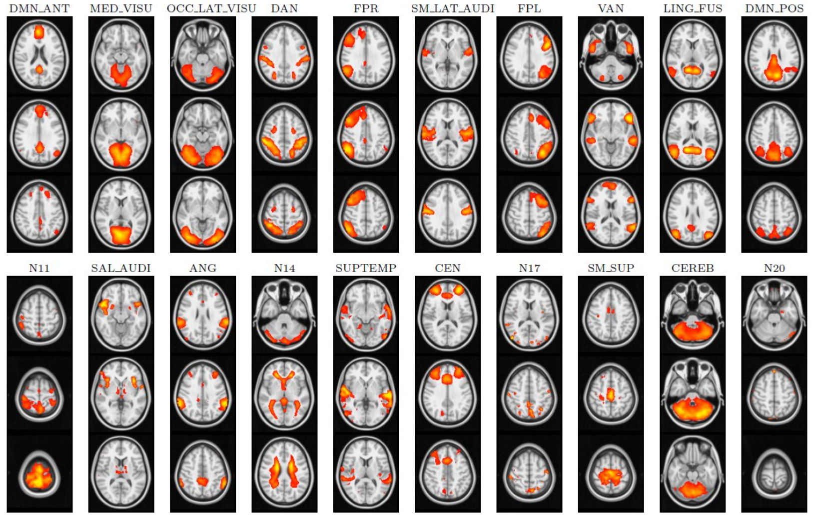
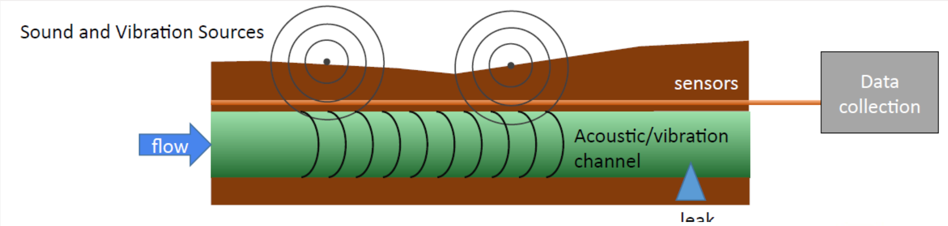


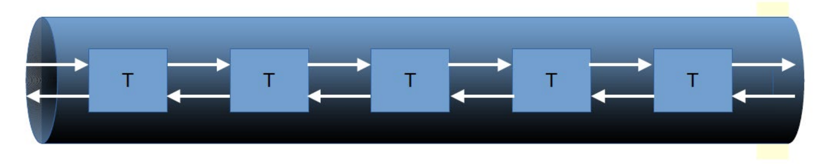
Figure 3: Spatial maps of the 20 active brain networks found through the ICA decomposition. Each image consists of 3 relevant horizontal slices of the brain, where the spatial map is indicated by the red color scale.

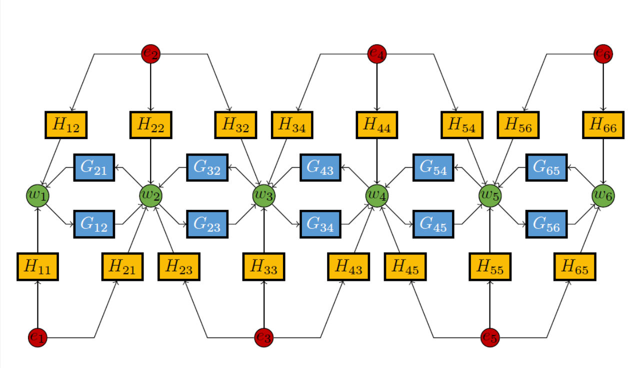
[1] R. van Esch, S. Shi, A. Bernas, S. Zinger, A. Aldenkamp, PVdH, *Computers in Biology and Medicine*, *Vol. 127*, December 2020

TU/e

Leak detection in gas pipelines with acoustic sensors







- Use operational data to detect changes in network model dynamics
- Map model changes to physical causes

Slide figures from Arne Dankers, Hifi Engineering Inc., Calgary, Canada (ECC Tutorial session 2021)

Additional projects

- Distributed identified model-based predictive control of a building climate system (Museum Hermitage Amsterdam)[1]

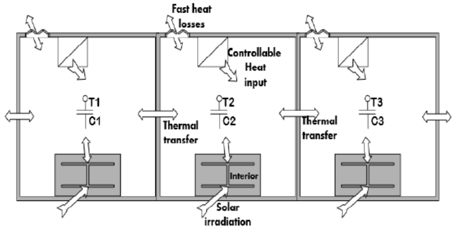


Fig. 1: Thermal model of three zones in Hermitage museum.

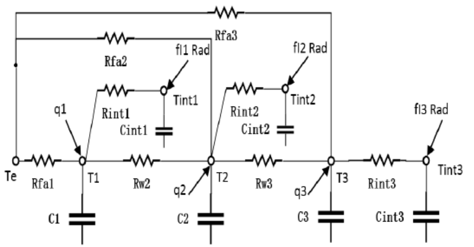


Fig. 2: Simplified thermal model.

- Data-driven modeling and control of a four-area power network[2]

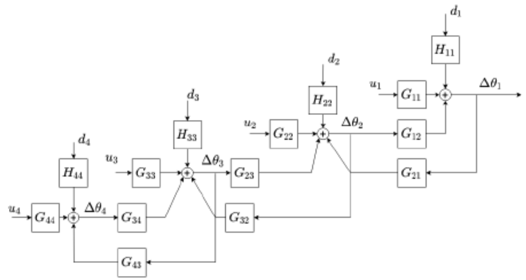


Fig. 2: Transfer function representation of four-area network.

[1] X. Chen, J.H.A. Ludlage, M. Lazar (2019)

[2] A. Anupama, T.R.V. Steentjes, M. Lazar (2021)