

Data-driven model learning in linear dynamic networks

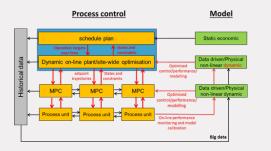
Paul Van den Hof

IFAC ACODS 2020 Advances in Control & Optimization of Dynamical Systems IIT Madras, Chennai, India, 16-19 February 2020 www.sysdynet.eu www.pvandenhof.nl p.m.j.vandenhof@tue.nl



Introduction – dynamic networks

Decentralized process control





Smart power grid

Pierre et al. (2012)

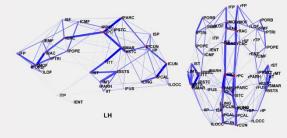


Autonomous driving



www.envidia.com

Brain network



P. Hagmann et al. (2008)

Hydrocarbon reservoirs

Mansoori (2014)



Introduction

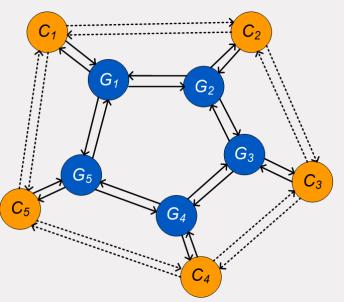
Overall trend:

- (Large-scale) interconnected systems
- With hybrid dynamics
- Distributed / multi-agent type monitoring, control and optimization problems
- Data is "everywhere", big data era
- Model-based operations require accurate/relevant models
- → Learning models from data (including physical insights when available)



Introduction

Distributed / multi-agent control:

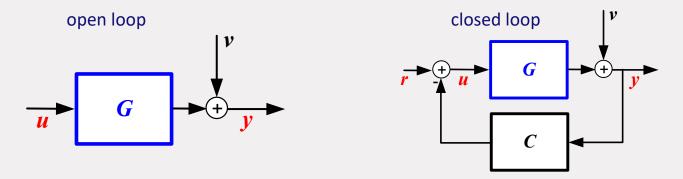


With both physical and communication links between systems G_i and controllers C_i

How to address data-driven modelling problems in such a setting?

Introduction

The classical (multivariable) identification problems^[1]:



Identify a model of G on the basis of measured signals u, y (and possibly r), focusing on *continuous LTI dynamics*.

We have to move from a simple and fixed configuration to deal with *structure* in the problem.

^[1] Ljung (1999), Söderström and Stoica (1989), Pintelon and Schoukens (2012)

Early contributors

Topology detection: Materassi, Innocenti, Salapaka, Yuan, Stan, Warnick, Goncalves, Sanandaji, Vincent, Wakin, Chiuso, Pillonetto exploring Granger causality, Bayesian networks, Wiener filters

Subspace algorithms for **spatially distributed systems** with identical modules (Fraanje, Verhaegen, Werner), or non-identical ones (Torres, van Wingerden, Verhaegen, Sarwar, Salapaka, Haber)

Here: focus on structural aspects in identification setups.



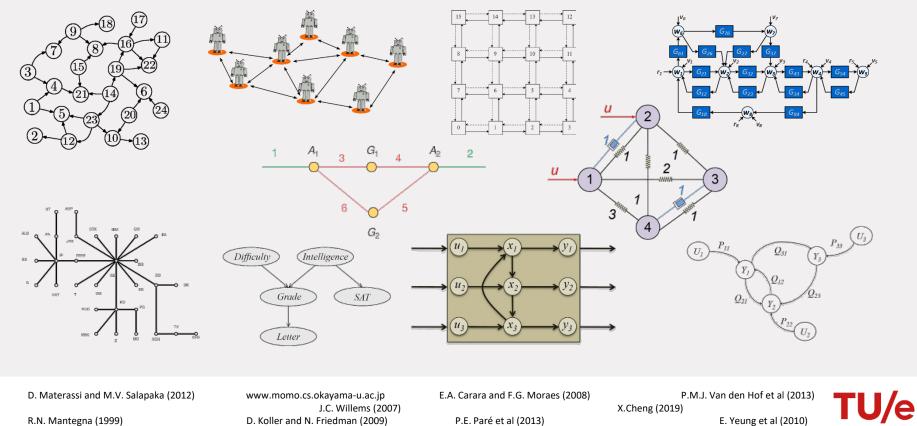
Contents

- Introduction and motivation
- How to model a dynamic network?
- Single module identification
- Global network identification
- Physical networks
- Extensions Discussion



Dynamic networks for data-driven modeling

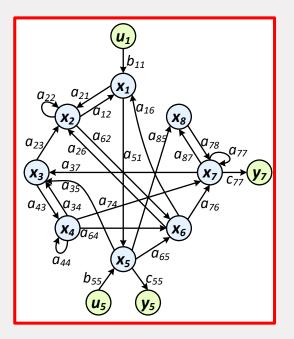
Network models



R.N. Mantegna (1999)

9

Network models



 $r_1 \qquad v_3 \qquad v_5$ $w_1 \qquad G_{21} \qquad w_2 \qquad G_{32} \qquad w_3 \qquad w_4 \qquad G_{54} \qquad w_5$ $G_{21} \qquad G_{32} \qquad G_{32} \qquad G_{45} \qquad G_{45} \qquad G_{45} \qquad G_{45} \qquad G_{45} \qquad G_{45} \qquad G_{15} \qquad G_{15}$

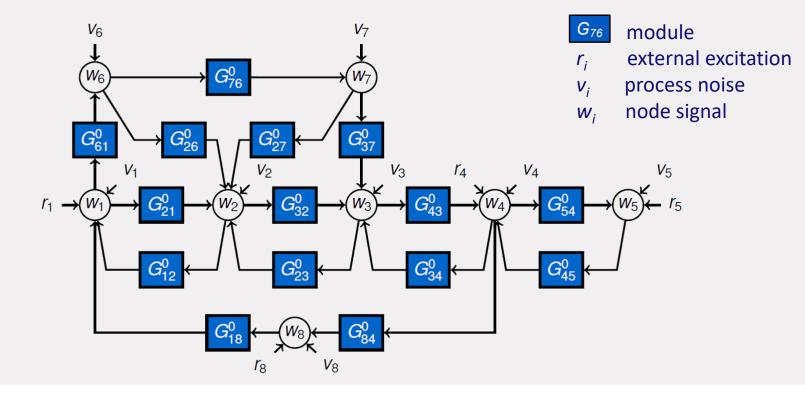
Module representation ^[2]

State space representation ^[1]

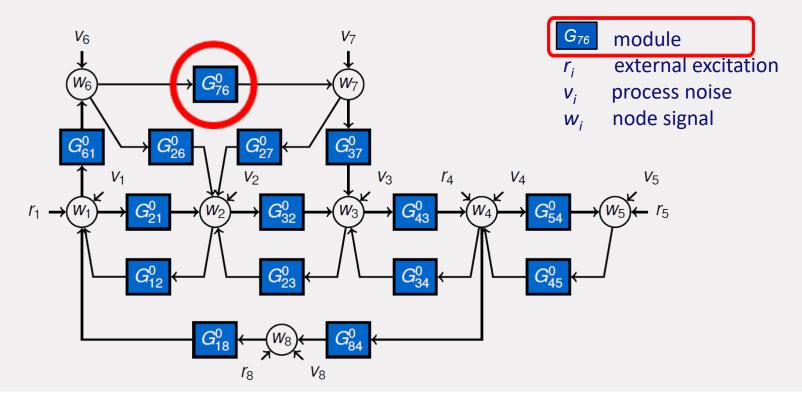
[1] Goncalves, Warnick, Sandberg, Yeung, Yuan, Scherpen,...

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

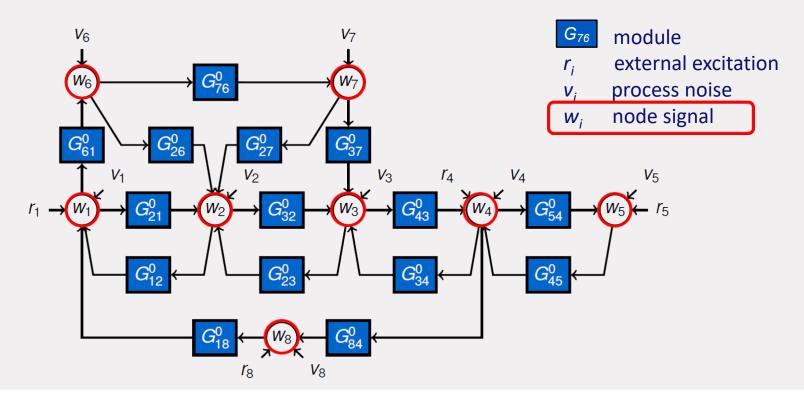




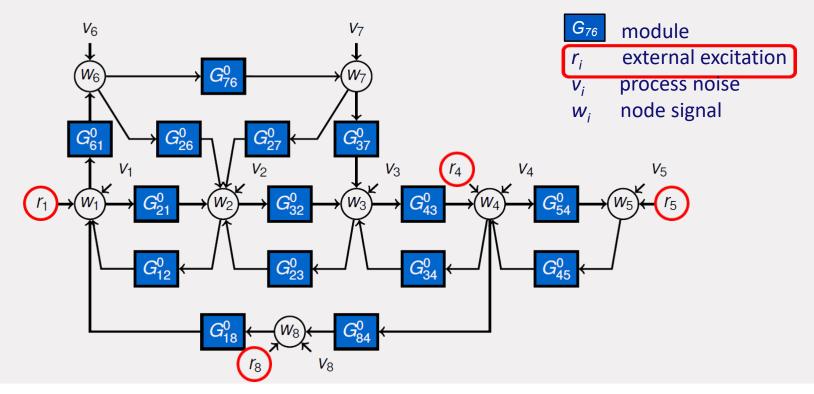
TU/e



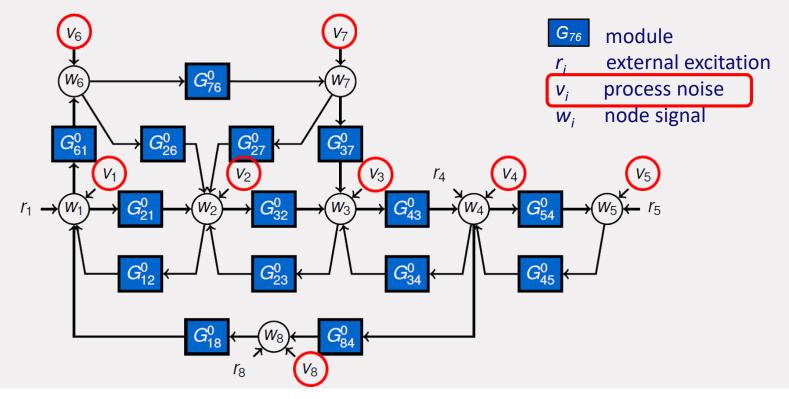
TU/e

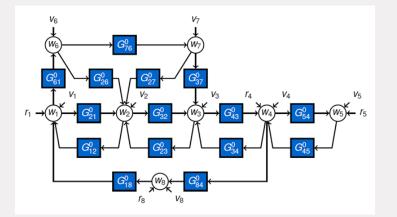












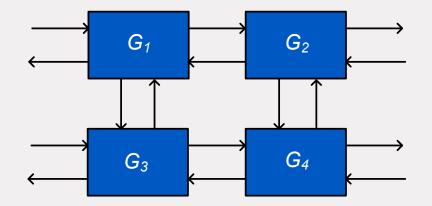
Assumptions:

- Total of L nodes
- Network is well-posed and stable
- Modules are dynamic LTI, may be unstable
- Disturbances are stationary stochastic and can be correlated

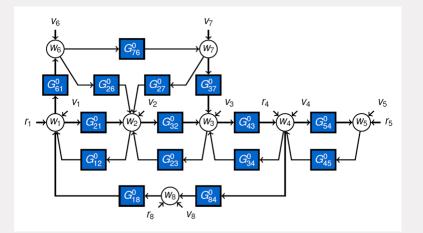
$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} = \begin{bmatrix} 0 & G_{12}^0 & \cdots & G_{1L}^0 \\ G_{21}^0 & 0 & \cdots & G_{2L}^0 \\ \vdots & \ddots & \ddots & \vdots \\ G_{L1}^0 & G_{L2}^0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} + R^0 \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_K \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ r_K \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_L \end{bmatrix}$$
$$w(t) = G^0(q)w(t) + R^0(q)r(t) + v(t)$$

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

Setup covers the situation of bilaterally coupled (physical) systems:





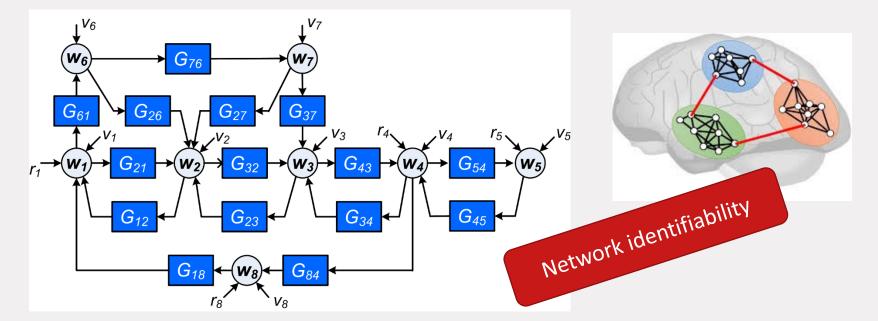


Many new data-driven modeling questions can be formulated

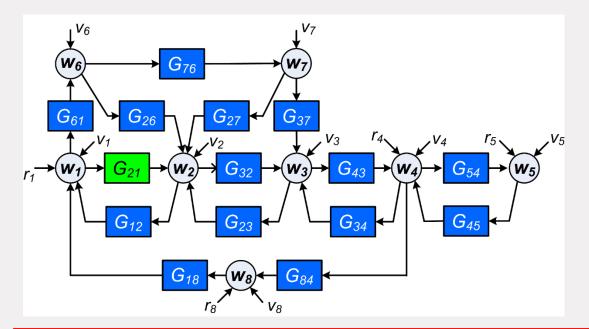
Measured time series:

 $\{w_i(t)\}_{i=1,\dots L}; \ \{r_j(t)\}_{j=1,\dots K}$



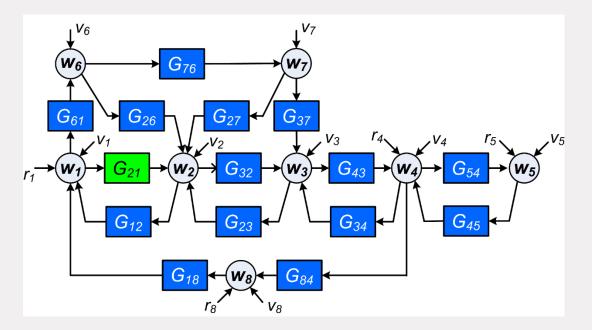


Under which conditions can we estimate the topology and/or dynamics of the full network?



How/when can we learn a local module from data (with known/unkown network topology)? Which signals to measure?

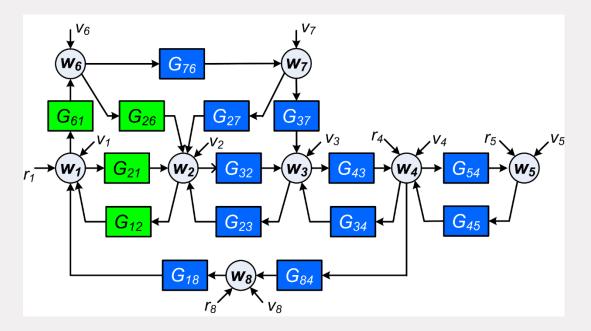




Where to optimally locate sensors and actuators?

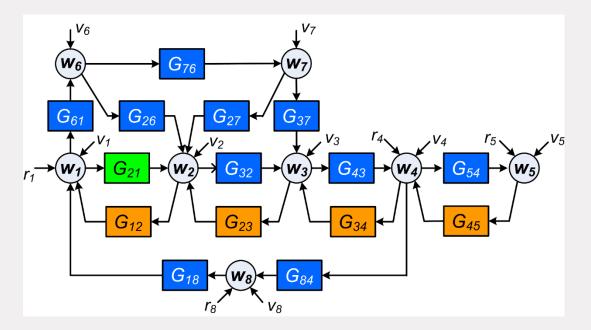


21

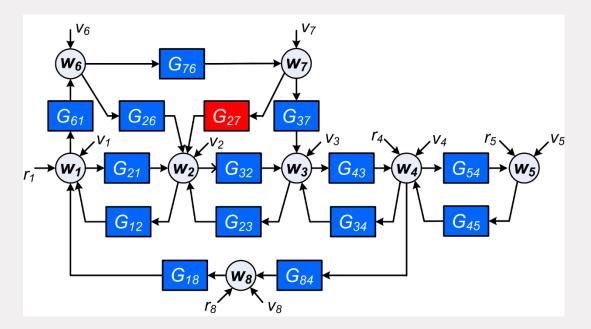


Same questions for a subnetwork



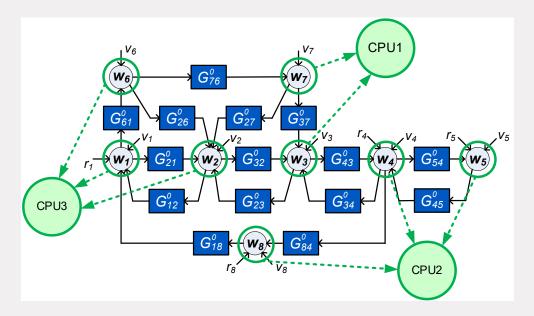


How can we benefit from known modules?



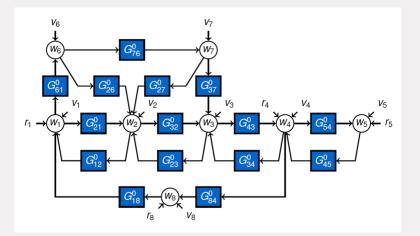
Fault detection and diagnosis; detect/handle nonlinear elements





Can we distribute the computations?



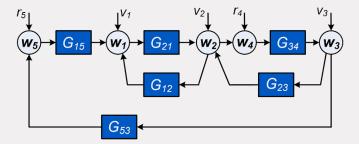


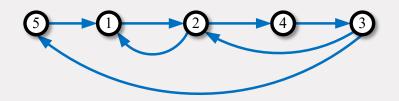
Measured time series: $\{w_i(t)\}_{i=1,\cdots L}; \ \{r_j(t)\}_{j=1,\cdots K}$

Many new data-driven modeling questions can be formulated

- Identification of a local module (known topology)
- Identification of the full network
- Topology estimation
- Identifiability
- Sensor and excitation allocation
- Fault detection
- User prior knowledge of modules
- Distributed identification
- Scalable algorithms

Dynamic network setup - graph

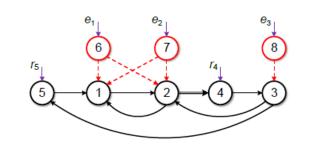




Nodes are vertices; modules/links are edges

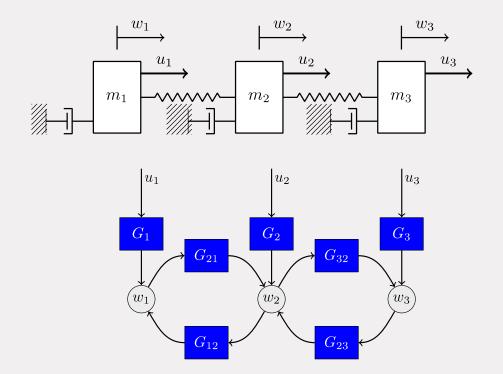
Extended graph:

including the external signals and disturbance correlations



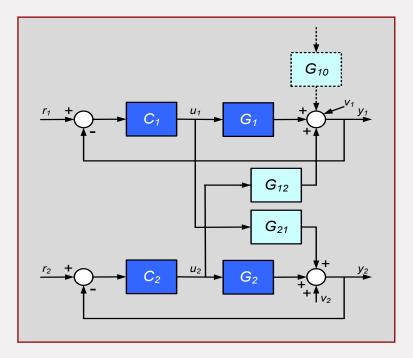


Application: Networks of (damped) oscillators



- Power systems, vehicle platoons, thermal building dynamics, ...
- Spatially distributed
- Bilaterally coupled
- No central coordination \Rightarrow local identification problems

Single module identification - Example



Decentralized MPC 2 interconnected MPC loops

Target: Identify interaction dynamics G_{21}, G_{12}

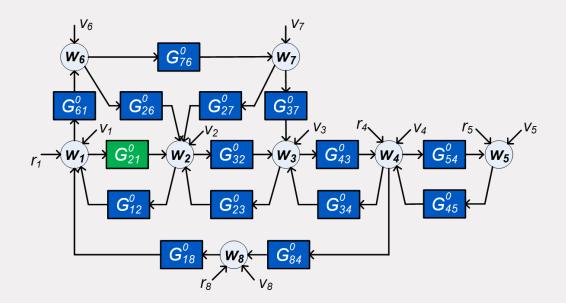
Addressed by Gudi & Rawlings (2006) for the situation $G_{12} = 0$ (no cycles)



Contents

- Introduction and motivation
- How to model a dynamic network?
- Single module identification
- Global network identification
- Physical networks
- Extensions Discussion

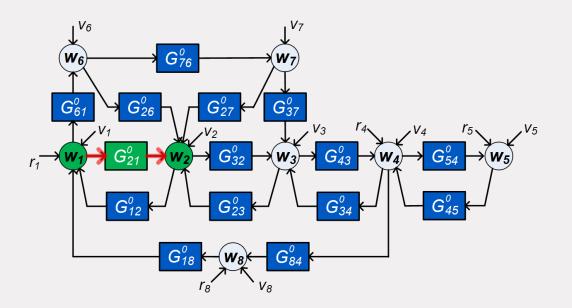




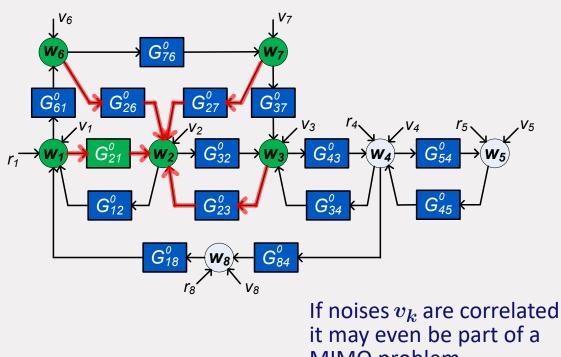
For a network with known topology:

- Identify G⁰₂₁ on the basis of measured signals
- Which signals to measure? Preference for local measurements
- When is there enough excitation / data informativity?

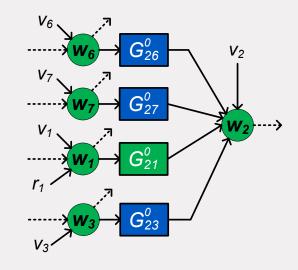




Naïve approach: identify based on w_1 and w_2 : in general does not work.

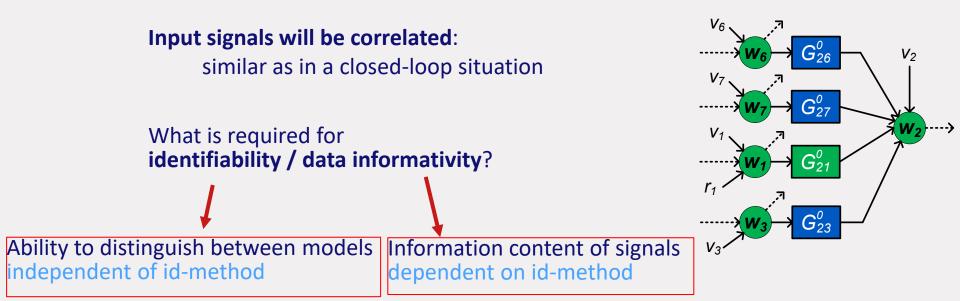


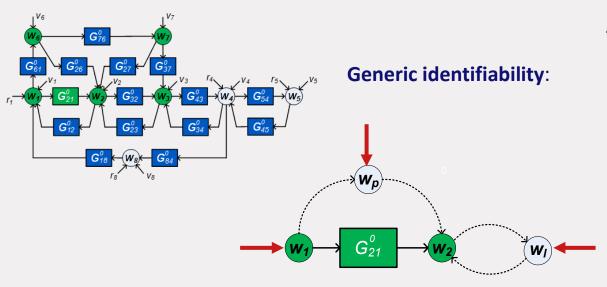
Identifying G_{21}^0 is part of a 4-input, 1-output problem



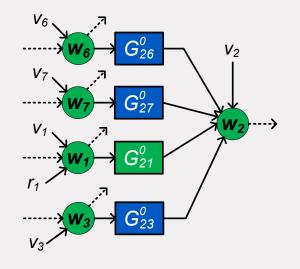
MIMO problem

Identifying G_{21}^0 is part of a 4-input, 1-output problem





Identifying G_{21}^0 is part of a 4-input, 1-output problem



All parallel paths, and loops around the output, plus input w_1 should have an independent external signal r or v

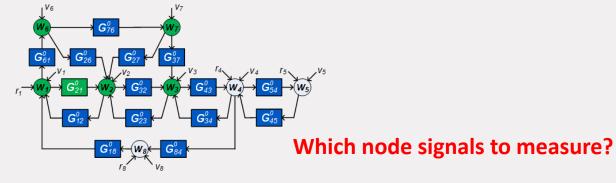
Weerts et al., Automatica 2018, CDC 2018
Bazanella et al. CDC2017; Hendrickx et al., IEEE-TAC, 2019.

[3] Dankers et al., TAC 2016[4] Shi et al., IFAC 2020 submitted.



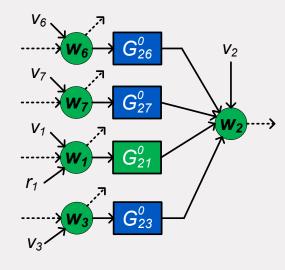
36

Single module identification



Dependent on

- *v* signals uncorrelated or not
 - Excitation conditions satisfied through *r* and/or *v*-signals



Typical solution: • One additional measured signal for each parallel path/loop

- Additional signals if excitation is through v signals
- Variation in available algorithms / options

Dankers et al., TAC 2016
Hendrickx et al., IEEE-TAC, 2019.

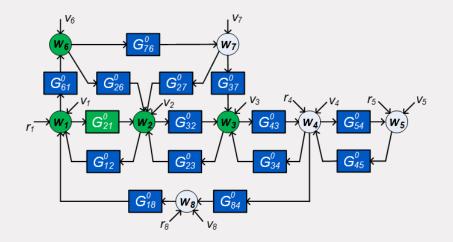
[3] Gevers et al. SYSID 2018[4] Bazanella et al., CDC2019

[5] PVdH, Ramaswamy, CDC2019[6] Shi et al., IFAC 2020 submitted.

Single module identification

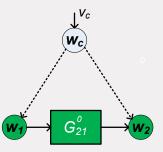
one signal per parallel path/loop:

With a 3-input, 1 output model we can consistently identify G_{21}^0



When excitation is through disturbance signals $oldsymbol{v}$:

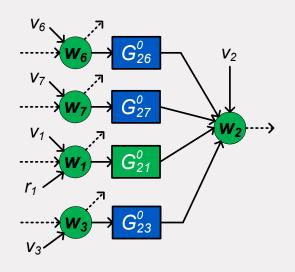
- dealing with confounding variables, ^{[1][2]} i.e. correlated disturbances on inputs and outputs
- can be addressed by adding inputs/outputs to the estimation problem ^[3]





Single module identification

Typical solution:

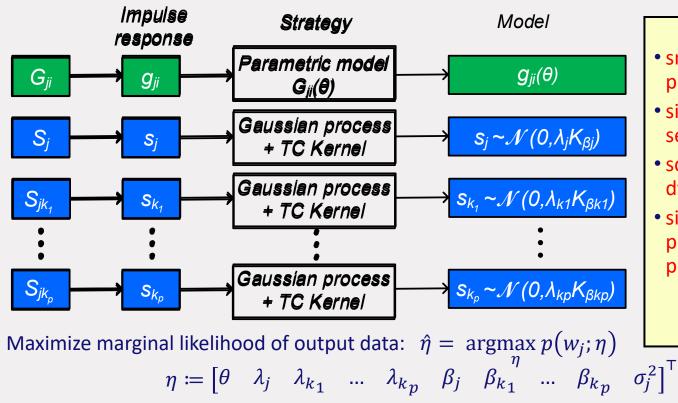


- MISO (sometimes MIMO) estimation problem
- to be solved by any (closed-loop) identification algorithm, e.g. direct/indirect method

Machine learning in local module identification e_i MISO identification with all modules parameterized Brings in two major problems : Vi Large number of parameters to estimate Gii Wi Model order selection step for each module (CV, AIC, BIC) W_{k_1} For 5 modules, combinations = 244,140,625 W_{k_2} **Increases variance HSANCE** Computationally challenging W_{k} We need only the target module. No NUISANCE!



Machine learning in local module identification

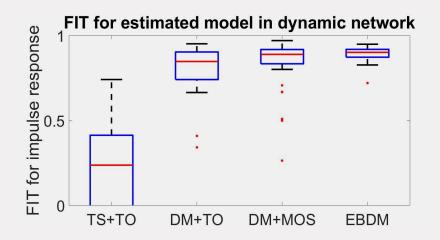


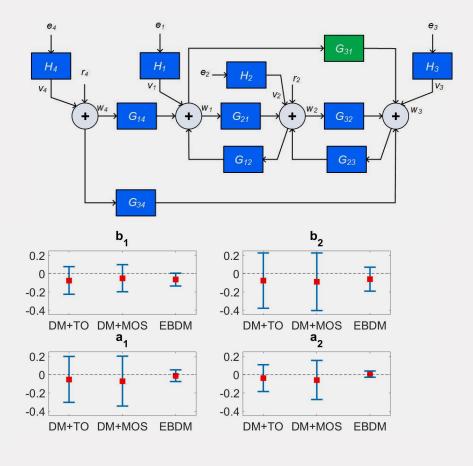
- smaller no. of parameters
- simpler model order selection step
- scalable to large dynamic networks
- simpler optimization problems to estimate parameters

Everitt et al., Automatica 2017.
K.R. Ramaswamy et al., CDC 2018.

Numerical simulation

- Identify G₃₁ given data
- ▶ 50 independent MC simulation
- Data = 500





Summary single module identification

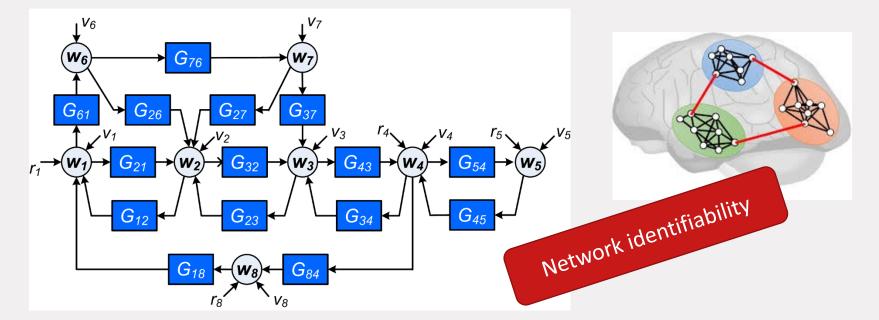
- Path-based conditions for **network identifiability** (where to excite?)
- Graph tools for checking conditions
- Degrees of freedom in selection of measured signals sensor selection
- Methods for **consistent** and **minimum variance** module estimation, and effective (scalable) algorithms
- A priori known modules can be accounted for



Contents

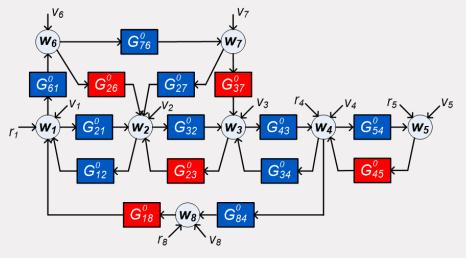
- Introduction and motivation
- How to model a dynamic network?
- Single module identification
- Global network identification
- Diffusively coupled physical networks
- Extensions Discussion

Full network identification



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

Network identifiability



blue = unknown red = known

Question: Can different dynamic networks be *distinguished* from each other from measured signals *w*, *r*?



Network identifiability

The identifiability problem:

The network model:

$$w(t) = G(q)w(t) + R(q)r(t) + \underbrace{H(q)e(t)}_{v(t)}$$

can be transformed with any rational P(q):

 $\boldsymbol{P(q)}\boldsymbol{w(t)} = \boldsymbol{P(q)}\{\boldsymbol{G(q)}\boldsymbol{w(t)} + \boldsymbol{R(q)}\boldsymbol{r(t)} + \boldsymbol{H(q)}\boldsymbol{e(t)}\}$

to an equivalent model:

 $w(t) = ilde{G}(q)w(t) + ilde{R}(q)r(t) + ilde{H}(q)e(t)$

Nonuniqueness, unless there are structural constraints on G, R, H.

Weerts, Linder et al., Automatica, 2019, provis. accepted.
Bottegal et al., SYSID 2017



Network identifiability

Consider a network model set:

 $\mathcal{M} = \{(G(heta), R(heta), H(heta))\}_{ heta \in \Theta}$

representing structural constraints on the considered models:

- modules that are fixed and/or zero (topology)
- locations of excitation signals
- disturbance correlation

48

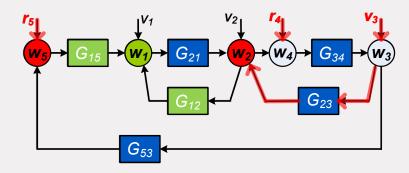
Generic identifiability of \mathcal{M} :

- There do not exist distinct equivalent models
- for almost all models in the set.

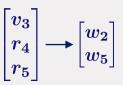


Example 5-node network

Conditions for identifiability **—** rank conditions on transfer function



Full row rank of

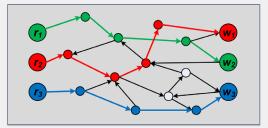


For the **generic case**, the rank can be calculated by a graph-based condition^{[1],[2]}:

Generic rank = number of vertex-disjoint paths

2 vertex-disjoint paths \rightarrow full row rank 2

The rank condition has to be checked for all nodes.



[1] Van der Woude, 1991

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

Synthesis solution for network identifiability

Allocating external signals for generic identifiability:

Cover the graph of the network model set by a set of disjoint pseudo-trees
Pseudo-trees:

Tree with root in green

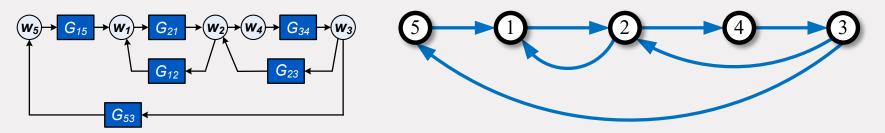
Cycle with outgoing trees; Any node in cycle is root

Edges are disjoint and all out-neighbours of a node are in the same pseudo-tree

2. Assign an independent external signal (r or e) at a root of each pseudo-tree.

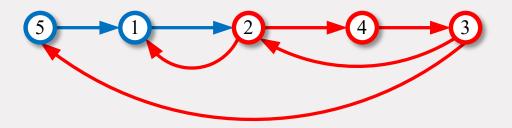
This guarantees generic identifiability of the model set.

Where to allocate external excitations for network identifiability?



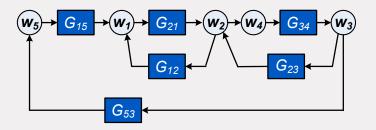
All indicated modules are parametrized

Two disjoint pseudo-trees

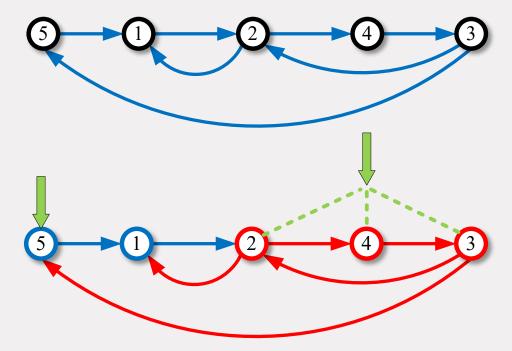




Where to allocate external excitations for network identifiability?



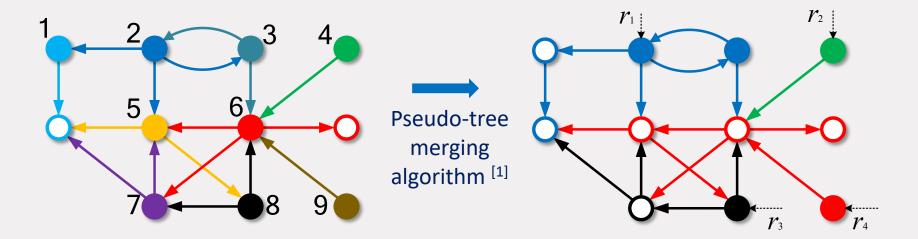
Two independent excitations guarantee generic network identifiability





[1] X. Cheng, S. Shi and PVdH, CDC 2019.

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, CDC 2019.

Summary identifiability of full network

Identifiability of network model sets is determined by

- Presence and location of external signals, and
- Correlation of disturbances
- Topology of parametrized modules
- Graphic-based tool for synthesizing allocation of external signals

Extensions:

• Situations where not all node signals are measured ^[1]



[1] Bazanella, CDC 2019.

Algorithms for identification of full network

(Prediction error) identification methods will typically lead to large-scale **non-convex** optimization problems

Convex relaxation algorithms are being developed^[1] as well as machine learning tools

Topology identification

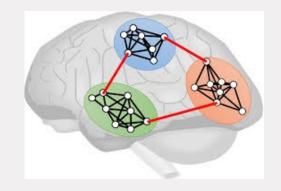
- Topology resulting from full dynamic model
- Alternative: non-parametric models (Wiener filters ^[1]) or kernel-based approaches ^{[2][3]}
- modeling module dynamics by Gaussian processes,

kernel with 2 parameters for each dynamic module

• Optimizing likelihood of the data as function of parameters and topology:

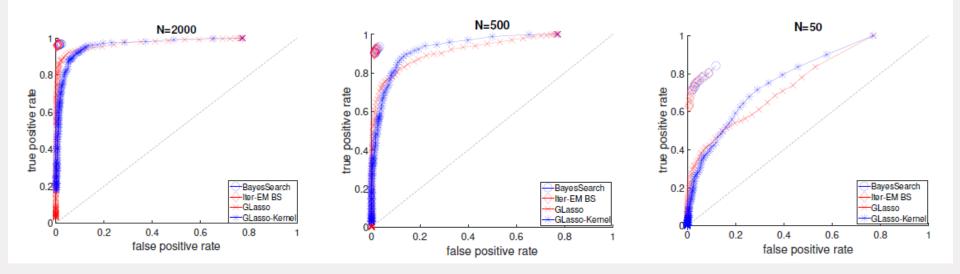
 $p(\{w(t)\}_{t=1}^N| heta,\mathcal{G})|$

• Forward-backward search over topologies + empirical Bayes (EM) for parameters





Topology identification



50 MC realizations of network with 6 nodes.



[1] Shi, Bottegal, PVdH, ECC 2019

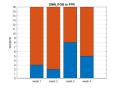
Neurodynamic effect of listening to Mozart music

DMN_ANT

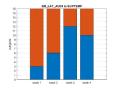
MED_VISU

OCC_LAT_VISU DAN

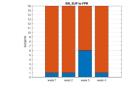
Identifying changes in network connections in the brain, after intensely listening for one week



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



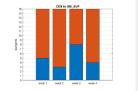
(d) Connection from the lateral sensori-motor network to the superior temporal gyrus.



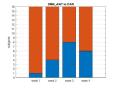
(b) Connection from the sensori-motor superior network to the fronto-parietal right network.



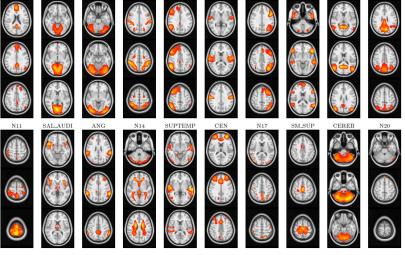
(e) Connection from the dorsal attention network to the angular gyrus.



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



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



FPR

SM_LAT_AUDI FPL

VAN

LING_FUS

DMN_POS

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.



Algorithms for identification of full network

Particular feature for larger networks:

Modeling disturbances as a **reduced rank process**: (cf dynamic factor analysis^[1])

Consequences for **estimation**^[3]:

H di

 $\dim(e) < \dim(v)$

- Optimization becomes a constrained quadratic problem with ML properties for Gaussian noise
- Reworked Cramer Rao lower bound
- Some parameters can be estimated variance free \rightarrow regularization effect

59





Contents

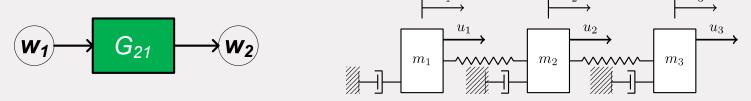
- Introduction and motivation
- How to model a dynamic network?
- Single module identification
- Global network identification
- Physical networks
- Extensions Discussion



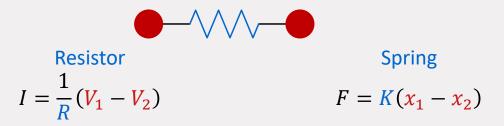
Physical networks

Back to the basics of physical interconnections

In connecting physical systems, there is often no predetermined direction of information [1]

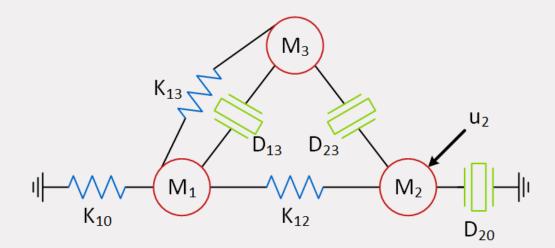


Example: resistor / spring connection in electrical / mechanical system:



Difference of node signals drives the interaction: diffusive coupling

Diffusively coupled physical network



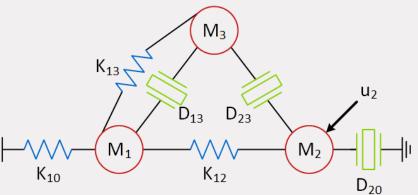
Equation for node *j*:

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



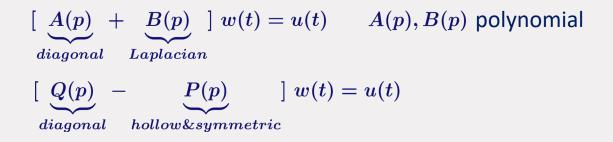
Mass-spring-damper system

- Masses M_j
- Springs K_{jk}
- Dampers D_{jk}
- Input u_j



$$\begin{bmatrix} M_{1} & & \\ & M_{2} & \\ & & M_{3} \end{bmatrix} \begin{bmatrix} \ddot{w}_{1} \\ \ddot{w}_{2} \\ \ddot{w}_{3} \end{bmatrix} + \begin{bmatrix} 0 & & \\ & D_{20} & \\ & & 0 \end{bmatrix} \begin{bmatrix} \dot{w}_{1} \\ \dot{w}_{2} \\ \dot{w}_{3} \end{bmatrix} + \begin{bmatrix} K_{10} & & \\ & 0 & \\ & & 0 \end{bmatrix} \begin{bmatrix} w_{1} \\ w_{2} \\ w_{3} \end{bmatrix} + \begin{bmatrix} K_{12} + K_{13} & -K_{12} & -K_{13} \\ -K_{12} & K_{12} & 0 \\ -K_{13} & 0 & K_{13} \end{bmatrix} \begin{bmatrix} w_{1} \\ w_{2} \\ w_{3} \end{bmatrix} + \begin{bmatrix} 0 & & \\ & -K_{12} & K_{12} & 0 \\ -K_{13} & 0 & K_{13} \end{bmatrix} \begin{bmatrix} w_{1} \\ w_{2} \\ w_{3} \end{bmatrix} = \begin{bmatrix} 0 \\ u_{2} \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} A(p) \\ diagonal \end{bmatrix} + \begin{bmatrix} B(p) \\ Laplacian \end{bmatrix} w(t) = u(t) \qquad A(p), B(p) \text{ polynomial } p = \frac{d}{dt}$$

Mass-spring-damper system



This fully fits in the earlier module representation:

$$w(t) = Gw(t) + \underbrace{Rr(t) + He(t)}_{Q^{-1}(p)u(t)}$$

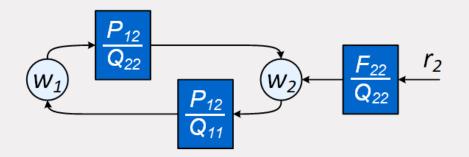
with the additional condition that:

 $G(p) = Q(p)^{-1}P(p)$ Q(p), P(p) polynomial P(p) symmetric, Q(p) diagonal



Module representation

Consequences for node interactions:



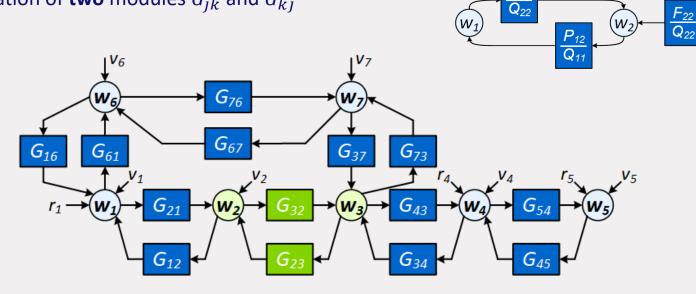
- Node interactions come in pairs of modules
- Where numerators are the same

Framework for network identification remains the same

Symmetry can simply be incorporated in identification

Local network identification

Identification of **one** physical interconnection Identification of **two** modules G_{jk} and G_{kj}



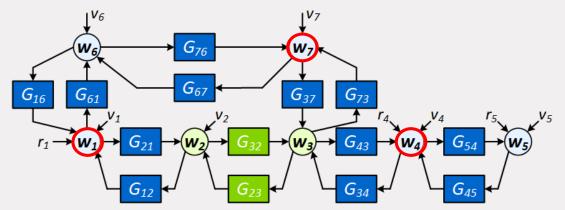
P₁₂



 r_2

Immersion conditions

For simultaneously identifying two modules in one interconnection:



The parallel path and loops-around-the-output condition, now simplifies to:

Measuring/exciting all neighbouring nodes of w_2 and w_3 leads to a solution



E.E.M. Kivits et al., CDC 2019.

Summary physical networks

• Physical networks fit within the module framework (special case)

- no restriction to second order equations

- Earlier identification framework can be utilized
- Local identification is well-addressed (and stays really local)
- Framework is fit for representing cyber-physical systems
 (combining physical bi-directional links, and cyber uni-directional links).





Extensions - Discussion

Extensions - Discussion

- Including sensor noise [1]
 - Errors-in-variabels problems can be more easily handled in a network setting
- Distributed estimation (MISO models) [2]
 - Communication constraints between different agents
 - Recursive (distributed) estimator converges to global optimizer (more slowly)
- Experiment design ^{[3],[4]}
 - design of least costly experiments

[1] Dankers et al., Automatica, 2015.

[2] Steentjes et al., IFAC-NECSYS, 2018.

[3] Gevers and Bazanella, CDC 2015.[4] Morelli, Bombois et al., ECC 2019;



Summary

• Dynamic network modeling:

intriguing research topic with many open questions

- The (centralized) LTI framework is only just the beginning
- Further move towards data-aspects related to distributed control
- and large-scale aspects
- and bring it to real-life applications

Acknowledgements



Lizan Kivits, Shengling Shi, Karthik Ramaswamy, Tom Steentjes, Mircea Lazar, Jobert Ludlage, Mannes Dreef, Tijs Donkers, Giulio Bottegal, Maarten Schoukens, Xiaodong Cheng

Co-authors, contributors and discussion partners:





Harm Weerts

Xavier Bombois Peter Heuberger Donatelllo Materassi Manfred Deistler Michel Gevers Jonas Linder Sean Warnick Alessandro Chiuso Hakan Hjalmarsson Miguel Galrinho Martin Enqvist





73

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, December 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, P.M.J. Van den Hof and A.G. Dankers (2018). Single module identifiability in linear dynamic networks. Proc. 57th IEEE CDC 2018.
- K.R. Ramaswamy, G. Bottegal and P.M.J. Van den Hof (2018). Local module identification in dynamic networks using regularized kernel-based methods. Proc. 57th CDC, 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. Provisionally accepted by Automatica. ArXiv 1901.00348.
- P.M.J. Van den Hof, K.R. Ramaswamy, A.G. Dankers and G. Bottegal. Local module identification in dynamic networks with correlated noise: the full input case. Proc. 2019 CDC.
- K.R. Ramaswamy and P.M.J. Van den Hof (2019). A local direct method for module identification in dynamic networks with correlated noise. Provisionally accepted by IEEE Trans. Automatic Control. ArXiv:1908.00976.
- X. Cheng, S. Shi and P.M.J. Van den Hof (2019). Allocation of excitation signals for generic identifiability of linear dynamic networks. Proc. 2019 CDC. ArXiv 1910.04525.
- K.R. Ramaswamy, P.M.J. Van den Hof and A.G. Dankers (2019). Generalized sensing and actuation schemes for local module identification in dynamic networks. Proc. 2019 CDC.
- E.M.M. Kivits and P.M.J. Van den Hof (2019). A dynamic network approach to identification of physical systems. Proc. 2019 CDC.
- S. Shi, G. Bottegal and P.M.J. Van den Hof (2019). Bayesian topology identification of linear dynamic networks. Proc. 2019 ECC, Napels, Italy, pp. 2814-2819.







The end