Data-driven modeling in linear dynamic networks

Paul M.J. Van den Hof

10th Harry Nicholson Distinguished Lecture in Control Engineering
The University of Sheffield, UK, 21 May 2019

www.sysdynet.eu
www.pvandenhof.nl
p.m.j.vandenhof@tue.nl
Introduction – dynamic networks

Decentralized process control

Smart power grid

Autonomous driving

Metabolic network

Brain network

Hydrocarbon reservoirs
Introduction

Overall trend:

• (Large-scale) interconnected systems
• With hybrid dynamics (continuous / switching)
• Distributed / multi-agent type monitoring, control and optimization problems
• Data is “everywhere”, big data era
• Modelling problems will need to consider this
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\textsuperscript{[1]}:

Identify a plant model $\hat{G}$ 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.

\textsuperscript{[1]}Ljung (1999), Söderström and Stoica (1989), Pintelon and Schoukens (2012)
Contents

• Introduction and motivation
• How to model a dynamic network?
• Single module identification – known topology
• Network identifiability
• Diffusively coupled physical networks
• Extensions - Discussion
Dynamic networks for data-driven modeling
Dynamic networks

State space representations
(Goncalves, Warnick, Sandberg, Yeung, Yuan, Scherpen,…)

Module representation
(VdH, Dankers, Materassi, Gevers, Bazanella,…)

Dynamic network setup

$r_i \quad$ external excitation
$v_i \quad$ process noise
$w_i \quad$ node signal
Dynamic network setup

$r_i$ external excitation
$v_i$ process noise
$w_i$ node signal
Dynamic network setup

$V_6$  $W_6$  $G^0_{61}$  $G^0_{26}$  $G^0_{27}$  $G^0_{37}$  $W_7$  $V_7$

$r_1$  $W_1$  $G^0_{21}$  $G^0_{32}$  $G^0_{43}$  $W_4$  $G^0_{54}$  $W_5$  $r_5$

$v_1$  $v_2$  $v_3$  $r_4$  $v_4$  $v_5$

$r_8$  $W_8$  $G^0_{18}$  $G^0_{23}$  $G^0_{34}$  $G^0_{45}$  $G^0_{84}$

$r_i$ external excitation
$v_i$ process noise
$w_i$ node signal
Dynamic network setup

$r_i$ external excitation
$v_i$ process noise
$w_i$ node signal
Dynamic network setup

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

$$w(t) = G^0(q)w(t) + R^0(q)r(t) + v(t)$$

Dynamic network setup

Many new identification 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
- Scalable algorithms
Contents

• Introduction and motivation
• How to model a dynamic network?
• **Single module identification – known topology**
• Network identifiability
• Diffusively coupled physical networks
• Extensions - Discussion
Single module identification - known topology
Single module identification

For a network with known topology:

- Identify $G_{21}^0$ on the basis of measured signals
- Which signals to measure? Preference for local measurements
Single module identification
Single module identification

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

4 input nodes to be measured: Can we do with less?

Network immersion \[1\]

- An immersed network is constructed by removing node signals, but leaving the remaining node signals invariant
- Modules and disturbance signals are adapted
- Abstraction through variable elimination (Kron reduction\[2\] in network theory).

Immersion
Immersion

When does immersion leave $G_{21}^0$ invariant?
Immersion

When does immersion leave $G_{21}^0$ invariant?

Parallel paths and loops around the output

There should be no parallel paths and loops around the output that run through removed nodes only.

Single module identification

parallel paths, and loops around the output
Single module identification

parallel paths, and loops around the output
Single module identification

parallel paths, and loops around the output
Single module identification

**parallel paths**, and **loops around the output**
Single module identification

Choose $w_6$ as an additional input (to be retained)
Single module identification

parallel paths, and loops around the output
Single module identification

parallel paths, and loops around the output
Single module identification

parallel paths, and loops around the output
Single module identification

Choose $\omega_3$ as an additional input, to be retained
Single module identification

Conclusion:

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

The immersion reasoning is *sufficient* but not *necessary* to arrive at a consistent estimate, see e.g. Linder and Enqvist [1], Bazanella et al. [2], Ramaswamy et al. [3]

Conclusion:

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

For a consistent and minimum variance estimate (direct method) there is one additional condition:

- absence of **confounding variables**, \(^1\)[2] i.e. correlated disturbances on inputs and outputs

---

\(^1\) J. Pearl, *Stat. Surveys, 3*, 96-146, 2009

\(^2\) A.G. Dankers et al., *Proc. IFAC World Congress*, 2017.
Confounding variables in the MISO case

- $w_7$ (not measured) now acts as a disturbance
Confounding variables in the MISO case

- $w_7$ (not measured) now acts as a disturbance
- Confounding variable if there is a path from $w_7$ to an input
- Can be solved by measuring $w_7$ and including it as input
Confounding variables in the MISO case

- $w_7$ (not measured) now acts as a disturbance
- Confounding variable if there is a path from $w_7$ to an input
- Can be solved by measuring $w_7$ and including it as input
- Or blocking the paths from $w_7$ to inputs/outputs by measured nodes, to be used as additional inputs.

Relation with d-separation in graphs (Materassi & Salapaka)

A.G. Dankers et al., IFAC World Congress, 2017.
Confounding variables in the MISO case

Can we always address confounding variables in this way?

No

If $v_2$ and $v_1$ are correlated then:
A MIMO approach with predicted outputs $w_2$ and $w_1$ can solve the problem

Summary single module identification

- Methods for **consistent** and **minimum variance** module estimation
- Graph tools for checking conditions
- Degrees of freedom in selection of measured signals – sensor selection
- A priori known modules can be accounted for
Contents

• Introduction and motivation
• How to model a dynamic network?
• Single module identification – known topology
• **Network identifiability**
• Diffusively coupled physical networks
• Extensions - Discussion
Network Identifiability
Network identifiability

Question: Can different dynamic networks be *distinguished* from each other from measured signals $w_i, r_i$?

Starting assumption: all signals $w_i, r_i$ that are present are measured.
Network identifiability

Network: \[ w = G^0 w + R^0 r + H^0 e \quad \text{cov}(e) = \Lambda^0, \quad \text{rank} \, p \]
\[ \text{dim}(r) = K \]

The network is defined by: \((G^0, R^0, H^0, \Lambda^0)\)
a network model is denoted by: \(M = (G, R, H, \Lambda)\)
and a network model set by:
\[ \mathcal{M} = \{M(\theta) = (G(\theta), R(\theta), H(\theta), \Lambda(\theta)), \theta \in \Theta\} \]

represents prior knowledge on the network models:
- topology
- disturbance correlation
- known modules
- the signals used for identification
Network identifiability

\[ w = (I - G^0)^{-1}[R^0 r + H^0 e] \]

Denote: \( w = T^0_{wr} r + \bar{v} \)

Objects that are uniquely identified from data \( r, w \): \( T^0_{wr}, \Phi^0_{\bar{v}} \)

**Definition**

A network model set \( \mathcal{M} \) is network identifiable from \( (w, r) \) at \( M_0 = M(\theta_0) \) if for all models \( M(\theta_1) \in \mathcal{M} \):

\[
\begin{align*}
T_{wr}(q, \theta_1) &= T_{wr}(q, \theta_0) \\
\Phi_{\bar{v}}(\omega, \theta_1) &= \Phi_{\bar{v}}(\omega, \theta_0)
\end{align*}
\]

\( \implies M(\theta_1) = M(\theta_0) \)
Network identifiability

Theorem – identifiability for general model sets

For each node signal $w_j$, let $\mathcal{P}_j$ be the set of in-neighbours of $w_j$ that map to $w_j$ through a parametrized module.

Then, under fairly general conditions,

$\mathcal{M}$ is network identifiable from $(w, r)$ at $M_0 = M(\theta_0)$ if and only if for all $j$:

- Each row of $[G(\theta) \quad H(\theta) \quad R(\theta)]$ has at most $K + p$ parametrized entries

- The transfer matrix from external inputs $(r, e)$ that are non-parametrized in $w_j$ to $\mathcal{P}_j$ has full row rank.
Example 5-node network

There are noise-free nodes, and $v_1$ and $v_2$ are expected to be correlated.

$\mathcal{M}$ with $H(\theta) = \begin{bmatrix} H_{11}(\theta) & H_{12}(\theta) & 0 \\ H_{21}(\theta) & H_{22}(\theta) & 0 \\ 0 & 0 & H_{33}(\theta) \end{bmatrix}$, $R = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 \end{bmatrix}$
Example 5-node network

There are noise-free nodes, and $v_1$ and $v_2$ are expected to be correlated

$$H = \begin{bmatrix}
  H_{11}(\theta) & H_{12}(\theta) & 0 \\
  H_{21}(\theta) & H_{22}(\theta) & 0 \\
  0 & 0 & H_{33}(\theta)
\end{bmatrix}, \quad R = \begin{bmatrix}
  0 & 0 & 0 \\
  0 & 0 & 0 \\
  0 & 1 & 0 \\
  0 & 0 & 1
\end{bmatrix}$$
Example 5-node network

There are noise-free nodes, and \( v_1 \) and \( v_2 \) are expected to be correlated

\[
\mathcal{M} \text{ with } H(\theta) = \begin{bmatrix}
H_{11}(\theta) & H_{12}(\theta) & 0 \\
H_{21}(\theta) & H_{22}(\theta) & 0 \\
0 & 0 & H_{33}(\theta)
\end{bmatrix}, \quad R = \begin{bmatrix}
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 1
\end{bmatrix}
\]
Example 5-node network

There are noise-free nodes, and $v_1$ and $v_2$ are expected to be correlated

\[ M \text{ with } H(\theta) = \begin{bmatrix} H_{11}(\theta) & H_{12}(\theta) & 0 \\ H_{21}(\theta) & H_{22}(\theta) & 0 \\ 0 & 0 & H_{33}(\theta) \end{bmatrix} , \quad R = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \]
Example 5-node network

There are noise-free nodes, and $v_1$ and $v_2$ are expected to be correlated

$\mathcal{M}$ with $H(\theta) = \begin{bmatrix} H_{11}(\theta) & H_{12}(\theta) & 0 \\ H_{21}(\theta) & H_{22}(\theta) & 0 \\ 0 & 0 & H_{33}(\theta) \end{bmatrix}$, $R = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$
Example 5-node network

If we restrict the structure of $G(\theta)$:

$$G(\theta) = 
\begin{bmatrix}
0 & G_{12}(\theta) & 0 & 0 & G_{15}(\theta) \\
G_{21}(\theta) & 0 & G_{23}(\theta) & 0 & 0 \\
0 & 0 & 0 & G_{34}(\theta) & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & G_{53}(\theta) & 0 & 0
\end{bmatrix}$$

$$[H \ R] = 
\begin{bmatrix}
H_{11}(\theta) & H_{12}(\theta) & 0 & 0 & 0 \\
H_{21}(\theta) & H_{22}(\theta) & 0 & 0 & 0 \\
0 & 0 & H_{3}(\theta) & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}$$

First condition:
Number of parametrized entries in each row < $K+p = 5$
Example 5-node network

\[ G(\theta) = \begin{bmatrix} 0 & G_{12}(\theta) & 0 & 0 & G_{15}(\theta) \\ G_{21}(\theta) & 0 & G_{23}(\theta) & 0 & 0 \\ 0 & 0 & 0 & G_{34}(\theta) & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & G_{53}(\theta) & 0 & 0 \end{bmatrix} \]

\[ [H \ R] = \begin{bmatrix} H_{11}(\theta) & H_{12}(\theta) & 0 & 0 & 0 \\ H_{21}(\theta) & H_{22}(\theta) & 0 & 0 & 0 \\ 0 & 0 & H_{3}(\theta) & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \]

Rank condition:
Row 1: Full row rank of transfer:
\[ \begin{bmatrix} v_3 \\ r_4 \\ r_5 \end{bmatrix} \rightarrow \begin{bmatrix} w_2 \\ w_5 \end{bmatrix} \]
Example 5-node network

Verifying the rank condition for $w_1$:

$j = 1$ : Evaluate the rank of the transfer matrix

\[
\begin{bmatrix}
  v_3 \\ r_4 \\ r_5
\end{bmatrix}
\rightarrow
\begin{bmatrix}
  w_2 \\ w_5
\end{bmatrix}
\]
Example 5-node network

Verifying the rank condition for $w_1$:

\[ j = 1 : \text{Evaluate the rank of the transfer matrix} \begin{bmatrix} v_3 \\ r_4 \\ r_5 \end{bmatrix} \rightarrow \begin{bmatrix} w_2 \\ w_5 \end{bmatrix} \]
**Example 5-node network**

Verifying the rank condition for $w_1$:

For the generic case, the rank can be calculated by a graph-based condition\([1],[2],[3]\):

**Generic rank = number of vertex-disjoint paths**

\[ \begin{bmatrix} v_3 \\ r_4 \\ r_5 \end{bmatrix} \rightarrow \begin{bmatrix} w_2 \\ w_5 \end{bmatrix} \]

2 vertex-disjoint paths $\rightarrow$ full row rank 2

---

[2] Hendrickx, Gevers & Bazanella, CDC 2017  
[3] Weerts et al., CDC 2018
Graph-based synthesis solution for full network

Decompose network in **disjoint pseudo-trees**:

- Connected directed graphs, where nodes have maximum indegree 1
- Edges are disjoint and all out-neighbours of a node are in the same pseudo-tree

- Any network can be decomposed into a set of disjoint pseudo-trees
Graph-based synthesis solution for full network

Result\cite{1}

A network is generically identifiable if

- It can be decomposed in $K$ disjoint pseudo-trees, and
- There are $K$ independent external signals entering at a root of each pseudo-tree

Two typical (disjunct) pseudo-trees:

Tree with root in green

Cycle with outgoing trees; Any node in cycle is root

\[1\] X. Cheng, S. Shi and PVdH, CDC 2019, submitted.
Where to allocate external excitations for network identifiability?

Two disjunct pseudo-trees
Where to allocate external excitations for network identifiability?

Two independent excitations guarantee network identifiability

Algorithm available for merging pseudo-trees.
Summary identifiability

Identifiability of network model sets is determined by

- Presence and location of external signals, and
- Correlation of disturbances
- Prior knowledge on modules

So far:

- All node signals assumed to be measured
- Fully applicable to the situation $p < L$ (i.e. reduced-rank noise)
- Identifiability of the full network model – conditions per row/output node
- Extensions towards identifiability of a single module [1],[2]

[2] Weerts et al., CDC 2018
Contents

• Introduction and motivation
• How to model a dynamic network?
• Single module identification – known topology
• Network identifiability
• **Diffusively coupled physical networks**
• Extensions - Discussion
Diffusively coupled 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:

\[
I = \frac{1}{R} (V_1 - V_2)
\]

\[
F = K (x_1 - x_2)
\]

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 $u_j$

\[
\begin{bmatrix}
M_1 & M_2 \\
M_2 & M_3
\end{bmatrix}
\begin{bmatrix}
\dot{\ddot{w}}_1 \\
\dot{\ddot{w}}_2 \\
\dot{\ddot{w}}_3
\end{bmatrix}
+ \begin{bmatrix} 0 \\
D_{20}
\end{bmatrix}
\begin{bmatrix}
\dddot{w}_1 \\
\dddot{w}_2 \\
\dddot{w}_3
\end{bmatrix}
+ \begin{bmatrix} K_{10} \\
0
\end{bmatrix}
\begin{bmatrix}
w_1 \\
w_2 \\
w_3
\end{bmatrix}
+ \begin{bmatrix} D_{13} & 0 & -D_{13} \\
0 & D_{23} & -D_{23} \\
-D_{13} & -D_{23} & D_{13} + D_{23}
\end{bmatrix}
\begin{bmatrix}
\dot{\ddot{w}}_1 \\
\dot{\ddot{w}}_2 \\
\dot{\ddot{w}}_3
\end{bmatrix}
+ \begin{bmatrix} K_{12} + K_{13} & -K_{12} & -K_{13} \\
-K_{12} & K_{12} & 0 \\
-K_{13} & 0 & K_{13}
\end{bmatrix}
\begin{bmatrix}
w_1 \\
w_2 \\
w_3
\end{bmatrix}
= \begin{bmatrix} 0 \\
u_2
\end{bmatrix}
\]

\[
\begin{bmatrix}
A(p) & B(p)
\end{bmatrix}
\begin{bmatrix}
d \\
dt
\end{bmatrix}
= u(t) \quad A(p), B(p) \text{ polynomial}\]
Mass-spring-damper system

\[
\begin{bmatrix}
    A(p) & B(p) \\
    \text{diagonal} & \text{Laplacian}
\end{bmatrix} w(t) = u(t) \quad A(p), B(p) \text{ polynomial}
\]

\[
\begin{bmatrix}
    Q(p) & -P(p) \\
    \text{diagonal} & \text{hollow}
\end{bmatrix} w(t) = u(t)
\]

- \( Q_{11} = M_1 p^2 + D_{13} p + (K_{10} + K_{12} + K_{13}) \)
- \( Q_{22} = M_2 p^2 + (D_{20} + D_{23}) p + K_{12} \)
- \( Q_{33} = M_3 p^2 + (D_{13} + D_{23}) p + K_{13} \)

\( Q_{jj} : \) elements related to node \( w_j \):

\( P_{ji} = P_{ij} : \) elements related to interconnection

\[
P = \begin{bmatrix}
    0 & K_{12} & D_{13} p + K_{13} \\
    K_{12} & 0 & D_{23} p \\
    D_{13} p + K_{13} & D_{23} p & 0
\end{bmatrix}
\]
Module representation

\[
\begin{bmatrix}
\underline{Q(p)} & -\underline{P(p)}
\end{bmatrix}
\begin{bmatrix}
w(t)
\end{bmatrix}
= Fr(t) + C(p)e(t)
\]

with the additional condition that:

\[
w(t) = Q^{-1}Pw(t) + Q^{-1}Fr(t) + Q^{-1}C(p)e(t)
\]

This fully fits in the earlier module representation:

\[
w(t) = Gw(t) + Rr(t) + He(t)
\]

with the additional condition that:

\[
G(p) = Q(p)^{-1}P(p) \quad Q(p), P(p) \text{ polynomial}
\]

\[
P(p) \text{ symmetric, } Q(p) \text{ diagonal}
\]
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 $G_{kj}$

$$G_{jk} = Q_{jj}^{-1}P_{jk} \quad \text{and} \quad G_{kj} = Q_{kk}^{-1}P_{kj} \quad \text{with} \quad P_{jk} = P_{kj}$$
Immersion conditions

For simultaneously identifying two modules in one interconnection:

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

All neighbouring nodes of $w_2$ and $w_3$ need to be retained/measured.

E.E.M. Kivits et al., CDC 2019 submitted.
Summary diffusively coupled physical networks

- Physical networks fit within the module framework (special case)
  - no restriction to second order equations
- Identification algorithms and identifiability analysis can be utilized
- Local identification is well-addressed (and stays really local)
- Framework is fit for representing cyber-physical systems
Extensions - Discussion
Extensions - Discussion

- **Identification algorithms to deal with reduced rank noise** [1]
  - number of disturbance terms is larger than number of white sources
  - Optimal identification criterion becomes a constrained quadratic problem with ML properties for Gaussian noise
  - Reworked Cramer Rao lower bound
  - Some parameters can be estimated variance free

- **Including sensor noise** [2]
  - Errors-in-variables problems can be more easily handled in a network setting

---

Extensions - Discussion

- **Machine learning tools for estimating large scale models** [1,2]
  - Choosing correctly parametrized model sets for all modules is impractical
  - Use of Gaussian process priors for kernel-based estimation of models

- **From centralized to distributed estimation (MISO models)** [3]
  - Communication constraints between different agents
  - Recursive (distributed) estimator converges to global optimizer (more slowly)

Discussion

• **Dynamic network identification:** 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, Giulio Bottegal, Maarten Schoukens, Xiaodong Cheng

Co-authors, contributors and discussion partners:

Arne Dankers
Harm Weerts

Xavier Bombois
Peter Heuberger
Donatello Materassi
Manfred Deistler
Michel Gevers
Jonas Linder
Sean Warnick
Alessandro Chiuso
Hakan Hjalmarsson
Miguel Galrinho
Further reading


Papers available at www.pvandenhof.nl
The end