# Identifiability of dynamic networks with noisy and noise-free nodes

Paul Van den Hof

Coworkers: Harm Weerts, Arne Dankers

21 November 2016, TU Vienna



European Research Council

Particular Structure Contracting Technische Universiteit Eindhoven University of Technology

Where innovation starts

TU/

### **Dynamic network**





## Introduction – relevant identification questions



Question: Can the dynamics/topology of a network be uniquely determined from measured signals  $w_i$ ,  $r_i$ ? Question: Can different dynamic networks be *distinguished* from each other from measured signals  $w_i$ ,  $r_i$ ?



## Introduction

When are models essentially different (in view of identification)? |e

In classical PE identification: Models are indistinguishable (from data) if their predictor filters are the same:

$$\hat{y}(t|t-1) = \underbrace{H(q)^{-1}G(q)}_{W_u(q)} u(t) + \underbrace{[1-H(q)^{-1}]}_{W_y(q)} y(t)$$

 $(G_1, H_1)$  and  $(G_2, H_2)$  are indistinguishable iff

$$\left\{ egin{array}{c} H_1^{-1}G_1 = H_2^{-1}G_2 \ 1 - H_1^{-1} = 1 - H_2^{-1} \end{array} \Leftrightarrow \left\{ egin{array}{c} G_1 = G_2 \ H_1 = H_2 \end{array} 
ight.$$

TU/e Technische Universiteit Eindhoven University of Technology

H

G

## Introduction

For a parametrized model set (model structure):

$$\hat{y}(t|t-1; heta) = \underbrace{H(q, heta)^{-1}G(q, heta)}_{W_u(q, heta)} u(t) + \underbrace{[1-H(q, heta)^{-1}]}_{W_y(q, heta)} y(t) \qquad heta \in \Theta$$

parameter values can be distinguished if

$$\begin{cases} G(\theta_1) = G(\theta_2) \\ H(\theta_1) = H(\theta_2) \end{cases} \end{cases} \Longrightarrow \theta_1 = \theta_2 \quad \text{for all } \theta \in \Theta$$

This property is generally known as the property of identifiability of the model structure



## Introduction

So there are two different bijective mappings involved:



#### Reason:

- Freedom in network structure
- Freedom in presence of excitation and disturbances



## **Network Setup**



#### **Assumptions:**

- Total of *L* nodes
- Network is well-posed and stable
- All  $w_m, m = 1, \cdots L,$ and present  $r_m$  are measured

Technische Universiteit **Eindhoven** University of Technology

- Modules may be unstable
- Modules are strictly proper (can be generalized)

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} = \begin{bmatrix} 0 & G_{12}^0 & \cdots & G_{1L}^0 \\ G_{21}^0 & 0 & \cdots & G_{2L}^0 \\ \vdots & \ddots & \ddots & \vdots \\ G_{L1}^0 & G_{L2}^0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} + R^0 \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_K \end{bmatrix} + H^0 \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{bmatrix}$$

$$e \text{ white noise}$$

## **Network Setup**

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} = \begin{bmatrix} 0 & G_{12}^0 & \cdots & G_{1L}^0 \\ G_{21}^0 & 0 & \cdots & G_{2L}^0 \\ \vdots & \cdots & \cdots & \vdots \\ G_{L1}^0 & G_{L2}^0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} + R^0 \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_K \end{bmatrix} + H^0 \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{bmatrix}$$

#### **Different situations:**

- *p*=*L*: Full rank noise process that disturbs every measured node
- *p*<*L*: Singular noise process, with the distinct options:
  - a) All nodes are noise disturbed
  - b) Some nodes noise-free; other nodes have full rank noise



## **Network Setup**

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} = \begin{bmatrix} 0 & G_{12}^0 & \cdots & G_{1L}^0 \\ G_{21}^0 & 0 & \cdots & G_{2L}^0 \\ \vdots & \cdots & \cdots & \vdots \\ G_{L1}^0 & G_{L2}^0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} + R^0 \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_K \end{bmatrix} + H^0 \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{bmatrix}$$

#### **Different situations:**

- *p*=*L*: Full rank noise process that disturbs every measured node
- *p*<*L*: Singular noise process, with the distinct options:

a) All nodes are noise disturbed

b) Some nodes noise-free; other nodes have full rank noise

Common situation in PE identification:  $H^0$  square and monic

Non-common situation:  $H^0$  non-square



### **Network identification setup**

#### **Network predictor:**

$$\hat{w}(t|t-1) = \mathbb{E}\{w(t) \mid w^{t-1}, r^t\}$$

with  $w^{t-1} = \{w(0), w(1), \cdots w(t-1)\}$ 

The network is defined by:  $(G^0, R^0, H^0)$ a network model is denoted by: M = (G, R, H)and a network model set by:

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

Models manifest themselves in identification through their predictor



Decompose the node signals

$$w(t) = egin{bmatrix} w_a(t) \ w_b(t) \end{bmatrix}$$

with  $w_a(t)$  noisy and  $w_b(t)$  noise-free (a priori known).

such that 
$$v(t) = \begin{bmatrix} H_a(q) \\ 0 \end{bmatrix} e(t)$$



## **Network identification setup**

Predictor can be written as:

$$\hat{w}(t|t-1) = W(q) egin{bmatrix} w(t) \ r(t) \end{bmatrix}$$

Problem with noise free-nodes:

Filter W(q) is non-unique, due to the noise-free nodes in w(t)

The predictor filter can be made unique when removing the noise-free signals as inputs

$$\begin{cases} \hat{w}(t|t-1) = P(q) \begin{bmatrix} w_a(t) \\ r(t) \end{bmatrix} \\ \hat{w}_b(t|t-1) = w_b(t) \end{cases}$$



#### When can network models be distinguished through identification?

Two optional directions to continue:

1. The philosophical path (Plato)

2. The pragmatic path (Aristoteles)





# Network identifiability (Philosphical path)

#### **Generalized notion:**

Consider an identification criterion J determining:

 $J(z,\mathcal{M})$ 

with z measured data,  $\mathcal{M}$  a model set, and  $J(z, \mathcal{M}) \subset \mathcal{M}$ , the set of identified models based on data z

#### **Definition:**

Model set  $\mathcal{M}$  is network identifiable (w.r.t. J) at  $M_0 \in \mathcal{M}$  if in  $\mathcal{M}$ there does not exist a model  $M_1 \neq M_0$  that always appears together with  $M_0$  in  $J(z, \mathcal{M})$ 

 $\mathcal{M}$  is network identifiable (w.r.t. J) if it is network identifiable at all  $M_0 \in \mathcal{M}$ 

Van den Hof, 1989, 1994 Weerts et al. (2016)



# Network identifiability (Philosphical path)

#### Identification criterion for the situation of noise-free nodes:

The identification criterion:

$$J\left(\mathrm{z},\mathcal{M}
ight) = \left\{ egin{arg min $\overline{\mathbb{E}}$ $arepsilon_a^T(t, heta)\Lambda^{-1}arepsilon_a(t, heta)$}\ {}_{M( heta)}^{H( heta)} \stackrel{ heta\in\Theta}{\in} \ {}^{\mathrm{subject to: }} arepsilon_b(t, heta) = 0 \ \ {
m for all } t. \end{array} 
ight\}$$

Noise-free nodes can be predicted exactly



## Network identifiability (merging of the paths)

Theorem 1 (or Definition 1)

Denote 
$$T(q)$$
 as the transfer function  $\begin{pmatrix} e \\ r \end{pmatrix} \to w$   
 $T(q) = (I - G(q))^{-1}U(q)$  with  $U(q) := \begin{bmatrix} H_a(q) & R_a(q) \\ 0 & R_b(q) \end{bmatrix}$ 

and let  $T(q, \theta)$  be its parametrized version

Then the network model set  $\mathcal{M}$  is network identifiable at  $M_0 = M(\theta_0)$ (w.r.t. J) if for all models  $M(\theta_1) \in \mathcal{M}$ :

$$T(q, \theta_1) = T(q, \theta_0) \Longrightarrow M(\theta_1) = M(\theta_0)$$

Goncalves and Warnick, 2008; Weerts et al, SYSID2015; Weerts et al. ALCOSP 2016; Gevers and Bazanella, 2016.



 ${\mathcal M}$  is network identifiable (w.r.t. J) if it is network identifiable at all models  $M \in {\mathcal M}$  .

#### **Question of identifiability means:**

Can the models in a network model set be distinguished from each other through identification

- a) With respect to a single model  $M_0 = M(\theta_0)$
- b) With respect to all models in the set



#### Theorem 2

 $\mathcal{M}$  is network identifiable at  $M_0 = M(\theta_0)$  (w.r.t. J) if there exists a square, nonsingular and fixed Q(q) such that

 $U(q, heta) Q(q) = egin{bmatrix} D(q, heta) & F(q, heta) \end{bmatrix} \qquad egin{array}{cc} U(q) \coloneqq egin{bmatrix} H_a(q) & R_a(q) \ 0 & R_b(q) \end{bmatrix}$ 

With  $D(q, \theta)$  square, and such that  $D(q, \theta)$  is diagonal and full rank for all

$$heta\in\Theta_0:=\{ heta\in\Theta\mid T(q, heta)=T(q, heta_0)\}$$

 $\mathcal{M}$  is network identifiable (w.r.t. J) if the above holds for  $\Theta_0 = \Theta$ 

Goncalves and Warnick, 2008; Weerts et al, SYSID2015; Gevers and Bazanella, 2016; Weerts et al., ArXiv 2016;



## **Closed-loop example**

This classical closed-loop system has a noise-free node



$$\mathcal{M} ext{ with } G( heta) = egin{bmatrix} 0 & G_0( heta) \ -C( heta) & 0 \end{bmatrix}, \ H( heta) = egin{bmatrix} H_a( heta) \ 0 \end{bmatrix}, \ R = egin{bmatrix} 0 \ 1 \end{bmatrix}$$

 $\begin{bmatrix} H(\theta) & R(\theta) \end{bmatrix}$  is diagonal and full rank  $\rightarrow$  identifiability



## **Example correlated noises**



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



#### **Example correlated noises**



We can not arrive at a diagonal structure in  $\begin{bmatrix} H( heta) & R( heta) \end{bmatrix}$ 



#### **Observations:**

- a) A simple test can be performed to check the condition
- b) The condition is typically fulfilled if each node  $w_j$  is excited by either an external excitation  $r_j$  or a noise  $v_j$  that are uncorrelated with the external signals on other nodes.
- c) The result is rather **conservative**:
  - 1. Restricted to situation where  $U(q, \theta)$  is full row rank
  - 2. Does not take account of structural properties of  $G(q, \theta)$  e.g. modules/controllers that are known a priori



#### **Theorem 3 – identifiability in case of structure restrictions**

Assumptions:

- a) Each parametrized entry in  $M(q, \theta)$  covers the set of all proper rational transfer functions
- b) All parametrized elements in  $M(q, \theta)$  are parametrized independently

Then the network model set  $\mathcal{M}$  is network identifiable at  $M_0 = M(\theta_0)$  (w.r.t. J) if and only if:

- Each row *i* of  $\begin{bmatrix} G(\theta) & U(\theta) \end{bmatrix}$  has at most K+p parametrized entries
- For each row i,  $\check{T}_i(q, \theta_0)$  has full row rank

where:  $\check{T}_i(q, \theta_0)$  is the submatrix of  $T(q, \theta_0)$ , composed of those rows *j* that correspond to elements  $G_{ij}(q, \theta)$  that are parametrized

Goncalves and Warnick, 2008; Weerts et al., ArXiv 2016;



#### **Theorem 3 – identifiability in case of structure restrictions**

Assumptions:

- a) Each parametrized entry in  $M(q, \theta)$  covers the set of all proper rational transfer functions
- b) All parametrized elements in  $M(q, \theta)$  are parametrized independently

Then the network model set  $\mathcal{M}$  is network identifiable (w.r.t. J) if and only if:

- Each row *i* of  $\begin{bmatrix} G(\theta) & U(\theta) \end{bmatrix}$  has at most K+p parametrized entries
- For each row  $i, \ \check{T}_i(q, \theta)$  has full row rank for all  $\theta \in \Theta$

where:  $\check{T}_i(q, \theta_0)$  is the submatrix of  $T(q, \theta_0)$ , composed of those rows *j* that correspond to elements  $G_{ij}(q, \theta)$  that are parametrized



#### Corollary – situation of $U(\theta)$ full row rank

If  $U(\theta)$  is full row rank for  $(\theta = \theta_0 / \forall \theta \in \Theta)$ 

Then  $\mathcal{M}$  is network identifiable (at  $M(\theta_0)$ ) if and only if:

• Each row *i* of  $\begin{bmatrix} G(\theta) & U(\theta) \end{bmatrix}$  has at most *K*+*p* parametrized entries

The number of parametrized transfer functions that map into a node  $w_i$  should not exceed the total number of excitation+noise signals in the network.



## **Example correlated noises (continued)**



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

$$G( heta) = egin{bmatrix} 0 & G_{12}( heta) & 0 & 0 & G_{15}( heta) \ G_{21}( heta) & 0 & 0 & 0 & 0 \ 0 & 0 & 0 & G_{34}( heta) & 0 & 0 \ 0 & 0 & 0 & G_{34}( heta) & 0 & 0 \ 0 & 0 & 0 & H_3( heta) & 0 & 0 \ 0 & 0 & 0 & H_3( heta) & 0 & 0 \ 0 & 0 & 0 & 1 & 0 \ 0 & 0 & 0 & 0 & 1 \ 0 & 0 & 0 & 0 & 1 \ \end{bmatrix} U( heta) = egin{bmatrix} H_{11}( heta) & H_{12}( heta) & 0 & 0 & 0 \ H_{21}( heta) & H_{22}( heta) & 0 & 0 & 0 \ 0 & 0 & 0 & H_3( heta) & 0 & 0 \ 0 & 0 & 0 & 1 & 0 \ 0 & 0 & 0 & 1 & 0 \ 0 & 0 & 0 & 0 & 1 \ \end{bmatrix}$$

Node/row 1 has 4 unknowns < K+p = 5 Node/row 2 has 3 unknowns (2 from noise model) < K+p

#### → identifiable!

## **Example: identifiability at a particular model**



T:r
ightarrow w

	1	0		1	0
$T_1 =$	$oldsymbol{A}$	1	$T_2 =$	(A+1)B	1
	AB+1	B		A+1	0



## Example: identifiability at a particular model

$$\mathcal{M} ext{ with } G( heta) = egin{bmatrix} 0 & G_{12}( heta) & G_{13}( heta) \ G_{21}( heta) & 0 & G_{23}( heta) \ G_{31}( heta) & G_{32}( heta) & 0 \end{bmatrix} ext{ and } R = egin{bmatrix} 1 & 0 \ 0 & 1 \ 1 & 0 \end{bmatrix}$$

Can we find a unique model that satisfies  $T = (I - G(\theta))^{-1}R$ 

Uniqueness of the solution depends on the system

The model set is network identifiable in system 1 but not in system 2



## Result

When is the model identifiable? Evaluate:  $(I - G(\theta))T = R$ 

#### **Condition 1**



## **Example 1 continued**

#### The reason there is no identifiability







- Concept of network identifiability has been introduced and extended beyond the classical PE assumptions (all measurements noisy)
   "can models be distinguished in identification?"
- The network transfer functions T remain the objects that can be uniquely identified from data
- Results lead to verifiable conditions on the network structure / parametrization / presence of external signals
- The framework is fit for extending it to the general situation of singular / reduced-rank noise





#### The general reduced-rank case

Decompose the node signals

$$w(t) = egin{bmatrix} w_a(t) \ w_b(t) \end{bmatrix}$$

with  $v_a(t)$  full-rank noise and  $v_b(t)$  driven by the same noise as  $v_a(t)$  (a priori known).

$$\begin{bmatrix} v_a(t) \\ v_b(t) \end{bmatrix} = \begin{bmatrix} H_a(q) \\ H_b(q) \end{bmatrix} e(t)$$

With  $H_a(q)$  monic and  $\lim_{z
ightarrow\infty}H_b(z)=\Gamma$ 



#### **Example rank-reduced noise**



This system has a rank-reduced noise

$$\mathcal{M} ext{ with } G( heta) = egin{bmatrix} 0 & G_{12}( heta) \ G_{21}( heta) & 0 \end{bmatrix}, \ H( heta) = egin{bmatrix} H_1( heta) \ H_2( heta) \end{bmatrix}, \ R = I$$



## Suitable identification criterion

Identification criterion for the situation of rank-reduced noise:

The identification criterion:

$$J\left(\mathrm{z},\mathcal{M}
ight) = egin{cases} rgmin_{M( heta)} & ar{\mathbb{E}} \ arepsilon_a^T(t, heta) \Lambda_a^{-1} arepsilon_a(t, heta) \ & \ M( heta) \ heta \in \Theta \ & \ ext{subject to: } arepsilon_b(t, heta) = \Gamma( heta) arepsilon_a(t, heta) \ & \ ext{for all } t. \ \end{pmatrix}$$

The constraint accounts for the fact that  $w_a(t)$  and  $w_b(t)$  are driven by the same noise

#### Identifiability condition remains unchanged:

$$T(q, heta_1)=T(q, heta_0)\Longrightarrow M( heta_1)=M( heta_0)$$
ere  $T(q)$  is the transfer function  $igg(egin{array}{c} e \ r \ \end{pmatrix}
ightarrow w$ 

Previous results remain valid

Van den Hof et al., ACC 2017, submitted; Weerts et al, IFAC 2017, submitted

/ Electrical Engineering - Control Systems

wh



## Suitable identification criterion

If the identifiability condition is satisfied then **consistency** can be shown provided that the external excitation signals are persistently exciting of a sufficient high order.

Weerts et al, IFAC 2017, submitted

## **Relaxed identification criterion**

#### **Relaxation of the constraint identification criterion**

$$egin{aligned} J\left(\mathrm{z},\mathcal{M}
ight) &= rgmin_{M( heta)} \min_{ heta \in \Theta} ar{\mathbb{E}} \, arepsilon_a^T(t, heta) \Lambda_a^{-1} arepsilon_a(t, heta) + \ &+ \lambda [arepsilon_b(t, heta) - \Gamma( heta) arepsilon_a(t, heta)]^T [arepsilon_b(t, heta) - \Gamma( heta) arepsilon_a(t, heta)]^T [arepsilon_b(t, heta) - \Gamma( heta) arepsilon_a(t, heta)]^T arepsilon_b(t, heta) - \Gamma( heta) arepsilon_a(t, heta)]^T arepsilon_a(t, heta) - \Gamma( heta) arepsilon_a(t, heta)]^T arepsilon_b(t, heta) - \Gamma( heta) arepsilon_a(t, heta)]^T arepsilon_a(t, heta) - \Gamma( heta) arepsilon_a(t, heta)]^T arepsilon_a(t, heta) - \Gamma( heta) arepsilon_a(t, heta) - \Gamma( heta) arepsilon_a(t, heta)]^T arepsi$$

$$egin{aligned} &J\left(\mathrm{z},\mathcal{M}
ight) = rgmin_{M( heta)} egin{aligned} &ar{\mathbb{E}} \ arepsilon^T(t, heta) \ W( heta) \ arepsilon(t, heta) \end{aligned} \ & egin{aligned} &W( heta) = egin{bmatrix} &\Lambda_a^{-1} & 0 \ 0 & 0 \end{bmatrix} + \lambda egin{bmatrix} &\Gamma^T( heta) \ &-I \end{bmatrix} egin{bmatrix} &\Gamma( heta) \ &-I \end{bmatrix}, &\lambda > 0 \end{aligned}$$

Equivalent to the constrained criterion for  $\lambda \to \infty$ 



#### Simulation example



Identification for various weights

 $Q_I: W = I$  $Q_{10}: W( heta)$  with  $\lambda = 10$  $Q_{10^3}: W( heta)$  with  $\lambda = 10^3$  $Q_\infty: W( heta)$  with  $\lambda o \infty$ 

In some situations variance-free estimates occur





4

Ŧ



• The reduced-rank case can be treated equally well !



# Bibliography

- 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.
- B. Günes, A. Dankers and P.M.J. Van den Hof (2014). Variance reduction for identification in dynamic networks. Proc. 19th IFAC World Congress, 24-29 August 2014, Cape Town, South Africa, pp. 2842-2847.
- A.G. Dankers, P.M.J. Van den Hof, P.S.C. Heuberger and X. Bombois (2012). Dynamic network structure identification with prediction error methods - basic examples. Proc. 16th IFAC Symposium on System Identification (SYSID 2012), 11-13 July 2012, Brussels, Belgium, pp. 876-881.
- A.G. Dankers, P.M.J. Van den Hof and X. Bombois (2014). An instrumental variable method for continuous-time identification in dynamic networks. Proc. 53rd IEEE Conf. Decision and Control, Los Angeles, CA, 15-17 December 2014, pp. 3334-3339.
- H.H.M. Weerts, A.G. Dankers and P.M.J. Van den Hof (2015). Identifiability in dynamic network identification. Proc.17th IFAC Symp. System Identification, 19-21 October 2015, Beijing, P.R. China.
- 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. Automatic Control, 61 (4)*, pp. 937-952, April 2016.
- H.H.M. Weerts, P.M.J. Van den Hof and A.G. Dankers (2016). Identifiability of dynamic networks with part of the nodes noise-free. Proc. 12th IFAC Intern. Workshop ALCOSP 2016, June 29 July 1, 2016, Eindhoven, The Netherlands.
- H.H.M. Weerts, P.M.J. Van den Hof and A.G. Dankers (2016). Identifiability of dynamic networks with noisy and noise-free nodes. <u>ArXiv:1609.00864</u> [CS.sy]
- P.M.J. Van den Hof, H.H.M. Weerts and A.G. Dankers (2016). Prediction error identification with rank-reduced output noise. Submitted to 2017 American Control Conference, 24-26 May 2017, Seattle, WA, USA.
- H.H.M. Weerts, P.M.J. Van den Hof and A.G. Dankers (2016). Identification of dynamic networks with rank-reduced process noise. Submitted to 2017 IFAC World Congress, 9-14 July 2017, Toulouse, France.

Papers available at www.pvandenhof.nl/publications.htm

