# Errors-in-Variables Identification in Dynamic Networks - Consistency Results for an Instrumental Variable Approach $^\star$

Arne Dankers [a] Paul M.J. Van den Hof [b] Xavier Bombois [c] Peter S.C. Heuberger [d]

[a] *Delft Center for Systems and Control, Delft University of Technology, The Netherlands (email: adankers@hifieng.com).*

[b] *Dept. of Electrical Engineering, Eindhoven University of Technology, The Netherlands (email: p.m.j.vandenhof@tue.nl).*

[c] *Laboratoire Ampère, École Centrale de Lyon, France (email: xavier.bombois@ec-lyon.fr).*

[d] *Dept. of Mechanical Engineering, Eindhoven University of Technology, The Netherlands (email: p.s.c.heuberger@tue.nl).*

## Abstract

In this paper we consider the identification of a linear module that is embedded in a dynamic network using noisy measurements of the internal variables of the network. This is an extension of the errors-in-variables (EIV) identification framework to the case of dynamic networks. The consequence of measuring the variables with sensor noise is that some prediction error identification methods no longer result in consistent estimates. The method developed in this paper is based on a combination of the instrumental variable philosophy and closed-loop prediction error identification methods, and leads to consistent estimates of modules in a dynamic network. We consider a flexible choice of which internal variables need to be measured in order to identify the module of interest. This allows for a flexible sensor placement scheme. We also present a method that can be used to validate the identified model.

## 1 Introduction

Many systems in engineering can be modelled as dynamic networks, as e.g. power systems, telecommunication systems, and distributed control systems. Models of these networks are important either for prediction, simulation, controller design or fault detection. Since sensors are becoming more ubiquitous and cheaper the result is that data can be collected from many variables in a dynamic network, and a system identification approach for modelling particular modules in the dynamic network, becomes attractive. Using this approach it is important to be mindful of the fact that every measurement is contaminated with *sensor noise.*

The literature on dynamic network identification can be split into two categories based on whether the interconnection structure of the network is known or not. In the case that the interconnection structure is not known, the network structure together with the dynamics typically needs to be estimated. The majority of the papers in this category are based on the concept of *Granger Causality* (Granger, 1980). In Caines & Chan (1975); Gevers &

Anderson (1981) it is shown that it is possible to distinguish between open and closed-loop data generating systems. The reasoning is extended to more complex interconnection structures using a non-parametric approach (Materassi & Innocenti, 2010; Materassi & Salapaka, 2012); using a Bayesian approach (Chuiso & Pillonetto, 2012); and using a parametric approach supplemented by $\ell_0$ regularization (Seneviratne & Solo, 2012; Yuan et al., 2011), $\ell_1$ regularization (Friedman et al., 2010), and compressed sensing (Sanandaji et al., 2012). In these papers it is assumed that each node in the network is driven by an unknown, independent stochastic process, each variable is measured without sensor noise, and every variable in the network is measured. It is shown that under these conditions topology detection is possible.

For many networks in engineering, the interconnection structure is known. This knowledge can be incorporated in the identification problem. A type of interconnection structure that results from the discretization of partial differential equations is a *spatially distributed system* where each node is connected only to its direct neighbors. Methods for identifying such systems are presented in Massioni & Verhaegen (2008); Haber & Verhaegen (2012); Ali et al. (2011), where common assumptions are that each subsystem is identical, known external excitation signals are present at each node, and no process

noise is present in the networks. Because such networks are typically very large, emphasis is on improving computational speed of the identification algorithms.

Identification in networks with a general interconnection structure have been investigated in Van den Hof et al. (2013); Dankers et al. (2014a, 2013a),where methods are presented to consistently identify a single transfer function embedded in a dynamic network. It is shown that by knowing the interconnection structure assumptions on the correlation of process noise can be relaxed, and that there is considerable flexibility in which variables need to be measured. In these papers the measurements are assumed to be sensor noise free.

Variance issues of identified models in a structured network have been addressed in Wahlberg et al. (2009); Gunes et al. (2014); Everitt et al. (2013, 2014) where is shown that "extra" measurements can be used to reduce the variance of the estimated transfer function. These papers assume that there is no process noise, and known external excitation and sensor noise are both present.

In this paper we consider a very general framework that covers all the cases discussed in the previous literature review, where there may or may not be known external excitation present, there is both (correlated) process noise and (correlated) sensor noise present, the modules making up the network are not identical, and not all internal variables of the network are measurable. Moreover, we do not make assumptions on the whiteness of the sensor noise. The main assumption that we make is that the interconnection structure of the network is known. We address the following question: under what conditions is it possible to consistently identify a particular module embedded in a dynamic network when only noisy measurements of a subset of the internal variables of the network are available? This is an extension of the so-called Errors-in-Variables (EIV) framework to the case of dynamic networks.

In the system identification literature, the open loop EIV problem has been extensively studied, see e.g. Söderström (2007, 2012). The main conclusion in these papers is that either prior knowledge about the system or a controlled experimental setup is required to ensure consistent estimates. This latter condition concerns either periodic excitation (Söderström & Hong, 2005; Schoukens et al., 1997; Pintelon & Schoukens, 2012) or repeated experiments Schoukens et al. (1997); Pintelon & Schoukens (2012). The closed-loop EIV problem has been studied in Söderström et al. (2013); Pintelon & Schoukens (2012) where it is shown that the plant is identifiable if a noise-free and sufficiently exciting reference signal is available.

In the extension of this problem to the dynamic network case fruitful use can be made of additionally measured signals in the network that can serve as instrumental variables, thereby enabling a considerable simplification of the EIV problem. The method presented in this paper is based on *instrumental variable* (IV) reasoning.

The IV method was developed in the econometrics field (Wright, 1928), where the method has been applied to static networks (*structural equation models* in statistics) (Angrist et al., 1996). In the econometrics literature IV methods are recognized to have three main advantages when aiming to obtain consistent estimates:

**(1).** Presence of sensor noise on the input (*explanatory variable* in economics) is no problem (Durbin, 1954);

**(2).** Confounding variables (*omitted variables* in econometrics), i.e. unknown or unmeasured variables for which there is a path to both the output and an input, are no problem (Angrist & Krueger, 2001; Becker, 2010);

**(3).** Presence of algebraic loops in the data generating system (*simultaneity* in econometrics) is no problem (Becker, 2010).

In this paper we show that the same advantages can be converted to the situation of a dynamic network, and moreover that the choice of candidate instrumental variable signals actually can be widened.

In the system identification literature IV methods are also extensively used for identification in open-loop (Wong & Polak, 1967; Söderström & Stoica, 1983), and closed-loop systems (Söderström et al., 1988; Söderström & Stoica, 1989; Gilson & Van den Hof, 2005). Again, IV methods have been recognized to be robust to the presence of (particular) sensor noise on the input Söderström & Hong (2005); Thil & Gilson (2011). In this paper we generalize the IV method such that it is possible to obtain consistent estimates of a transfer function embedded in a dynamic network where all predictor inputs are measured with (colored) sensor noise, and the instrumental signal(s) are contaminated too.

In Section 2 background material on dynamic networks, prediction error identification and IV methods is presented. In Sections 3 and 4 the main result is presented for two different cases of IV signals. In Section 5 results are generalized for a flexible choice of predictor inputs, while in Section 6 a practical implementation of the method is proposed. In Section 7 a method is presented to validate the obtained model. [1]

## 2 Background

### 2.1 Dynamic Networks

The specific identification framework considered in this paper is based on Van den Hof et al. (2013). A dynamic network is built up of $L$ elements, related to $L$ scalar *internal variables* $w_j$, $j = 1, \ldots, L$. Each internal variable is defined by:

$$w_j(t) = \sum_{k \in \mathcal{N}_j} G_{jk}^0(q) w_k(t) + r_j(t) + v_j(t) \qquad (1)$$

---

[1] This paper is based, in part, on the preliminary results of Dankers et al. (2014b).

where $G_{jk}^0$, $k \in \mathcal{N}_j$ is a proper rational transfer function, $q^{-1}$ is the delay operator, i.e. $q^{-1}w_j(t) = w_j(t-1)$ and,

- $\mathcal{N}_j$ is the set of indices of internal variables that are direct inputs to the transfer functions determining $w_j$, i.e. $k \in \mathcal{N}_j$ iff $G_{jk}^0 \neq 0$;
- $v_j$ is *process noise*, that is modeled as a realization of a stationary stochastic process with rational spectral density: $v_j = H_j^0(q)e_j$ where $e_j$ is a white noise process, and $H_j^0$ is a monic, stable, minimum phase transfer function;
- $r_j$ is an *external variable* that is known to the user, and may be manipulated by the user.

It may be that the noise and/or external variables are not present at some nodes. The network is defined by:

$$
\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} = \begin{bmatrix} 0 & G_{12}^0 & \cdots & G_{1L}^0 \\ G_{21}^0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & G_{L-1\,L}^0 \\ G_{L1}^0 & \cdots & G_{L\,L-1}^0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} + \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_L \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_L \end{bmatrix},
$$

where $G_{jk}^0$ is non-zero if and only if $k \in \mathcal{N}_j$ for row $j$, and $v_k$ (or $r_k$) is zero if it is not present. Using an obvious notation this results in the matrix equation:

$$
w = G^0 w + r + v. \tag{2}
$$

Each internal variable is measured with some measurement error or sensor noise:

$$
\tilde{w}_k(t) = w_k(t) + s_k(t), k = 1, \ldots, L
$$

where $\tilde{w}_k$ denotes the measurement of $w_k$, and $s_k$ is the *sensor noise*, which is represented by a stationary stochastic process with rational spectral density ($s_k$ is not necessarily white noise).

There exists a *path* from $w_i$ to $w_j$ if there exist integers $n_1, \ldots, n_k$ such that $G_{jn_1}^0 G_{n_1 n_2}^0 \cdots G_{n_k i}^0$ is non-zero.

The following assumption holds throughout the paper.

**Assumption 1** *General Conditions.*

(a) *The network is well-posed in the sense that all principal minors of $\lim_{z \to \infty}(I - G^0(z))$ are non-zero.*
(b) $(I - G^0)^{-1}$ *is stable.*
(c) *All process noise variables $v_k$ are uncorrelated to all sensor noise variables $s_\ell$.* [2]

---

[2] in this paper when two variables $x$ and $y$ are said to be uncorrelated, this will mean that the cross-correlation between $x$ and $y$, $R_{xy}(\tau)$ is zero for all $\tau$, where $R_{xy}$ is defined in (5).

The conditions ensure that there is a stable causal relationship from the external and noise variables to the internal variables.

## 2.2 Prediction Error Identification

Let $w_j$ denote the variable which is to be predicted. The one-step-ahead predictor for $w_j$ is then (Ljung, 1999):

$$
\hat{w}_j(t|t-1,\theta) = H_j^{-1}(q,\theta)\Big(\sum_{k \in \mathcal{D}_j} G_{jk}(q,\theta)\tilde{w}_k(t) + r_j(t)\Big) + \Big(1 - H_j^{-1}(q,\theta)\Big)\tilde{w}_j(t) \tag{3}
$$

where $H_j(q,\theta)$ is a (monic) noise model and $G_{jk}(q,\theta)$, $k \in \mathcal{D}_j$ are module models. The *predictor inputs* are those variables that are used to predict $w_j$. The set $\mathcal{D}_j$ denotes the set of indices of the measured internal variables that are chosen as predictor inputs, i.e. $\tilde{w}_k$ is a predictor input iff $k \in \mathcal{D}_j$. The prediction error with respect to the parameter $\theta$ is:

$$
\begin{aligned}
\varepsilon_j(t,\theta) &= \tilde{w}_j(t) - \hat{w}_j(t|t-1,\theta) \\
&= H_j(\theta)^{-1}\Big(\tilde{w}_j - \sum_{k \in \mathcal{D}_j} G_{jk}(\theta)\tilde{w}_k - r_j\Big),
\end{aligned} \tag{4}
$$

where the arguments $q$ and $t$ have been dropped for notational clarity. Usually the parameterized transfer functions $G_{jk}(\theta)$, $k \in \mathcal{D}_j$ and $H_j(\theta)$ are estimated by minimizing the sum of squared (prediction) errors. Let $\hat{\theta}_N$ denote the estimated parameter vector based on $N$ data points. If $\hat{\theta}_N \to \theta^0$ as $N \to \infty$ with probability 1, then the obtained estimates are *consistent*.

The following notation will be used throughout the remainder of this paper. The auto and cross correlation of vectors of variables $x$ and $y$ are defined as

$$
\begin{aligned}
R_x(\tau) &:= \bar{\mathbb{E}}[x(t)x^T(t-\tau)], \\
R_{xy}(\tau) &:= \bar{\mathbb{E}}[x(t)y^T(t-\tau)]
\end{aligned} \tag{5}
$$

respectively, where $\bar{\mathbb{E}}[\,\cdot\,] = \lim_{N \to \infty} \frac{1}{N}\sum_{t=0}^{N-1} \mathbb{E}[\,\cdot\,]$, and $\mathbb{E}$ denotes the expected value operator. The power spectral density and cross power spectral density are

$$
\Phi_x(\omega) := \mathcal{F}[R_x(\tau)] \text{ and } \Phi_{xy}(\omega) := \mathcal{F}[R_{xy}(\tau)]
$$

respectively, where $\mathcal{F}[\,\cdot\,]$ denotes the Fourier Transform.

## 2.3 Instrumental Variable Method

The instrumental variable (IV) method for identification of dynamical systems typically applies a model structure

where the one-step-ahead predictor (3) is written in a linear regression form:

$$\hat{w}_j(t|t-1,\theta) = \varphi^T(t)\theta$$

where $\varphi(t)$ is composed of measured signals (predictor inputs) and $\theta$ collects all unknown parameters. This is generally referred to as an ARX model structure (Ljung, 1999) when also past values of $\tilde{w}_j(t)$ are included in $\varphi(t)$. The prediction error can then be written as

$$\varepsilon_j(t,\theta) = w_j(t) - \phi^T(t)\theta. \qquad (6)$$

The parameter estimate is obtained by solving for

$$\frac{1}{N}\sum_{t=0}^{N-1}\varepsilon_j(t,\theta)z(t-\tau)=0, \ \tau=0,\ldots,n_\theta \qquad (7)$$

where $z$ is a measured *instrumental variable* that typically is required to be correlated to the predictor inputs and uncorrelated to the process noise $v_j$ (Söderström & Stoica, 1983), and $n_\theta$ is chosen sufficiently large, i.e. at least as large as the dimension of $\theta$.

In the closed-loop situation, the Basic Closed-Loop Instrumental Variable (BCLIV) method as described in Söderström et al. (1988); Gilson & Van den Hof (2005) is characterized by choosing an external reference signal $r$ as the instrumental variable and not considering any sensor noise. Under conditions on the excitation properties of $r$, and the general prediction error convergence assumptions, this IV estimate then provides a consistent estimate of the plant transfer function, see Gilson & Van den Hof (2005).

## 3 An IV Method for Dynamic Networks and sensor noise

In this section an extension of the BCLIV method is presented. The extension focuses on three aspects:

- We generalize the method so that it is able to identify a particular module embedded in a dynamic network, not just a closed-loop data generating system.
- We consider the situation that all measured variables can be subject to sensor noise.
- Rather than the classical case where only external variables are considered as candidate instrumental variables, we consider both internal and external variables as candidate instrumental variables.

A main theme in this paper is that for a dynamic network, there are many different variables present that can serve as potential instrumental variables. For instance, one can choose between several external and internal variables. In this paper we consider any measured or known variable that is not $\tilde{w}_j$ or a predictor input as a potential instrumental variable. In other words, the set of candidate instrumental variables is $\tilde{w}_\ell, \ell \in \{1,\ldots,L\} \setminus \{\mathcal{D}_j \cup \{j\}\}$ and $r_\ell, \ell \in \mathcal{R}$. Let $\mathcal{X}_j$ and $\mathcal{I}_j$ denote the set of indices of external and internal variables respectively chosen as instrumental variables (i.e. $r_\ell$ is an instrumental variable iff $\ell \in \mathcal{X}_j$ and $\tilde{w}_\ell$ is an instrumental variable iff $\ell \in \mathcal{I}_j$). Since predictor inputs and $\tilde{w}_j$ are not considered as allowable instrumental variables it must be that $\mathcal{I}_j \cap \{\mathcal{D}_j \cup \{j\}\} = \emptyset$.

The variables that are selected as instrumental variables are placed in a vector of instrumental variables, denoted $z$, where the elements of $z$ can be excitation signals, measured variables, or linear combinations of both. The particular choice will depend on which variables are available/measured, and which choice ensures that $z$ is persistently exciting of sufficiently high order.

In any IV method, it is essential that the instrumental variables and the predictor inputs are correlated. In the case of dynamic networks it is not automatically guaranteed that a candidate internal variable is correlated to (one or more of) the predictor inputs and/or $w_j$. The following lemma presents graphical conditions to check whether two variables are (not) correlated.

**Lemma 1** *Consider a dynamic network as defined in Section 2.1 that satisfies Assumption 1. Let $z_\ell$ be an internal or external variable. Then $z_\ell$ and $w_k$ are not correlated if the following three conditions hold:*

*(a) There is no path from $z_\ell$ to $w_k$.*
*(b) There is no path from $w_k$ to $z_\ell$.*
*(c) There is no variable $w_p$, $p \notin \mathcal{D}_j \cup \mathcal{I}_j \cup \{j\}$ such that there are paths from $w_p$ to both $z_\ell$ and to $w_k$.* $\qquad \square$

The proof can be found in Appendix A. Note that the absence of correlation can be stated in terms of the interconnection structure of the network, but the presence of correlation cannot be guaranteed by the structure since it depends on the dynamics of the paths. However, Lemma 1 can still be used to guide the user to choose appropriate instrumental variables that are correlated to the predictor inputs.

In order to formulate an IV method in a dynamic network setting, it is necessary to move to a multiple input, single output ARX model structure. In that situation the modules and noise model are parameterized as:

$$G_{jk}(q,\theta) = \frac{B_{jk}(q,\theta)}{A_j(q,\theta)}, \text{ and } H_j(q,\theta) = \frac{1}{A_j(q,\theta)}, \quad (8)$$

for all $k \in \mathcal{D}_j$, where

$$B_{jk}(q,\theta) = q^{-n_k^{jk}}(b_0^{jk} + b_1^{jk}q^{-1} + \cdots + b_{n_b^{jk}}^{jk}q^{n_b^{jk}}),$$

$$A_j(q,\theta) = 1 + a_1q^{-1} + \cdots + a_{n_a}q^{-n_a},$$

Note that all modules $G_{jk}$, $k \in \mathcal{D}_j$ have the same denominator, and that $B_{jk}(\theta)$ is a polynomial of order $n_b^{jk}$

and $A_j(\theta)$ is a polynomial of order $n_a$. For notational convenience all polynomials $B_{jk}(\theta)$ will be assumed to be of the same order, denoted $n_b$. Let each module $G_{jk}^0$, $k \in \mathcal{D}_j$ be expressed as $\frac{B_{jk}^0}{A_{jk}^0}$. Then, from (1), $w_j$ can be expressed using transfer functions with a common denominator as follows:

$$w_j(t) = \frac{1}{\breve{A}_j^0(q)} \sum_{k \in \mathcal{N}_j} \breve{B}_{jk}^0(q) w_k(t) + r_j(t) + v_j(t) \quad (9)$$

where

$$\breve{A}_j^0(q) = \prod_{n \in \mathcal{N}_j} A_{jn}^0(q) \text{ and } \breve{B}_{jk}^0(q) = \prod_{n \in \mathcal{N}_j \setminus k} B_{jk}^0(q) A_{jn}^0(q).$$

Then with (8) the one-step-ahead prediction error is

$$\begin{aligned}
\varepsilon_j(\theta) &= A_j(q,\theta)\tilde{w}_j(t) - \sum_{k \in \mathcal{D}_j} B_{jk}(q,\theta)\tilde{w}_k(t) \\
&= \tilde{w}_j(t) - \left[ \tilde{\phi}_{k_1}^T(t) \cdots \tilde{\phi}_{k_n}^T(t) \ \tilde{\phi}_j^T(t) \right] \theta \\
&= \tilde{w}_j - \tilde{\phi}^T(t)\theta. \quad (10)
\end{aligned}$$

where $\tilde{\phi}_{k_i}^T(t) = [\tilde{w}_{k_i}(t) \cdots \tilde{w}_{k_i}(t-n_b)]$, $\tilde{\phi}_j^T(t) = [-\tilde{w}_j(t-1) \cdots -\tilde{w}_j(t-n_a)]$ and $\theta$ is a vector of parameters:

$$\theta = [b_0^{jk_1} \cdots b_0^{jk_n} \cdots b_{n_b}^{jk_1} \cdots b_{n_b}^{jk_n} \ a_1 \cdots a_{n_a}]^T.$$

The mechanism that forms the foundation for the IV method applied to dynamic networks is presented in the following Proposition.

**Theorem 2** *Consider a dynamic network as defined in Section 2.1 that satisfies Assumption 1. Consider the prediction error (10). Choose the set of predictor inputs $\{\tilde{w}_k, k \in \mathcal{D}_j\}$, being all measured signals that are direct inputs of the transfer functions determining $w_j$, i.e. $\mathcal{D}_j = \mathcal{N}_j$, and let $d = card(\mathcal{D}_j)$. Choose a set of instrumental variable signals $\{r_\ell, \ell \in \mathcal{X}_j\}$ and $\{\tilde{w}_\ell, \ell \in \mathcal{I}_j\}$ such that $\mathcal{I}_j \cap \{\mathcal{D}_j \cup \{j\}\} = \emptyset$. The equivalence relation*

$$\begin{aligned}
&\left\{ R_{\varepsilon z}(\tau) = 0, \text{ for } \tau = 0, \dots n_s \right\} \\
&\iff \left\{ G_{jk}(q,\theta) = G_{jk}^0(q), \forall k \in \mathcal{D}_j \right\} \quad (11)
\end{aligned}$$

*holds for any finite $n_s \geq \lceil (n_a + dn_b)/length(z(t)) \rceil$ if the following conditions are satisfied:[3]*

(a) *If $v_j$ is present, then there is no path from $w_j$ to any $w_\ell$, $\ell \in \mathcal{I}_j$*

(b) *The $(dn_b + n_a) \times (n_z length(z))$ matrix*

$$\bar{R} = \bar{\mathbb{E}}\left[ \tilde{\phi}(t)[z^T(t) \cdots z^T(t-n_z)] \right]$$

*is full row rank, where $\tilde{\phi}(t)$ is defined in (10).*
(c) *Each sensor noise $s_\ell$, $\ell \in \mathcal{I}_j$ is uncorrelated to all $s_k$, $k \in \mathcal{D}_j$.*
(d) *If $v_j$ is present, then it is uncorrelated to all $v_m$ with a path to any one of the instrumental variables $\tilde{w}_\ell$, $\ell \in \mathcal{I}_j$.*
(e) *The parameterization is flexible enough, i.e. there exists at least one $\theta^\star$ such that $G_{jk}(q,\theta^\star) = G_{jk}^0(q)$, $\forall k \in \mathcal{D}_j$.* □

The proof can be found in Appendix B. Most importantly, the presence of sensor noise does not affect the validity implication (11) (as long as Condition (c) holds). Condition (a) puts a restriction on which internal variables are candidate instrumental variables. All external signals qualify as well as all internal signals that do not serve as input/ouput of the predictor model and that are not in a loop that passes through $w_j$. Note that the process noise terms on each predictor input variable can be correlated. However, by Condition (d) the process noise term $v_j$ must be uncorrelated to all noise terms affecting the instrumental variables. This condition can also help in guiding a proper choice of instrumental variables.

Condition (b) of Theorem 2 can be further analyzed using the concept of persistence of excitation. If no instrumental variables are selected (or available) then Condition (b) is not satisfied. Next, suppose that none of the selected instrumental variables satisfy all conditions of Lemma 1. This does not guarantee that the matrix $\bar{R}$ of Condition (b) of Theorem 2 is full rank. For the condition to hold, the vector $z$ of instrumental variables is required to be persistently exciting of sufficiently high order, as shown in the following lemma.

**Lemma 3** *Consider the situation of Theorem 2 and*

$$\bar{R} = \bar{\mathbb{E}}\left[ \tilde{\phi}(t) \left[ z^T(t) \ z^T(t-1) \cdots z^T(t-n_z) \right] \right]$$

*where $z$ is the vector of instrumental variables, and $\tilde{\phi}(t)$ is defined in (10). Let $n_\theta$ denote the size of the vector $\tilde{\phi}(t)$. The matrix $\bar{R}$ generically has rank $n_\theta$ (i.e. full row rank) if the following conditions hold:*

(a) *$n_z \cdot length(z(t)) \geq n_\theta$.*
(b) *$z$ is persistently exciting of order $\lceil n_\theta/length(z) \rceil$.[4]*
(c) *The parameter vector $\theta$ such that $G_{jk}(q,\theta) = G_{jk}^0$, $\forall k \in \mathcal{D}_j$ is unique.*

---

[3] $\lceil x \rceil$ denotes the ceiling function, i.e. the smallest integer that is larger than or equal to $x$

[4] A vector signal $z$ is persistently exciting of order $n$ if the $n \times n$ block Toeplitz matrix of the autocorrelation function $R_z(\tau)$ is positive definite (Söderström & Stoica, 1989).

(d) *No instrumental variable is uncorrelated to all $w_k$, $k \in \mathcal{D}_j \cup \{j\}$ according to the conditions of Lemma 1.*

The proof follows the same reasoning as that of Lemma 4.1 and Theorem 4.1 in Söderström & Stoica (1983).

The main point of Lemma 3 is that as long as the instruments are correlated to the predictor inputs and $w_j$, and are persistently exciting of sufficiently high order, then Condition (b) of Theorem 2 generically holds. There is no explicit restriction on the number of instrumental variables, as long as the chosen $z$ is persistently exciting of sufficiently high order. However, if only one internal variable is selected as the instrumental variable, then by Condition (b) of Lemma 3 $z$ must be persistently exciting of order $n_\theta$. Whereas, if two internal variables are selected as instrumental variables then by the same condition $z$ need only be persistently exciting of order $n_\theta/2$. Thus Condition (b) may implicitly place a restriction on the required number of instrumental variables.
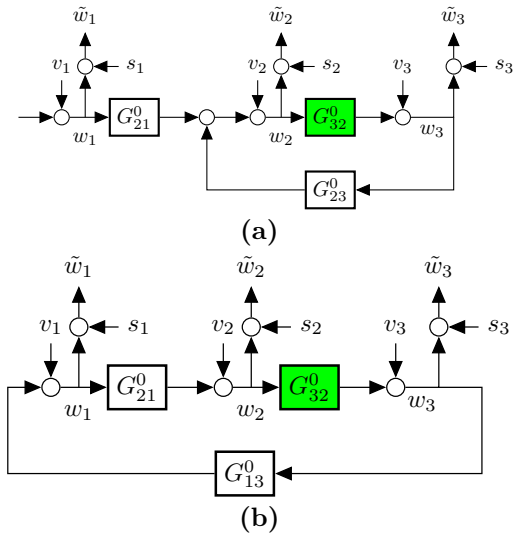


**(a)**



**(b)**

Fig. 1. Closed loop data generating systems

**Example 1** *Consider the data generating system shown in Fig. 1a. Suppose that the objective is to obtain a consistent estimate of $G_{32}^0$. Thus, $\{j\} = \{3\}$, and $\mathcal{N}_3 = \{2\}$. There is only one variable left as a candidate instrumental variable (since it must be that $\{\mathcal{D}_j \cup \{j\}\} \cap \mathcal{I}_j = \emptyset$), i.e. $\tilde{w}_1$ must be chosen as the instrumental variable. Since there is no path from $w_3$ to $w_1$, Condition (a) of Theorem 2 holds. Moreover, Condition (b) of Theorem 2 generically holds because the instrumental variable is persistently exciting of sufficiently high order (since $v_1$ is white noise) and because there is a path from $w_1$ to both $w_2$ and $w_3$ (i.e. there is a path from the instrumental variable to the predictor inputs and $w_j$). If the remaining conditions of Theorem 2 hold, then the implication (11) holds.*  □

**Example 2** *Consider the data generating system shown in Fig. 1b. Suppose that the objective is to obtain a consistent estimate of $G_{32}^0$. In this case it is not possible to*

*satisfy Condition (a) of Theorem 2, since there exists a path from output signal $w_3$ to the only candidate instrumental signal $w_1$.*  □

The following algorithm shows how the implication of Theorem 2 can be exploited to obtain an estimate of a module embedded in a dynamic network.

**Algorithm 1** *Objective: obtain an estimate of $G_{jk}^0$.*

1. *Choose the set of predictor inputs $\{w_{k_n}, k_n \in \mathcal{N}_j\}$. (i.e. $\mathcal{D}_j = \mathcal{N}_j$).*
2. *Choose the sets $\mathcal{I}_j$ and $\mathcal{X}_j$ of instrumental variables. Construct $z$, the vector of instrumental variables.*
3. *Choose an ARX model structure and construct the prediction error (10).*
4. *Find a solution, $\hat{\theta}_N$ to the set of equations*

$$\frac{1}{N} \sum_{t=0}^{N-1} \varepsilon_j(t,\theta) z^T(t-\tau) = 0, \ for \ \tau = 0, \ldots, n_z, \quad (12)$$

*where $n_z \cdot length\big(z(t)\big) \geq n_a + dn_b$.*

Let $\hat{R}_{\varepsilon z}(\tau)$ denote the function in (12). Under weak general convergence conditions of the Prediction-Error Identification methods (Ljung, 1999) it follows that

$$\mathbb{E}[\hat{R}_{\varepsilon z}(\tau)] \to R_{\varepsilon z}(\tau) \text{ as } N \to \infty \quad (13)$$

and that the solution to (12), denoted $\hat{\theta}_N$ tends to $\theta_0$ as $N \to \infty$. Thus, the estimates of $G_{jk}^0$, $k \in \mathcal{N}_j$ obtained by Algorithm 1 are consistent if the conditions presented in Theorem 2 are satisfied. In Step 4 of Algorithm 1 $\hat{\theta}_N$ can be obtained by linear regression. This follows from (12) which is affine in $\theta$.

In the next section this method is generalized to allow for instrumental variables to appear in a loop around the output signal $w_j$ (i.e. relax Condition (a) of Theorem 2).

## 4 Generalized Instrumental Variable Approach

In the previous section the set of candidate instrumental variables is restricted by Condition (a) of Theorem 2. In this section a method is presented for which all external variables and all internal variables $w_\ell$, $\ell \notin \mathcal{D}_j \cup \{j\}$ are candidate instrumental variables.

The key difference in this method is that a Box-Jenkins model structure is used instead of an ARX model structure in order to correctly model the process noise also. The price for the increased applicability is that the estimates of $G_{jk}^0$ $k \in \mathcal{N}_j$ can no longer be obtained by solving a linear regression problem.

The main reason that a path from output $w_j$ to an instrumental variable $\tilde{w}_i$ causes a problem is because then the

projections of the predictor inputs onto the instrumental variable(s) are correlated to the output noise. This is equivalent to the closed-loop identification problem where the plant input is correlated to the output noise, and several solutions are available (Forssell & Ljung, 1999; Van den Hof et al., 2013). The Direct Method deals with the problem by exactly modeling the noise.

Here it is shown that this idea can be extended to the IV framework, so that all (measured) internal variables $\tilde{w}_\ell$ $\ell \in \{1, \ldots, L\} \setminus \{\mathcal{D}_j \cup \{j\}\}$ are candidate instrumental variables. Note that the idea is to exactly model the process noise term $v_j$, and not the sensor noise (or a sum of the two). The sensor noise is dealt with using the instrumental variable mechanism.

The Box-Jenkins (BJ) model structure is represented by:

$$G_{jk}(q,\theta) = \frac{B_{jk}(q,\theta)}{F_{jk}(q,\theta)}, k \in \mathcal{D}_j \text{ and } H_j(q,\theta) = \frac{C_j(q,\theta)}{D_j(q,\theta)}, \quad (14)$$

where $F_{jk}(\theta)$, $B_{jk}(\theta)$, $C_j(\theta)$, $D_j(\theta)$ are polynomials in $q^{-1}$ of orders $n_f$, $n_b$, $n_c$ and $n_d$ respectively (similar to Section 3 for notational convenience we assume that all module tranfer functions have the same orders). The polynomials $F_{jk}, C_j,$ and $D_j$ are monic. In addition we will assume that the numerators and denominators in all $G_{jk}(q,\theta)$ and $H_j(q,\theta)$ are the same order, i.e. $n_c = n_d$ and $n_f = n_b$. Similar to (10) denote the regression matrix $\phi(t)$ as

$$\phi^T(t) = [\phi_{k_1}^T(t) \cdots \phi_{k_d}^T(t) \phi_j^T(t)], \{k_1, \ldots, k_d\} = \mathcal{D}_j,$$
$$\phi_k^T(t) = [w_k(t) \cdots w_k(t-n_x)], \ k \in \mathcal{D}_j,$$
$$\phi_j^T(t) = [w_j(t-1) \cdots w_j(t-n_x)], \quad (15)$$

where $n_x$ is a finite integer. In the following proposition it is shown that by changing the model structure from ARX to BJ, the fundamental mechanism on which the IV methods are based still holds.

**Theorem 4** *Consider the same situation as in Theorem 2 with the prediction error (4), the model structure given by (14), and $\mathcal{D}_j = \mathcal{N}_j$. Let $z$ denote the vector of instrumental variables:*

$$z^T(t) := \begin{bmatrix} \tilde{w}_{\ell_1}(t-d_1) & \cdots & \tilde{w}_{\ell_n}(t-d_n) & r_{m_1}(t) & \cdots & r_{m_p}(t) \end{bmatrix}$$

*where $d_k$, $k = 1, \cdots n$, are chosen so that $R_{e_j \tilde{w}_{\ell_k}}(\tau) = 0$ for $\tau \geq d_k$. The equivalence relation*

$$\left\{ R_{\varepsilon z}(\tau) = 0, \ for \ \tau = 0, \ldots, n_s \right\}$$
$$\iff \begin{cases} G_{jk}(q,\theta) = G_{jk}^0(q), \ \forall k \in \mathcal{D}_j, \\ H_j(q,\theta) = H_j^0(q) \end{cases} \quad (16)$$

*holds if the following conditions are satisfied:*

(a) *The matrix $M = \bar{\mathbb{E}}\big[\phi(t)[z^T(t-n_x) \cdots z^T(t-n_x - n_z)]\big]$ has full row rank, where $\phi$ is defined in (15), $n_z$ is any finite integer large enough so that $M$ has full row rank, and $n_x = 2dn_f + 2n_c$.*

(b) $n_s \geq n_z + n_x$.

(c) *Every sensor noise variable $s_k$, $k \in \mathcal{D}_j \cup \{j\}$ is uncorrelated to every $s_\ell$, $\ell \in \mathcal{I}_j$.*

(d) *The process noise variable $v_j$ is uncorrelated to all $v_k$ with a path to either $w_j$ or any $w_\ell$, $\ell \in \mathcal{I}_j$.*

(e) *The parameterization is chosen flexible enough, i.e. there exists a parameter $\theta$ such that $G_{jk}(q,\theta) = G_{jk}^0(q)$, $\forall k \in \mathcal{D}_j$, and $H_j(q,\theta) = H_j^0(q)$.* □

The proof can be found in Appendix C. The requirement on $R_{e_j \tilde{w}_{\ell_k}}(\tau)$ reflects that there should not be an algebraic connection between $e_j(t)$ and $\tilde{w}_{\ell_k}(t-d_k)$. If there is a delay in the path from $w_j$ to the instrumental variable $\tilde{w}_{\ell_k}$, then the condition is satisfied for $d_k = 0$. If there is no delay in that path then one can choose $d_{\ell_k} > 0$.

By Condition (e) the process noise must be exactly modelled. This condition is a signature of the Direct closed-loop method Forssell & Ljung (1999); Van den Hof et al. (2013). Thus, the mechanism proposed in Theorem 4 is a hybrid between the Direct closed-loop method and an instrumental variable method. Recall that in Theorem 2 exact noise modeling was not required.

In Theorem 4 Condition (a) is a condition on the excitation of the data. A necessary condition for Condition (a) to hold is that no instrumental variable satisfies all the conditions of Lemma 1 for all $w_k$, $k \in \mathcal{D}_j \cup \{j\}$. The requirement that the matrix of Condition (a) must have full row rank implicitly puts a constraint on $n_m$. In particular it must be that the matrix has at least as many columns as rows.

**Example 3** *Consider again the situation of Example 2. Suppose that there is a delay in $G_{13}^0$. Choose, $\{j\} = \{3\}$, $\mathcal{N}_2 = \{2\}$. Choose $w_1$ as the instrumental variable, i.e. $z(t) = \tilde{w}_1(t)$, $\mathcal{I}_j = \{1\}$, and $\mathcal{X}_j = \emptyset$. Note that due to the delay in $G_{13}^0$, the instrumental variable is a function of strictly delayed versions of $v_j$, as required by Theorem 4. By Lemma 1, since there is a path from $w_1$ to both $w_2$ and $w_3$ the necessary conditions for Condition (a) to hold are satisfied. If the remaining conditions of Theorem 4 are satisfied, then the implication (16) holds. Since the signal $w_1$ appears in a loop with $w_3$, instrumental variable $\tilde{w}_1$ does not satisfy the conditions of Theorem 2.* □

**Example 4** *Consider the network shown in Fig. 2. Suppose that the objective is to obtain a consistent estimate of $G_{21}^0$. Thus, $\{j\} = \{2\}$, and $\mathcal{N}_2 = \{1, 4, 6\}$. A possible choice for the set of instrumental variables is $\tilde{w}_3(t)$, $\tilde{w}_5(t)$, $\tilde{w}_7(t)$ and $\tilde{w}_8(t)$ (i.e. $\mathcal{I}_2 = \{3, 5, 7, 8\}$). In this case if $z(t) = [\tilde{w}_3(t) \ \tilde{w}_5(t) \ \tilde{w}_7(t) \ \tilde{w}_8(t)]^T$ is persistently exciting of sufficiently high order, and the remaining conditions of Theorem 4 hold, then the implication (16) holds.*
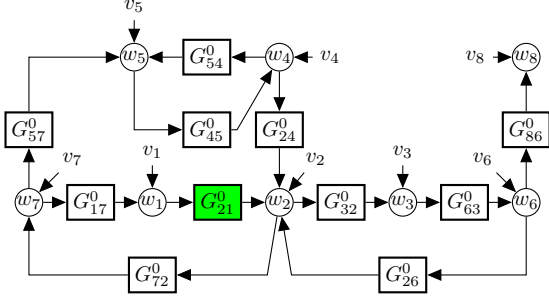
Fig. 2. Example of a dynamic network. The sensor noise is not shown, but is still assumed to be present on every measured variable. The labels of the $w_i$'s have been placed inside the summations indicating that the output of the sum is $w_i$.

*Another option for choosing the instrumental variables is to only use $\tilde{w}_7$. In this case if $z'(t) = [\tilde{w}_7(t)\ \tilde{w}_7(t-n_g-1)\ \tilde{w}_7(t-2n_g-1)\ \tilde{w}_7(t-3n_g-1)]$ is persistently exciting of sufficiently high order, and the remaining conditions of Theorem 4 hold, then the implication (16) holds. Other options for choosing $z$ are also possible depending on the persistence of excitation of $z$.* □

The result of Theorem 4 can be cast in the following identification algorithm.

**Algorithm 2** *Objective: obtain an estimate of $G_{jk}^0$.*

1. *Choose the set of predictor inputs as $\mathcal{D}_j = \mathcal{N}_j$.*
2. *Choose the set $\mathcal{I}_j$ and construct the vector of instrumental variables, $z$.*
3. *Choose a Box-Jenkins model structure, (14), and construct the prediction error (4).*
4. *Find a solution to the set of equations*

$$\frac{1}{N}\sum_{t=0}^{N-1}\varepsilon(t,\theta)z(t-\tau) = 0,\ for\ \tau = 1,\ldots,n_s.\quad (17)$$

By Theorem 4 and the reasoning of (13) it follows that the estimate obtained using Algorithm 2 is consistent, as long as all the conditions of the proposition are satisfied.

We now highlight the major properties of the two methods presented in Sections 3 and 4 in Table 1.

| Method Section 3 | Method Section 4 |
|---|---|
| MISO ARX | MISO BJ |
| Convex | Non-convex |
| No path from $w_j$ to IV signals | IV signals have path from $w_j$ |
| No noise modelling | Full noise modelling |

Table 1
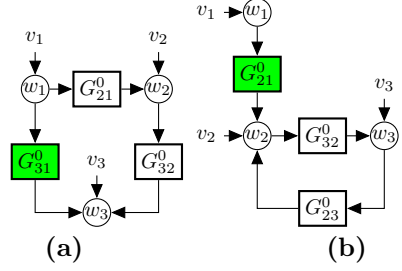 Overview of key properties of the two presented identification methods.



Fig. 3. Simple dynamic networks used to illustrate the effect of predictor input selection .

In the following section the choice $\mathcal{D}_j = \mathcal{N}_j$ will be relaxed, in other words it will be allowed to remove particular signals from the set of predictor inputs, and include signals that are not in $\mathcal{N}_j$.

## 5  Predictor Input Selection

In this paper thus far, the required set of predictor inputs for identifying $G_{ji}^0$ has been fixed to include all internal variables with direct connections to $w_j$, i.e. $\tilde{w}_k$, $k \in \mathcal{N}_j$. This is an overly restrictive requirement. There is a strong motivation to seek less restrictive conditions. For instance it may be that several internal variables are difficult or expensive to measure and it would be preferred to avoid the necessity of measuring these variables.

The problem of predictor input selection has been analyzed in Dankers et al. (2013a,b) where conditions are derived for consistent identification of $G_{ji}^0$ for the direct and two-stage method of identification in dynamic networks. In this section those ideas are extended to the methods of Sections 3 and 4.

The conditions can be used to determine whether it is possible to consistently identify a particular transfer function given a set of measured internal variables. Conversely, the conditions can also be used to design a sensor placement scheme to ensure that a particular transfer function can be consistently estimated. The sensor placement scheme could be designed to minimize the number of sensors, or it could be designed to avoid using a particular variable that is difficult to measure. The cost of the increased flexibility in the choice of predictor inputs is that instead of consistently estimating $G_{jk}^0$, $k \in \mathcal{N}_j$, only the module of interest $G_{ji}^0$ is consistently estimated.

When attempting to identify $G_{ji}^0$, $w_i$ must be included as a predictor input. Interestingly, the other variables that are included as predictor inputs affect the dynamics that are identified between $w_i$ to $w_j$. There are two phenomena that result in a change in the object being identified, illustrated in the following two examples.

Firstly, consider the system shown in Figure 3a. Suppose that the objective is to identify $G_{31}^0$. If $w_1$ and $w_3$

8

are both included as predictor inputs, then the dynamics from $w_1$ to $w_3$ are $G_{31}^0$. However, if only $w_1$ is chosen as a predictor input, then the identified transfer function is an estimate of $G_{31}^0 + G_{21}^0 G_{32}^0$, i.e. a sum of both parallel paths from $w_1$ to $w_3$.

Secondly, consider the system shown in Figure 3b. Suppose that the objective is to identify $G_{21}^0$. If $w_1$ and $w_3$ are both included as predictor inputs, then the dynamics from $w_1$ to $w_2$ are $G_{21}^0$. However, if only $w_1$ is chosen as a predictor input, then the identified transfer function is an estimate of $G_{21}^0/(1 - G_{32}^0 G_{23}^0)$, i.e. the target transfer function includes the sensitivity function of the (unmodelled) loop on the output.

The results of these two examples, can be generalized in the following property and proposition.

**Property 1** *Consider the internal variables $w_i$, $w_j$ and the set of indices of predictor inputs, $\mathcal{D}_j$. Let $\mathcal{D}_j$ satisfy the following conditions:*

*(a) $i \in \mathcal{D}_j$, $j \notin \mathcal{D}_j$,*
*(b) every loop from $w_j$ to $w_j$ passes through a $w_k$, $k \in \mathcal{D}_j$,*
*(c) every path from $w_i$ to $w_j$, excluding the path $G_{ji}^0$, passes through a $w_k$, $k \in \mathcal{D}_j$.* □

**Proposition 5 (Dankers et al. (2013b))** *Consider a dynamic network as defined in Section 2.1 that satisfies Assumption 1, and a set of predictor inputs $\mathcal{D}_j$. Denote $\mathcal{Z}_j$ as the index set of all variables $k$, such that $k \notin \{\mathcal{D}_j, j\}$. Then $w_j(t)$ can be uniquely written as*

$$w_j(t) = \sum_{k \in \mathcal{D}_j} \breve{G}_{jk}^0(q) w_k(t) + \sum_{k \in \mathcal{Z}_j} \breve{F}_{jk}^0(q) \begin{bmatrix} r_k(t) \\ v_k(t) \end{bmatrix} + \\ + r_j(t) + v_j(t) \tag{18}$$

*and $\breve{G}_{ji}^0(q) = G_{ji}^0(q)$ if $\mathcal{D}_j$ satisfies Property 1.* □

Note that Proposition 5 does not say anything about the actual transfer functions between the predictor inputs $w_k$ and $w_j$, $k \neq i$. Thus, if the set of chosen predictor inputs has Property 1 then it is possible to obtain an estimate of $G_{ji}^0$. This is formalized in the following proposition which is a generalization of Theorem 2.

**Theorem 6** *Consider the situation of Theorem 2, except for the case that the set of predictor inputs $\mathcal{D}_j$ is not predefined anymore.*
*If $\mathcal{D}_j$ satisfies Property 1, then the implication*

$$\left\{ R_{\varepsilon z}(\tau) = 0, \text{ for } \tau = 0, \ldots n_s \right\} \Rightarrow \left\{ G_{ji}(q, \theta) = G_{ji}^0(q) \right\}$$

*holds for any finite $n_s \geq \lceil (n_a + dn_b)/length(z(t)) \rceil$ if the following conditions are satisfied:*

*(a) Conditions (a) - (c) of Theorem 2 hold, (thus there is no path from $w_j$ to any of the instrumental variables.)*

(b) $\mathcal{X}_j \subset \mathcal{D}_j$, *i.e. only external variables that directly excite predictor inputs are selected as instrumental variables.*
(c) *There is no noise term $v_m$ that has both of the following two properties: 1) there is a path from $v_m$ to $w_j$ that passes only through nodes $w_k$, $k \in \mathcal{Z}_j$, and 2) there is a path from $v_m$ to an instrumental variable.*
(d) *All process noise variables are uncorrelated.*
(e) *The parameterization is flexible enough, i.e. there exists a $\theta$ such that $R_{\varepsilon z}(\tau, \theta) = 0$ for $\tau = 0, \cdots n_s$.* □

As was the case in Theorem 2 the fact that there cannot be a path from $w_j$ to any of the instrumental variables poses a severe restriction on the candidate instrumental variables. However, this restriction can be removed while still keeping the flexible choice of predictor inputs, as shown in the following proposition which is a generalization of Theorem 4.

**Theorem 7** *Consider the situation of Theorem 4, except for the case that the set of predictor inputs $\mathcal{D}_j$ is not predefined anymore.*
*If $\mathcal{D}_j$ satisfies Property 1, then the implication*

$$\left\{ R_{\varepsilon z}(\tau) = 0, \text{ for } \tau = 0, \ldots, n_s \right\} \\ \Rightarrow \left\{ G_{ji}(q, \theta) = G_{ji}^0(q) \right\} \tag{19}$$

*holds if the following conditions are satisfied:*

*(a) Conditions (a) - (c) of Theorem 4 hold.*
*(b) $\mathcal{X}_j \subset \mathcal{D}_j$, i.e. only external variables that directly excite predictor inputs are selected as instrumental variables.*
*(c) Consider every $v_m$ that has a path to $w_j$ that passes only through nodes $w_k$, $k \in \mathcal{Z}_j$. For every such $v_m$, it must hold that every path from $v_m$ to $w_\ell$, $\ell \in \mathcal{I}_j$, must pass through $w_j$ before reaching $w_\ell$.*
*(d) All process noise variables are uncorrelated.*
*(e) The parameterization is flexible enough, i.e. there exists a $\theta$ such that $R_{\varepsilon z}(\tau, \theta) = 0$.* □

The following example illustrates the flexibility that is allowed in choosing the set $\mathcal{D}_j$ (instead of forcing the user to use only the variables with direct paths to $w_j$ as predictor inputs, i.e. $\tilde{w}_k$, $k \in \mathcal{N}_j$).

**Example 5** *Consider the network shown in Fig. 4. Suppose that the objective is to obtain an estimate of $G_{32}^0$. First we must choose which internal variables to include as predictor inputs, i.e. we must choose $\mathcal{D}_3$ such that it has Property 1. By Condition (a) of Property 1 $\tilde{w}_2$ must be included as a predictor input. Next, we must check all loops from $w_3$ to $w_3$. All such loops pass through $w_2$, which is already chosen as a predictor input, so Condition (b) of Property 1 is satisfied. Next, we check all paths*
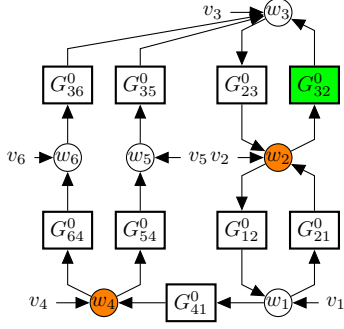
Fig. 4. Network analyzed in Example 5. For clarity the sensor noise is not shown, although it is still assumed to be present.

*from $w_2$ to $w_3$:*

$$w_2 \to w_1 \to w_4 \to w_5 \to w_3 \qquad w_2 \to w_1 \to w_4 \to w_6 \to w_3$$

*for instance. It can be seen that all paths from $w_2$ to $w_3$ (not including $G_{32}^0$) pass through $w_4$. Thus Condition (c) is satisfied if we include $\tilde{w}_4$ as a predictor input. Note that this is not the only choice of $\mathcal{D}_3$ that has Property 1.*

*For this choice of $\mathcal{D}_j$, the candidate instrumental variables are $\{\tilde{w}_1, \tilde{w}_5, \tilde{w}_6\}$. For all these candidates there is a path from $w_j$ to the candidate. Thus, Theorem 6 does not apply and we have to defer to Theorem 7. If measurements of all the candidate instrumental variables are available, we could choose to use them all, i.e. $\mathcal{I}_j = \{1, 5, 6\}$. However, for this choice of instrumental variables Condition (c) is not satisfied, since there is a path from $v_6$ (where $6 \in \mathcal{Z}_3$), that does not pass through the output $w_3$ before reaching the instrumental variable $w_6$. Thus, we choose the set $\mathcal{I}_j = \{1\}$. In this case Condition (c) is satisfied. If the remaining conditions of Theorem 7 are satisfied, then the implication (19) holds.* $\qquad\square$

## 6   Implementation of Algorithm 2

An attractive feature of the classical IV methods is that the estimates can be obtained by solving a linear regression problem. When making the move to a BJ model structure, as in the method proposed in Section 4, this property is lost. In this section an implementation of the method presented in Section 4 is presented.

We show that standard tools for identifying Box-Jenkins models can be used to obtain an estimate of the solution to (17). Recall from Section 4 that we are interested in finding $\theta$ such that

$$R_{\varepsilon z}(\tau, \theta) = 0 \text{ for } \tau = 0, \ldots, n_z.$$

This is equivalent to finding $\theta$ such that

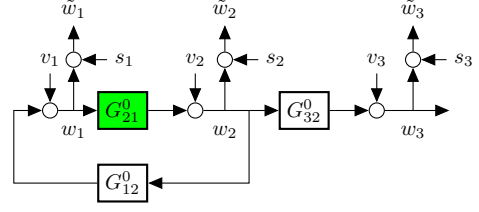$$\sum_{\tau=0}^{n_z} R_{\varepsilon z}^2(\tau, \theta) = 0. \qquad (20)$$



Fig. 5. Data generating system considered in Example 6

Since (20) is nonnegative for all $\theta$, finding $\theta$ such that (20) holds is equivalent to finding $\theta$ such that

$$\hat{\theta} = \arg\min_{\theta} \sum_{\tau=0}^{n_z} R_{\varepsilon z}^2(\tau, \theta). \qquad (21)$$

Note that (21) is a standard sum of squared errors objective function. Now, consider the expression for $R_{\varepsilon z}(\tau)$:

$$R_{\varepsilon z}(\tau) = \bar{\mathbb{E}}\Big[\Big(H_j^{-1}(\theta)\Big(w_j(t) - \sum_{k \in \mathcal{D}_j} G_{jk}(\theta)w_k(t)\Big)\Big)z(t-\tau)\Big]$$

$$= H_j^{-1}(q, \theta)\Big(R_{w_j z}(\tau) - \sum_{k \in \mathcal{D}_j} G_{jk}(q, \theta)R_{w_k z}(\tau)\Big). \quad (22)$$

The point is that (22) has the same form as the prediction error using a Box-Jenkins model structure (see (4)), where the "output" is $\hat{R}_{w_j z}(\tau)$ and the predictor "inputs" are $\hat{R}_{w_k z}(\tau)$, $k \in \mathcal{D}_j$. In practice $R_{w_j z}(\tau)$ and $R_{w_k z}(\tau)$ cannot be exactly computed. However, $R_{w_j z}(\tau)$ for instance can be approximated as:

$$\hat{R}_{w_j z}(\tau) = \frac{1}{N - \tau}\sum_{t=\tau}^{N} w_j(t)z(t - \tau).$$

Thus, we can compute $\hat{R}_{w_j z}(\tau)$ and $\hat{R}_{w_k z}(\tau)$ for $\tau = 0, \ldots, n_z$, $k \in \mathcal{D}_j$ resulting in a data set. Now standard identification tools (such as the `bj` function in the MAT-LAB identification toolbox) can be used to find $\theta$.

**Example 6** *Consider the system shown in Fig. 5. The objective is to obtain an estimate of $G_{21}^0$ using $\tilde{w}_1$, $\tilde{w}_2$ and $\tilde{w}_3$. Thus, the output is $w_2$, and the predictor input is $\tilde{w}_1$. This leaves $\tilde{w}_3$ as the only choice for instrumental variable. In this case Algorithm 1 does not apply since there is a path from $w_2$ to the instrumental variable $w_3$. Thus, we use Algorithm 2. All the noise variables $v_k$ and $s_k$, $k = 1, 2, 3$ are simulated as sequences of low-pass filtered white noise. 5000 data points are simulated. Results are shown in Fig. 6. The blue lines denote estimates that are obtained by ignoring the presence of sensor noise and applying the Direct Method of Van den Hof et al. (2013). Clearly these estimates are biased. The red lines denote estimates obtained using the implementation of Algorithm 2 presented in this section with $n_z = 1000$. The estimates appear consistent, as expected.* $\qquad\square$
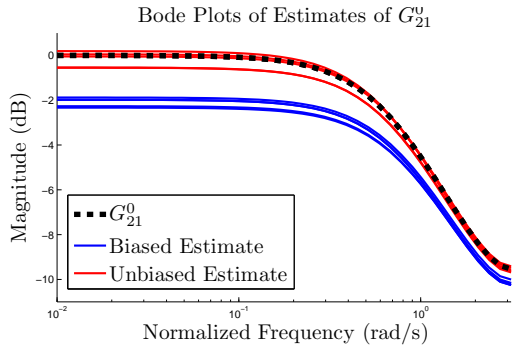
Fig. 6. Frequency responses related to the system of Fig. 5: $G_{21}^0$ (dashed), five realizations of estimated frequency responses using the Direct Method (blue) and the generalized IV Method (red).

## 7 Model Validation

Once a model is obtained, it is possible to express how confident one is that the obtained model is in fact the model that generated the data. Presumably, $\hat{R}_{\varepsilon z}(\tau, \hat{\theta}_N)$ is small for all $\tau \geq 0$. However, how can one be sure that it is small enough to be considered "very near to zero"? If the variance of $\hat{R}_{\varepsilon z}(\tau, \hat{\theta}_N)$ is known, then it is possible to say that $\hat{R}_{\varepsilon z}(\tau, \hat{\theta}_N)$ is zero with probability $p$. Then, by the implications (11) and (16), it follows that it is possible to address the quality of the estimate $G_{jk}(q, \hat{\theta})$.

The steps shown in Söderström & Stoica (1990, 1989); Ljung (1999) can be closely followed in order to obtain the variance of $\hat{R}_{\varepsilon z}^N(\tau, \hat{\theta}_N)$. The result is that

$$\sqrt{N} \hat{R}_{\varepsilon z}^N(\tau, \hat{\theta}_N) \in \mathrm{As}\mathcal{N}(0, P)$$

where $\mathrm{As}\mathcal{N}(0, P)$ means that as $N \to \infty$ the distribution of $\sqrt{N} \hat{R}_{\varepsilon z}^N(\tau, \hat{\theta}_N)$ tends to a normal distribution with zero mean and variance $P$, where (Ljung, 1999):

$$P = \sum_{\tau=-\infty}^{\infty} R_\varepsilon(\tau) R_z(\tau).$$

Let $n_\alpha$ denote the $\alpha$ level of the $\mathcal{N}(0, 1)$ distribution. Then it is possible to check if (Ljung, 1999)

$$\left| \hat{R}_{\varepsilon z}^N(\tau, \hat{\theta}) \right| \leq \sqrt{\frac{P}{N}} n_\alpha.$$

If the inequality holds, then the obtained model is the correct model with probability $\alpha$.

**Example 7** *Consider the same situation as in Example 6. Consider a confidence level of 95%. The results are shown in Fig. 7. From the figure, one can conclude with 95% confidence that there is no evidence in the data that the model is wrong.* □
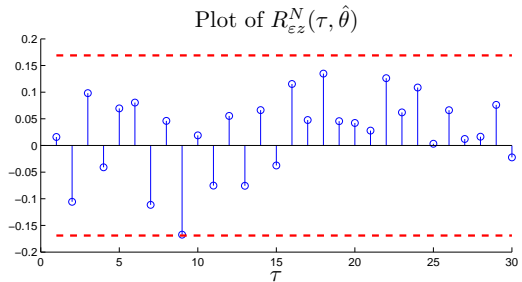


Fig. 7. Results of the validation test presented in Section 7 for the setup considered in Examples 6 and 7. The red dashed lines represent the 95% confidence intervals.

## 8 Conclusion

In this paper a novel IV-based method is presented to obtain consistent estimates of a module $G_{ji}^0$ embedded in a dynamic network using measurements of internal variables that are corrupted by sensor noise. Any variable (external or internal) that is not a predictor input or $w_j$ is a candidate instrumental variable. If instrumental variables are chosen that do not have a connected path from the module output $w_j$, the resulting algorithm is of a linear regression format. If there is such a connection then a nonlinear optimization algorithm results. In both cases consistency results are formulated that are based on excitation conditions on the appropriate input and IV signals, while allowing all measured signals to be measured with noise. The resulting EIV problem, that is known to be problematic in classical (open-loop and closed-loop) identification, becomes tractable now because of the additionally measured variables that are available in a dynamic network. Additionally predictor input selection results are formulated that allow a check on which choice of predictor input signals leads to consistent module estimates. This allows the experimenter to flexibly choose which particular variables to measure in a particular network, allowing choices that are based on sensor costs, least number of variables, etcetera.

## A Proof of Lemma 1

Consider first the following lemma. For a proof see Mason's Rules (Mason, 1953), or Van den Hof et al. (2013).

**Lemma 8** *Consider a dynamic network as defined in Section 2.1 that satisfies Assumption 1. Let $\mathcal{G}_{mn}^0$ be the $(m, n)$th entry of $(I - G_0)^{-1}$. If there are no paths from $n$ to $m$ then $\mathcal{G}_{mn}^0$ is zero.* □

Now follows the proof of Lemma 1.

**PROOF.** The proof proceeds by considering $z_\ell = w_\ell$ (the proof for $z_\ell = r_\ell$ is analogous). First both $w_\ell$ and $w_k$ are expressed in terms of process noise variables. Then Lemma 8 is used to prove the result. Using the notation

of Lemma 8, $w_\ell$ and $w_m$ can be expressed in terms of only process noise terms:

$$w_\ell(t) = \sum_{n=1}^{L} \mathcal{G}_{\ell n}^0(q) v_n(t) \text{ and } w_k(t) = \sum_{n=1}^{L} \mathcal{G}_{kn}^0(q) v_n(t).$$

Consequently the cross power spectral density $\Phi_{w_\ell w_k}$ is

$$\Phi_{w_\ell w_k}(\omega) = \sum_{\substack{n=1 \\ n \neq \ell, k}}^{L} \mathcal{G}_{\ell n}^0(e^{j\omega}) \Phi_{v_n}(\omega) \mathcal{G}_{kn}^0(e^{-j\omega}) +$$
$$\mathcal{G}_{\ell k}^0(e^{j\omega}) \Phi_{v_k}(\omega) \mathcal{G}_{kk}^0(e^{-j\omega}) + \mathcal{G}_{\ell \ell}^0(e^{j\omega}) \Phi_{v_\ell}(\omega) \mathcal{G}_{k\ell}^0(e^{-j\omega})$$

Suppose that none of the Conditions of Lemma 1 hold. By Lemma 8 and Condition (a), $\mathcal{G}_{k\ell}^0$ is zero. Thus the third term of $\Phi_{w_\ell w_k}(z)$ is zero. Similarly, by Condition (b) the second term is zero. By Condition (c) for each $n \in \{1, \ldots, L\} \setminus \{k, \ell\}$ either $\mathcal{G}_{\ell n}^0$ or $\mathcal{G}_{kn}^0$ is zero. Thus the first term of $\Phi_{w_\ell w_m}(z)$ is zero. Consequently, if none of the conditions hold, $w_k$ and $w_\ell$ are uncorrelated. $\square$

## B Proof of Theorem 2

**PROOF.** The proof proceeds by first deriving a simplified expression for $R_{\varepsilon z}(\tau)$ and then showing that this expression equals 0 for $\tau = 0, \ldots, n_s$ if and only if $\theta = \theta^0$. Using (10) $R_{\varepsilon z}(\tau)$ can be expressed as

$$\bar{\mathbb{E}}[\varepsilon(t)z(t-\tau)] = \bar{\mathbb{E}}\left[\left(\tilde{w}_j(t) - \theta^T \tilde{\phi}(t)\right) z(t-\tau)\right]$$

Both the predictor inputs and the instrumental variable have a component that is due to the sensor noise. However, By Condition (c) both these components can be removed from the expression of $R_{\varepsilon z}(\tau)$:

$$R_{\varepsilon z}(\tau) = \bar{\mathbb{E}}\Big[\big(w_j(t) + s_j(t) - \theta^T \phi(t) - \theta^T \phi_s(t)\big)$$
$$\cdot \big(z(t-\tau) + z_s(t-\tau)\big)\Big]$$
$$= \bar{\mathbb{E}}\Big[\big(w_j(t) - \theta^T \phi(t)\big) z(t-\tau)\Big]$$

where
$$\phi_s^T(t) = \Big[s_{d_1}(t) \cdots s_{d_1}(t-n_b) \cdots s_j(t-1) \cdots s_j(t-n_a)\Big]$$

and, similarly, $z_s(t)$ is a vector of all the measurement noise terms associated with the instrumental variables. From (9) $w_j(t)$ can be expressed as:

$$w_j(t) = \theta_0^T \phi(t) + A_j^0(q) v_j(t)$$

where $\theta^0 = [\breve{b}_{jk_1}^0 \cdots \breve{b}_{jk_d}^0 \ a_j^0]$ where $\breve{b}_{jk_i}^0$ is a vector of the coefficients of $\breve{B}_{jk_i}^0$, $k_i \in \mathcal{N}_j$ and $a_j^0$ is a vector of the

coefficients of $A_j^0$. Using this expression for $w_j$ in $R_{\varepsilon z}(\tau)$:

$$R_{\varepsilon z}(\tau) = \bar{\mathbb{E}}\Big[\big(\theta_0^T \phi(t) + A_j^0(q) v_j(t) - \theta^T \phi(t)\big) z(t-\tau)\Big]$$
$$= \bar{\mathbb{E}}\Big[\big(\Delta\theta\phi(t) + A_j^0(q) v_j(t)\big) z(t-\tau)\Big] \qquad (B.1)$$

where $\Delta\theta = \theta_0 - \theta$. Condition (a) states that there is no path from any predictor input to any variable chosen as an instrument. This implies that each $w_\ell$, $\ell \in \mathcal{I}_j$ is not a function of $v_j$. This statement can be proved using Lemma 8 as follows. First, using the notation of Lemma 8, express $w_\ell$ in terms of $v$:

$$w_\ell = \sum_{k=1}^{L} \mathcal{G}_{\ell k}^0 v_k. \qquad (B.2)$$

Since there is no path from $w_j$ to $w_\ell$, by Lemma 8 $\mathcal{G}_{\ell j}^0$ is zero. Thus, $w_\ell$ is not a function of $v_j$. Consequently, by Condition (d) $w_\ell$ and $v_j$ are uncorrelated. This leads to the following simplification of (B.1):

$$R_{\varepsilon z}(\tau) = \bar{\mathbb{E}}\Big[\Delta\theta\phi(t) z(t-\tau)\Big] \qquad (B.3)$$

This is the final expression for $R_{\varepsilon z}(\tau)$. It follows immediately from (B.3) that if $\theta = \theta_0$ then $R_{\varepsilon z}(\tau) = 0$ for all finite $\tau \geq 0$. It remains to be shown that if $R_{\varepsilon z}(\tau) = 0$ for $\tau = 0, \ldots, n_s$ then $\theta = \theta_0$. Consider the set of equations:

$$[R_{\varepsilon z}(0) \ R_{\varepsilon z}(1) \ \cdots \ R_{\varepsilon z}(n_s)] = 0$$

Then, using (B.3), it follows that

$$\Delta\theta[R_{\phi z}(0) \ \cdots \ R_{\phi z}(n_m) \ \cdots \ R_{\phi z}(n_s)] = 0. \qquad (B.4)$$

The matrix $[R_{\phi z}(0) \ R_{\phi z}(1) \ \cdots \ R_{\phi z}(n_m)]$ is either square or has more columns than rows. By Condition (b) it is full row rank. Consequently, the only solution to the equation is $\Delta\theta = 0$. This proves the result. $\square$

## C Proof of Theorem 4

The following Lemma is used in the proof of Theorem 4.

**Lemma 9** *Consider vector signals $w(t) \in \mathbb{R}^m$ and $z(t) \in \mathbb{R}^n$ and let the scalar signal $\varepsilon(t)$ be given by $\varepsilon(t) = M(q)w(t)$ with $M(q) = a^{-1}(q)B(q)$, with*

$$a(q) = 1 + a_1 q^{-1} + \cdots a_{n_a} q^{-n_a}, \text{ and}$$
$$B(q) = B_0 + B_1 q^{-1} + \cdots B_{n_b} q^{-n_b}.$$

*Denote the vector: $\phi^T(t) = [w^T(t) \ \cdots \ w^T(t - n_b)]$. If the matrix*

$$\bar{\mathbb{E}}\Big[\phi(t)[z^T(t - n_a) \ \cdots \ z^T(t - n_a - n_m)]\Big] \qquad (C.1)$$

*has full row rank for sufficiently large $n_m$, then the implication*

$$\{R_{\varepsilon z}(\tau) = 0 \;\; for\; \tau = 0, \dots, n_s \implies M(q) \equiv 0\}$$

*holds for any finite $n_s \geq n_a + n_m$.*  □

**PROOF.** Using the expression for $M(q)$ it follows that $a(q)\varepsilon(t) = B(q)w(t)$. Postmultiplying this expression with $z^T(t-\tau)$ and taking expectations, then leads to

$$a(q)R_{\varepsilon z}(\tau) = [B_0\; B_1\; \cdots B_{n_b}] \cdot \bar{\mathbb{E}}\phi(t)z^T(t-\tau).$$

Combining this equation for $\tau = n_a, n_a + 1, \cdots n_a + n_m$ and positioning them next to each other then leads to

$$[1\; a_1 \cdots a_{n_a}]\begin{bmatrix} R_{\varepsilon z}(n_a) & \cdots & R_{\varepsilon z}(n_a + n_m) \\ \vdots & & \vdots \\ R_{\varepsilon z}(0) & \cdots & R_{\varepsilon z}(n_m) \end{bmatrix} =$$
$$= [B_0\; B_1\; \cdots\; B_{n_b}]\bar{\mathbb{E}}\Big[\phi(t)[z^T(t-n_a)\cdots z^T(t-n_a-n_m)]\Big]$$

where $n_m$ is chosen large enough so that the matrix on the right hand side can have full row rank.
If $R_{\varepsilon z}(\tau) = 0$ for $\tau = 0, \dots, n_a + n_m$, then the matrix on the left hand side is 0, and under the condition that the matrix on the right hand side has full row rank, this implies that $[B_0\; B_1 \cdots B_{n_b}] = 0$ which implies that $M(q) \equiv 0$.  □

Now follows the proof of Theorem 4.

**PROOF.** First an expression for $R_{\varepsilon z}(\tau)$ is derived. Then it is shown that under the specified conditions, $R_{\varepsilon z}(\tau) = 0$ for $\tau = 0, \dots, n_s$ iff $G_{jk}(\theta) = G_{jk}^0$ and $H_j(\theta) = H_j^0$. Consider first an expression for the prediction error. Substitute the expressions for $\tilde{w}_j$ and $\tilde{w}_k$ into (4):

$$\varepsilon_j(\theta) = H_j^{-1}(\theta)\Big(w_j + s_j - \sum_{k \in \mathcal{N}_j} G_{jk}(\theta)(w_k + s_k)\Big) \quad \text{(C.2)}$$

The vector of instrumental variables is chosen to contain only node measurements and no external variables. This does not affect the reasoning, it simply reduces the notational burden:

$$z^T(t) = \big[\tilde{w}_{\ell_1}(t - d_1)\; \cdots\; \tilde{w}_{\ell_n}(t - d_n)\big] \quad \text{(C.3)}$$

where $\tilde{w}_{\ell_k}(t - d_k) = w_{\ell_k}(t - d_k) + s_{\ell_k}(t - d_k)$. No measurement chosen as an IV can be a predictor input ($\mathcal{N}_j \cap \mathcal{I}_j = \emptyset$ by the statement of the theorem). Thus,

no $s_\ell$ that appears in the instrumental variable vector $z$ (C.3), will appear in the expression for $\varepsilon_j$, (C.2). By Condition (c) each $s_k$, $k \in \mathcal{D}_j$ is uncorrelated to all $s_\ell$, $\ell \in \mathcal{I}_j$. Thus, combining (C.2) and (C.3) results in the following expression for $R_{\varepsilon z}$:

$$\bar{\mathbb{E}}[\varepsilon_j(t,\theta) \cdot z(t - \tau)] = \bar{\mathbb{E}}\Big[\Big(H_j^{-1}(\theta)\Big(w_j - \sum_{k \in \mathcal{N}_j} G_{jk}(\theta)w_k\Big)\Big)$$
$$\cdot \big[w_{\ell_1}(t-\tau-d_1)\; \cdots\; w_{\ell_n}(t-\tau-d_n)\big]\Big]. \quad \text{(C.4)}$$

Using the fact that $w_j = \sum_{k \in \mathcal{N}_j} G_{jk}^0 w_k + v_j$, and using the notation $\Delta G_{jk}(\theta) = G_{jk}^0 - G_{jk}(\theta)$ and $\Delta H_j(\theta) = H_j^{-1}(\theta) - H_j^{0^{-1}}$, we can rewrite the first component on the right hand side of (C.4) as:

$$H_j^{-1}(\theta)\Big(w_j - \sum_{k \in \mathcal{N}_j} G_{jk}(\theta)w_k\Big) =$$
$$\sum_{k \in \mathcal{N}_j} \frac{\Delta G_{jk}(q,\theta)}{H_j(\theta)}w_k(t) + \Delta H_j(q,\theta)v_j(t) + e_j(t).$$

Because each (delayed) instrumental variable $w_{\ell_k}(t - d_k)$ is only a function of delayed versions of $e_j$ (i.e. $R_{e_j w_{\ell_k}}(\tau) = 0$ for $\tau \geq d_k$ as stated in the body of the theorem), and because, by Condition (d) $v_j$ is uncorrelated to each of the noise terms that make up the IV's, the following simplification of (C.4) results:

$$R_{\varepsilon z}(\tau) = \bar{\mathbb{E}}\Big[\Big(\sum_{k \in \mathcal{N}_j} \frac{\Delta G_{jk}(q,\theta)}{H_j(q,\theta)}w_k(t) + \Delta H_j(q,\theta)v_j(t)\Big)$$
$$\cdot \big[w_{\ell_1}(t-\tau-d_1)\; \cdots\; w_{\ell_n}(t-\tau-d_n)\big]\Big] \quad \text{(C.5)}$$

which holds for all $\tau \geq 0$. Using a vector notation (C.5) can be expressed as:

$$R_{\varepsilon z}(\tau) = \bar{\mathbb{E}}\Big[\Delta X(q,\theta)^T$$
$$\cdot \begin{bmatrix} w_{k_1}(t) \\ \vdots \\ w_{k_d}(t) \\ v_j(t) \end{bmatrix}\big[w_{\ell_1}(t-\tau-d_1)\cdots w_{\ell_n}(t-\tau-d_n)\big]\Big]$$

where
$$\Delta X(q,\theta)^T = \Big[\frac{\Delta G_{jk_1}(q,\theta)}{H_j(q,\theta)}\; \cdots\; \frac{\Delta G_{jk_d}(q,\theta)}{H_j(q,\theta)}\; \Delta H_j(q,\theta)\Big]$$

and $\{k_1, \dots, k_d\} = \mathcal{D}_j$. The variable $v_j$ can be expressed in terms of internal variables as:

$$v_j = w_j - \sum_{k \in \mathcal{N}_j} G_{jk}^0(q)w_k$$

and so

$$\begin{bmatrix} w_{k_1}(t) \\ \vdots \\ w_{k_d}(t) \\ v_j(t) \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ -G^0_{jk_1}(q) & \cdots & -G^0_{jk_d}(q) & 1 \end{bmatrix} \begin{bmatrix} w_{k_1}(t) \\ \vdots \\ w_{k_d}(t) \\ w_j(t) \end{bmatrix}. \quad \text{(C.6)}$$

Denote the matrix in (C.6) as $J^0(q)$. Using this notation,

$$R_{\varepsilon z}(\tau) = \bar{\mathbb{E}}\Big[ \Delta X(q,\theta)^T J^0(q) w(t) \\ \cdot [w_{\ell_1}(t-\tau-d_1) \cdots w_{\ell_n}(t-\tau-d_n)] \Big] \quad \text{(C.7)}$$

where $w(t) = [w_{k_1}(t) \cdots w_{k_d}(t)\, w_j(t)]^T$. Note that (C.7) is valid for all finite $\tau \geq 0$.

First consider the 'if' statement. It must be shown that if $G_{jk}(q,\theta) = G^0_{jk}$, for all $k \in \mathcal{N}_j$ and $H_j(q,\theta) = H^0_j$, then $R_{\varepsilon z}(\tau) = 0$ for all finite $\tau \geq 0$. Let $\theta_0$ denote this particular parameter vector (such a parameter vector is guaranteed to exist by Condition (e)). Clearly, $\Delta G_{jk}(\theta_0) = 0$ and $\Delta H_j(\theta_0) = 0$. Thus, from (C.5), $R_{\varepsilon z}(\tau,\theta_0) = 0$, for $\tau = 0, \ldots, n_s$.

Now consider the 'only if' statement. It must be shown that if $R_{\varepsilon z}(\tau) = 0$, for $\tau = 0, \ldots, n_s$ then $G_{jk}(\theta) = G^0_{jk}$, for all $k \in \mathcal{N}_j$ and $H_j(\theta) = H^0_j$. From (C.7) the filtered $w(t)$-signal can be expressed as $M(q,\theta)w(t)$ where

$$M(q,\theta) = \Delta X(q,\theta)^T J^0 = x_d(q,\theta)^{-1} P(q,\theta)$$

with $x_d$ and $P$ polynomial. By Condition (a) of Theorem 4 the matrix $\bar{\mathbb{E}}[\phi(t)[z(t-n_x) \cdots z(t-n_x-n_z)]$ is full row rank where $\phi(t)$ is composed of delayed versions of $w(t)$. If the value of $n_x$ is chosen correctly and the result of Lemma 9 can be applied, then it follows that $\Delta X(q,\theta)J^0 = 0$. By evaluating this latter equation, incorporating Condition (e), it follows then immediately that $G_{jk}(q,\theta) = G^0_{jk}(q)$, and $H_j(q,\theta) = H^0_j(q)$. The only remaining item to address now is to determine the polynomial degrees of $x_d$ and $P$ for the appropriate application of Lemma 9.

To this end, let $G^0_{jk} = \frac{B^0_{jk}}{F^0_{jk}}$ and $H^0_j = \frac{C^0}{D^0}$, where $B^0_{jk}$ and $F^0_{jk}$ are co-prime polynomials (in $q^{-1}$) for each $k$, $C^0$ and $D^0$ are co-prime polynomials. Then $\Delta X$ can be expressed as:

$$\Delta X(q,\theta) = \frac{1}{x_d(q,\theta)}\Big[ p_{k_1}(q,\theta) \cdots p_{k_d}(q,\theta)\, p_j(q,\theta) \Big],$$

where

$$x_d(q,\theta) = F^0_{jk_1} F_{jk_1}(\theta) \cdots F^0_{jk_d} F_{jk_d}(\theta) C^0 C(\theta),$$
$$p_{k_n}(q,\theta) = \Big( B^0_{jk_n} F_{jk_n}(\theta) - B_{jk_n}(\theta) F^0_{jk_n} \Big)$$
$$\prod_{k_m \in \mathcal{D}_j \setminus \{k_n\}} F^0_{jk_m} F_{jk_m}(\theta) C(\theta) D^0, \quad \text{(C.8)}$$
$$p_j(q,\theta) = \Big( D(\theta)C^0 - D^0 C(\theta) \Big) \prod_{k_m \in \mathcal{D}_j} F^0_{jk_m} F_{jk_m}(\theta).$$

Multiplying $\Delta X$ by $J^0$ results in:

$$\Delta X(q,\theta)J^0 = \frac{1}{x_d(q,\theta)}\Big[ p_{k_1}(q,\theta) - j_{k_1}(q,\theta) \\ \cdots\ p_{k_d}(q,\theta) - j_{k_d}(q,\theta)\ p_j(q,\theta) \Big],$$

where $j_{k_n} = \frac{p_j B^0_{jk_n}}{F^0_{jk_n}}$ (which is a polynomial). From (C.8) it follows that $p_{k_n}$ has degree $\sum_{k \in \mathcal{D}_j} 2n_{f_k} + 2n_c$, which is the same as the degree of $j_{k_n}$, $p_j$ and $x_d$. As a result Lemma 9 can be applied with $n_a = n_b = n_x = \sum_{k \in \mathcal{D}_j} 2n_{f_k} + 2n_c$, and this is used in Condition (a), which proves the theorem. $\square$

## D  Proof of Theorem 6

**PROOF.** Recall the expression for $w_j$ from Proposition 5 and denote

$$\breve{v}_j := \sum_{k \in \mathcal{Z}_j} \breve{F}^0_{jk} v_k + v_j. \quad \text{(D.1)}$$

The proof will proceed by showing that Conditions (c) and (d) of Theorem 6 result in $\breve{v}_j$ being uncorrelated to all $v_m$ with a path to any one of the instrumental variables. Then all the conditions of Theorem 2 hold, and so the implication holds.

The expression for $\breve{F}^0_{j\mathcal{Z}}$ is $G_{j\mathcal{Z}}(I - G^0_{\mathcal{Z}\mathcal{Z}})^{-1}$. Applying Lemma 8 to $(I - G^0_{\mathcal{Z}\mathcal{Z}})^{-1}$ it follows that any $\breve{F}_{jk}$ is zero if there is no path from $v_k$, $k \in \mathcal{Z}_j$ to $w_j$ that passes only through nodes in $\mathcal{Z}_j$. Thus, by Condition (c) and (D.1), it follows that $v_j$ is not a function of any $v_m$ that has a path to $w_j$ that passes only through nodes in $\mathcal{Z}_j$.

Secondly, consider the the instrumental variables. Express the instrumental variables in terms of process noise and external variables:

$$w_\ell = \sum_{k=1}^{L} \mathcal{G}^0_{\ell k}(v_k + r_k) \quad \text{(D.2)}$$

By Lemma 8 and Condition (c) $w_\ell$ is not a function of $v_m$. In addition, by the same reasoning as in the proof of Theorem 2 $w_\ell$ is not a function of $v_j$. Thus it has been

shown that no instrumental variable is a function of any $v_m$ that has a path that passes only through nodes in $\mathcal{Z}_j$ to $w_j$. It follows by Condition (d) $\tilde{v}_j$ is uncorrelated to the instrumental variables, $w_\ell$, $\ell \in \tilde{\mathcal{I}}_j$.

Thus, by Theorem 2 and the expression for $w_j$ from Proposition 5 the implication of Theorem 6 holds.

## E Proof of Theorem 7

**PROOF.** The proof will proceed by showing that Conditions (c) and (d) result in Condition (d) of Theorem 4 being satisfied. Then, all conditions of Theorem 4 are satisfied, and so the implication holds. Recall the expression for $\breve{v}_j$ (D.1). From the same reasoning as in the proof of Theorem 6 it follows that $\breve{F}_{jk}^0$ is non-zero if there is a path from a $v_k$, $k \in \mathcal{Z}_j$ to $w_j$ that passes only through nodes in $\mathcal{Z}_j$. Now we will show that Conditions (c) and (d) ensure that $\tilde{v}_j$ is uncorrelated to all other process noise terms with a path to any instrumental variable. Each instrumental variable can be expressed as:

$$w_\ell = \sum_{k=1}^{L} \mathcal{G}_{\ell k}^0 v_k. \tag{E.1}$$

The instrumental variables can be expressed in terms of $\breve{v}_j$. Denote the set of variables $v_k$ that make up $\breve{v}_j$ as $\mathcal{A}_j$. Because there is no path from any $v_k$, $k \in \mathcal{A}_j$ to $w_\ell$ that does not pass through $w_j$, it follows that $w_\ell$ can be expressed as

$$w_\ell = \sum_{k \in \mathcal{B}} \mathcal{G}_{\ell k} v_k + \mathcal{G}_{\ell j} \breve{v}_j \tag{E.2}$$

where $\mathcal{B} = \{1, \ldots, L\} \setminus \mathcal{A}_j$. Because all process noise terms are uncorrelated, it follows that $\breve{v}_j$ is uncorrelated to all other noise terms $v_k$ that have a path to $w_\ell$.

Thus we have shown that Condition (d) of Theorem 7 holds. All remaining conditions of Theorem 7 hold, and so the implication holds.

## References

Ali, M., Popov, A., Werner, H., & Abbas, H. (2011). Identification of distributed systems with identical subsystems. In *Proceedings of the 18th IFAC World Congress* (pp. 5633–5638). Milan, Italy.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*, 444–455.

Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives*, *15*, 6985.

Becker, W. E. (2010). Issues of endogeneity and instrumental variables in economic education research *in An Online Handbook for the Use of Contemporary Econometrics in Economic Education Research*. American Economic Association. Accessed: October 1 2014.

Caines, P. E., & Chan, C. W. (1975). Feedback between stationary stochastic processes. *IEEE Transactions on Automatic Control*, *20*, 498–508.

Chuiso, A., & Pillonetto, G. (2012). A Bayesian approach to sparse dynamic network identification. *Automatica*, *48*, 1553–1565.

Dankers, A., Van den Hof, P. M. J., & Bombois, X. (2014a). An instrumental variable method for continuous-time identification in dynamic networks. In *Proceedings of 53th Conference on Decision and Control (CDC)*. Los Angeles, CA, USA. Accepted.

Dankers, A., Van den Hof, P. M. J., Bombois, X., & Heuberger, P. S. C. (2013a). Predictor input selection for two stage identification in dynamic networks. In *Proceedings of the European Control Conference (ECC)* (pp. 1422–1427). Zürich, Switzerland.

Dankers, A., Van den Hof, P. M. J., Bombois, X., & Heuberger, P. S. C. (2014b). Errors in variables identification in dynamic networks by an instrumental variable approach. In *Proceedings of 19th IFAC World Congress* (pp. 2335– 2340). Cape Town, South Africa.

Dankers, A., Van den Hof, P. M. J., & Heuberger, P. S. C. (2013b). Predictor input selection for direct identification in dynamic networks. In *Proceedings of the 52nd IEEE Conference on Decision and Control (CDC)* (pp. 4541–4546). Florence, Italy.

Durbin, J. (1954). Errors in variables. *Revue de l'Institut International de Statistique / Review of the International Statistical Institute*, *22*, 23–32.

Everitt, N., Hjalmarsson, H., & Rojas, C. (2013). A geometric approach to variance analysis of cascaded systems. In *Proceedings of the 52nd IEEE Conference on Decision and Control* (pp. 6496 – 6501). Florence, Italy.

Everitt, N., Hjalmarsson, H., & Rojas, C. (2014). Variance results for parallel cascade serial systems. In *Proceedings of 19th IFAC World Congress* (pp. 2317–2322). Cape Town, South Africa.

Forssell, U., & Ljung, L. (1999). Closed-loop identification revisited. *Automatica*, *35*, 1215–1241.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Applications of the lasso and grouped lasso to the estimation of sparse graphical models. Unpublished.

Gevers, M. R., & Anderson, B. D. O. (1981). Representing of jointly stationary stochastic feedback processes. *International Journal of Control*, *33*, 777–809.

Gilson, M., & Van den Hof, P. M. J. (2005). Instrumental variable methods for closed-loop system identification. *Automatica*, *41*, 241–249.

Granger, C. W. J. (1980). Testing for causality; a personal viewpoint. *Journal for Economic Dynamics and Control*, (pp. 329–352).

Gunes, B., Dankers, A., & Van den Hof, P. M. J. (2014). Variance reduction for identification in dynamic net-

works. In *Proceedings of 19th IFAC World Congress* (pp. 2842–2847). Cape Town, South Africa.

Haber, A., & Verhaegen, M. (2012). Identification of spatially distributed discrete-time state-space models. In *Proceedings of the 16th IFAC Symposium on System Identification* (pp. 410–415). Brussels, Belgium.

Ljung, L. (1999). *System Identification. Theory for the User*. (2nd ed.). Prentice Hall.

Mason, S. J. (1953). Feedback theory - some properties of signal flow graphs. *Proceedings of The IRE*, (pp. 1144–1156).

Massioni, P., & Verhaegen, M. (2008). Subspace identification of circulant systems. *Automatica*, *44*, 2825 – 2833.

Materassi, D., & Innocenti, G. (2010). Topological identification in networks of dynamical systems. *IEEE Transactions on Automatic Control*, *55*, 1860–1871.

Materassi, D., & Salapaka, M. V. (2012). On the problem of reconstructing an unknown topology via locality properties of the Wiener filter. *IEEE Transactions on Automatic Control*, *57*, 1765–1777.

Pintelon, R., & Schoukens, J. (2012). *System Identification, A Frequency Domain Approach*. (2nd ed.). Hoboken, New Jersey, USA: IEEE Press, John Wiley and Sons, Inc.

Sanandaji, B. M., Vincent, T. L., & Wakin, M. B. (2012). A review of sufficient conditions for structure identification in interconnected systems. In *Proceedings of the 16th IFAC Symposium on System Identification* (pp. 1623–1628). Brussels, Belgium.

Schoukens, J., Pintelon, R., Vandersteen, G., & Guillaume, G. (1997). Frequency-domain system identification using non-parametric noise models estimated from a small number of data sets. *Automatica*, *33*, 1073–1086.

Seneviratne, A., & Solo, V. (2012). Topology identification of a sparse dynamic network. In *Proceedings of the 51st IEEE Conference on Decision and Control (CDC)* (pp. 1518–1523). Maui, HI, USA.

Söderström, T. (2007). Errors-in-variables methods in system identification. *Automatica*, *43*, 939–958.

Söderström, T. (2012). System identification for the errors-in-variables problem. *Transactions of the Institute of Measurement and Control*, *34*, 780–792.

Söderström, T., & Hong, M. (2005). Identification of dynamic errors-in-variables systems with periodic data. In *Proceedings of the 16th IFAC World Congress* (pp. 809–814). Prague, Czech Republic.

Söderström, T., & Stoica, P. (1983). *Instrumental Variable Methods for System Identification*. LNCIS. Berlin: Springer.

Söderström, T., & Stoica, P. (1989). *System Indentification*. Hemel Hempstead, U.K.: Prentice-Hall Internatinal.

Söderström, T., & Stoica, P. (1990). On covariance function tests used in system identification. *Automatica*, *26*, 125–133.

Söderström, T., Stoica, P., & Trulsson, E. (1988). Instrumental variable methods for closed loop systems.

In *Proceedings of the 10th IFAC World Congress* (pp. 363–368). Munich, Germany.

Söderström, T., Wang, L., Pintelon, R., & Schoukens, J. (2013). Can errors-in-variables systems be identified from closed-loop experiments. *Automatica*, *49*, 681–684.

Thil, S., & Gilson, M. (2011). Survey of analytical IV estimates for errors-in-variables model identification. In *Proceeings of the 18th IFAC World Congress* (pp. 13098–13103). Milan, Italy.

Van den Hof, P. M. J., Dankers, A., Heuberger, P. S. C., & Bombois, X. (2013). Identification of dynamic models in complex networks with prediction error methods - basic methods for consistent module estimates. *Automatica*, *49*, 2994–3006.

Wahlberg, B., Hjalmarsson, H., & Mårtensson, J. (2009). Variance results for identification of cascade systems. *Automatica*, *45*, 1443 – 1448.

Wong, K., & Polak, E. (1967). Identification of linear discrete time systems using the instrumental variable method. *Automatic Control, IEEE Transactions on*, *12*, 707–718.

Wright, P. G. (1928). *The tariff on animal and vegetable oils*. New York, USA: The Macmillan company.

Yuan, Y., Stan, G., Warnick, S., & Goncalves, J. (2011). Robust dynamical network structure reconstruction. *Automatica*, *47*, 1230 – 1235.