2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

Predictor Input Selection for Two Stage Identification in Dynamic Networks

Arne Dankers, Paul M.J. Van den Hof, Xavier Bombois and Peter S.C. Heuberger

***Abstract*— Recently, the Two-Stage method has been proposed as a tool to obtain consistent estimates of modules embedded in dynamic networks [1], [2]. However, for this method the variables that are included in the predictor model are currently not considered as a user choice. In this paper it is shown that there is considerable freedom as to which variables can be included in the predictor model as inputs, and still obtain consistent estimates of the module of interest. Conditions that the choice of predictor inputs must satisfy are presented. The conditions could be used to find the smallest number of predictor inputs for instance. Algorithms are presented for checking the conditions and obtaining the estimates.**

I. Introduction

Obtaining models of complex dynamic networks from data is an increasingly important area of research. In many fields of science and engineering such as power systems, biological systems, flexible mechanical structures, economic systems, it is becoming possible to collect data at various locations, or of different variables that have dynamic interrelations (i.e. form an interconnected network). However, in some applications, although it is possible to take measurements at many different locations in the network, it may be expensive to do so. Therefore, there is a motivation to use the minimum number of required measurement locations in order to identify a particular module embedded in a network. Secondly, it may be unsafe, or practically infeasible to measure some variables in the network. Therefore, it would be preferable if it is not necessary to measure these variables in order to obtain estimates of the dynamics of interest.

The question addressed in this paper is: given a dynamic network with known interconnection structure, which variables must be included as inputs in the predictor model so that it is possible to obtain consistent estimates of a particular module of interest embedded in the network? Conditions are presented that the set of predictor inputs must satisfy. In this paper the conditions are derived for the Two-Stage Prediction-Error Method as described in [1], [2].

This problem can also be interpreted as determining which variables should be measured (or where should sensors be placed) in order to obtain consistent estimates of a particular module in the network. Using this interpretation, conditions on the sensor placement scheme are presented such that the Two-Stage Method results in consistent estimates of the module of interest.

If sensors are expensive then these conditions can be used to find the smallest number of required sensors. If obtaining data at particular locations in the network is difficult, then the conditions can be used to obtain a predictor input set which avoids using the particular variables.

There is a growing interest in dynamic network identification, ([3], [4], [5] and references therein). If the interconnection structure is not known, then all variables must be used in the predictor input (no choices based on the interconnection structure can be made). However, in the case where the interconnection structure is known, it becomes possible to choose the set of predictor inputs which are optimal in some sense. The results in this paper build on the results for the Two-Stage Prediction-Error Method found in [2], [1].

In Section II required background material is presented, Section III contains the main result and Section IV contains algorithms for a practical implementation of the result.

II. Background

In this section dynamic networks are formally defined, then the Prediction-Error framework and Two-Stage Method are presented, and finally some graph theory is presented.

A. Dynamic Networks and Problem Setup

A dynamic network consists of L internal (or node) variables $\{w_1, \ldots, w_L\}$ that are dynamically interrelated:

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

with G^0_{jk}, $k \in \mathcal{N}_j$ proper transfer functions, q^{-1} the delay operator ($qu(t) = u(t-1)$), and

- $\mathcal{N}_j$ is the set of indices of node variables with direct causal connections to w_j, i.e. $k \in \mathcal{N}_j$ if $G^0_{jk} \neq 0$;
- v_j is an unmeasured variable that is a realization of a stationary stochastic process with rational spectral density: $v_j = H^0_j(q)e_j$ where e_j is white noise, and H^0_j is a monic, stable, minimum phase filter;
- r_j is a external excitation term that is known to the user;

It may be that the disturbance term and/or the external excitation term are not present at some nodes.

All the node variables can be written in one equation as:

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

The work of Arne Dankers is supported in part by the National Science and Engineering Research Council (NSERC) of Canada

A. Dankers, X. Bombois are with the Delft Center for Systems and Control, Delft University of Technology, The Netherlands {a.g.dankers, x.j.a.bombois}@tudelft.nl

P.M.J. Van den Hof is with the Dept. of Electrical Engineering, Eindhoven University of Technology, The Netherlands p.m.j.vandenhof@tue.nl

P.S.C. Heuberger is with the Dept. of Mechanical Engineering, Eindhoven University of Technology, The Netherlands p.s.c.heuberger@tue.nl

978-3-952-41734-8/©2013 EUCA

Using an obvious notation this can be expressed as

$$\begin{aligned} w &= G^0 w + r + H^0 e \\ &= (I - G^0)^{-1}(r + v) \end{aligned} \tag{2}$$

where, if an external variable or disturbance term is absent at node i then the ith entry of r or v respectively is 0.

A *path* from $w_i \to w_j$ will be understood to mean that there are transfer functions such that $G_{jn_1}G_{n_1n_2}\cdots G_{n_k i}$ is nonzero. A *loop* is a path from $w_j \to w_j$.

Since (2) is assumed to describe some physical phenomenon, the equations should be realizable. To characterize the suitability of the equations in describing a physical system, the property of well-posedness is used [6]. The dynamic networks considered are assumed to satisfy the following general conditions.

Assumption 1:

(a) The network is well-posed in the sense that all principal minors of $(I - G^0(\infty))$ are non-zero. [1]
(b) $(I - G^0)^{-1}$ is stable.
(c) The noise process v has a positive semi-definite spectral density, $\Phi_v(\omega) \succeq 0$, not necessarily diagonal

B. Prediction Error Identification

The Prediction-Error framework is an identification framework that is based on the one-step-ahead predictor model. See [7] for a detailed description and analysis.

Let w_j denote the variable which is to be predicted. Let w_k, $k \in \mathcal{D}_j$ and r_k, $k \in \mathcal{R}_j^p$ denote the *predictor inputs* (the set of node variables and external variables that will be used to predict w_j). The one-step-ahead predictor for w_j is: [7]

$$\hat{w}_j(t|t-1,\theta) = \sum_{k\in\mathcal{D}_j} G_{jk}(q,\theta)w_k(t) + \sum_{k\in\mathcal{R}_j^p} F_{jk}(q,\theta)r_k(t). \tag{3}$$

Note that no noise model has been included in the predictor (in the next section it will be shown that noise models are not required for the Two-Stage Method). The *prediction error* is:

$$\begin{aligned} \varepsilon_j(t,\theta) &= w_j(t) - \hat{w}_j(t|t-1,\theta) \\ &= w_j(t) - \sum_{k\in\mathcal{D}_j} G_{jk}(q,\theta)w_k(t) - \sum_{k\in\mathcal{R}_j^p} F_{jk}(q,\theta)r_k(t) \end{aligned} \tag{4}$$

The unknown parameters, θ, are estimated by minimizing the sum of squared (prediction) errors (SSE):

$$V_j(\theta) = \frac{1}{N}\sum_{t=0}^{N-1} \varepsilon_j^2(t,\theta). \tag{5}$$

Under standard (weak) assumptions the estimated parameter vector $\hat{\theta}_N$ converges in the number of data N as [7]

$$\hat{\theta}_N \to \theta^* \text{ with probability 1 as } N \to \infty.$$

where

$$\theta^* = \arg\min_\theta \bar{\mathbb{E}}[\varepsilon_j^2(t,\theta)] \quad \text{and} \quad \bar{\mathbb{E}} := \lim_{N\to\infty} \frac{1}{N}\sum_{t=0}^{N-1}\mathbb{E},$$

and $\mathbb{E}$ is the expected value operator. If $G_{jk}(q,\theta^*) = G_{jk}^0$ the module transfer is said to be estimated *consistently*.

The main result of this paper is to derive rules which can help in choosing appropriate sets $\mathcal{D}_j$ and $\mathcal{R}_j^p$ such that it is possible to obtain a consistent estimate of a particular module G_{ji}^0 embedded in the network using the Two-Stage Method.

C. Two Stage Method

One problem in the identification in networks is that the noise of the predicted variable w_j may be correlated to the predictor inputs w_k, $k \in \mathcal{D}_j$ (if there are loops). Without special attention consistent estimates will not be obtained if the noise is correlated to the predictor inputs [8], [9].

The Two-Stage Method is one way to deal with this problem [8], [9], [1], [2]. The method relies on two facts:

1. Any internal variable w_k can be written as:

$$w_k = \sum_{m\in\mathcal{R}_k} F_{km}r_m + \sum_{m\in\mathcal{V}_k} H_{km}v_m = \sum_{m\in\mathcal{R}_k} w_k^{(r_m)} + \sum_{m\in\mathcal{V}_k} w_k^{(v_m)}$$

where $\mathcal{R}_k$ and $\mathcal{V}_k$ denote the sets of indices of all external and noise variables respectively that have a path to w_k. A new notation is introduced in the second equality. The term $w_k^{(r_m)}$ is referred to as the projection of w_k onto r_m.

2. The transfer function from $w_k^{(r_m)}$ to w_j is the same as the transfer function from w_k to w_j:

$$w_j = \sum_{k\in\mathcal{N}_j} G_{jk}^0 w_k^{(r_m)} + \sum_{k\in\mathcal{N}_j} G_{jk}^0 w_k^{(\perp r_m)} + v_j + r_j$$

where $w_k^{(\perp r_m)} = w_k - w_k^{(r_m)}$.

In the first stage, an estimate of $w_k^{(r_m)}$ is obtained. This is an open-loop problem since r_m is assumed to be uncorrelated to all noise sources and all other external variables.

In the second stage, an estimate of G_{jk}^0 is obtained by estimating the dynamics from $w_k^{(r_m)}$ to w_j. By the second fact, this transfer is the desired transfer.

The generalized procedure is summarized in the following algorithm adapted from [1], [2]. It has been slightly modified to emphasize that (in this paper) which inputs to include in the predictor model is a choice that the user must make.

Algorithm 1: Two Stage Method

1. Choose a set of external excitation signals $\{r_m\}$, $m \in \mathcal{R}_j^s$ to project onto.
2. Choose a set of internal and external variables w_k, $k \in \mathcal{D}_j$, and r_k, $k \in \mathcal{R}_j^p$ to include as inputs to the predictor.
3. Determine $w_k^{(\mathcal{R}_j^s)}$ for $k \in \mathcal{D}_j$, where $w_k^{(\mathcal{R}_j^s)}$ is the projection of w_k onto all r_m, $m \in \mathcal{R}_j^s$.
4. Construct the predictor

$$\hat{w}_j(t|t-1,\theta) = \sum_{k\in\mathcal{D}_j} G_{jk}(\theta)w_k^{(\mathcal{R}_j^s)} + \sum_{k\in\mathcal{R}_j^p} F_{jk}(\theta)r_k. \tag{6}$$

5. Obtain estimates of $G_{jk}(q,\theta^*)$ by minimizing the sum of squared prediction errors (5).

Recall the objective is to obtain consistent estimates of a particular module, G_{ji}^0. In [1] the reasoning is as follows. Choose a set of external variables to project onto that have

[1] This condition is adopted from [6] and imposes weak restrictions on allowable feed-through terms in the network but still allows for the occurrence of algebraic loops.

paths to w_i. Choose $\mathcal{D}_j = \mathcal{N}_j$ and $\mathcal{R}_j^p = \{j\}$ if r_j is present, otherwise choose $\mathcal{R}_j^p = \emptyset$. If a particular w_k, $k \in \mathcal{N}_j$ is not correlated to any r_m, $m \in \mathcal{R}_j^s$, remove k from $\mathcal{D}_j$. In [1] it is shown that for these choices of $\mathcal{R}_j^s$, $\mathcal{R}_j^p$ and $\mathcal{D}_j$ consistent estimates of G_{ji}^0 are possible using Algorithm 1.

This particular choice of predictor inputs works, but it is not the only possible choice. In this paper it will be shown that there is considerable flexibility in the choice of predictor inputs. The chosen set $\mathcal{D}_j$ need not even be a subset of $\mathcal{N}_j$.

D. Some Results From Graph Theory

Some useful concepts from graph theory are briefly presented. A graph G is made up of nodes interconnected by edges. The set of nodes of G is denoted $V(G)$.

Definition 1 (A-B path): Given a directed graph G and sets of nodes A and B. Denote the nodes in the graph x_i. A path $P = x_0 x_1 \cdots x_k$, where the x_i are all distinct, is an *A-B path* if $V(P) \cap A = \{x_0\}$, and $V(P) \cap B = \{x_k\}$. [10].

Definition 2 (A-B Separating Set): Consider a directed graph G. Given $A, B \subset V(G)$, a set $X \subseteq V(G)$ is an *A-B separating set* if the removal of the nodes in X results in a graph with no A-B paths. [10].

III. Predictor Input Selection

In this section conditions that the set of predictor inputs must satisfy are presented such that it is possible to obtain a consistent estimate of G_{ji}^0. This enables the user to choose a set of signals from a given data set such that the conditions are satisfied. Or, equivalently, it enables the user to place sensors in a network in order to collect the required data.

Before stating the main result, two guiding principles behind the main proposition will be illustrated using a few simple examples. The guiding principles for this paper are

1. choose $\mathcal{D}_j$ so that the expression of w_j in terms of w_k, $k \in \mathcal{D}_j$ is such that the transfer between w_i and w_j is G_{ji}^0 (since this will be the dynamics that will be estimated by the predictor) and
2. choose $\mathcal{R}_j^p$ such that the unmodeled component of w_j is uncorrelated to the predictor inputs $w_k^{(\mathcal{R}_j^s)}$ (i.e. the predictor inputs are uncorrelated to the "noise").

The following two examples illustrate the first point.

Example 1: Consider the dynamic network described by:

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{bmatrix} = \begin{bmatrix} 0 & G_{12}^0 & 0 & 0 & 0 \\ G_{21}^0 & 0 & G_{23}^0 & 0 & G_{25}^0 \\ G_{31}^0 & 0 & 0 & 0 & 0 \\ G_{41}^0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & G_{54}^0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{bmatrix} + \begin{bmatrix} v_1 + r_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 + r_5 \end{bmatrix}$$

Suppose the objective is to obtain an estimate of G_{21}^0. Expressing w_2 in terms of only w_1 plus a disturbance:

$$w_2 = (G_{21}^0 + G_{23}^0 G_{31}^0 + G_{25}^0 G_{54}^0 G_{41}^0) w_1 + G_{23}^0 v_3 + G_{25}^0 (v_5 + r_5) + G_{25}^0 G_{54}^0 v_4 + v_2 \quad (7)$$

where the relation between w_1 and w_2 is not G_{21}^0 as desired. If a predictor were constructed to predict w_2 with only w_1 as the input, then from (7), an estimate of the transfer $(G_{21}^0 + G_{23}^0 G_{31}^0 + G_{25}^0 G_{54}^0 G_{41}^0)$ would be obtained. For this example, apparently it is not sufficient to simply estimate the dynamics between w_1 and w_2. Expressing w_2 in terms of both w_1, w_5, and a disturbance results in:

$$w_2 = (G_{21}^0 + G_{23}^0 G_{31}^0) w_1 + G_{25}^0 w_5 + G_{23}^0 v_3 + v_2$$

where again, the relationship between w_1 and w_2 is not G_{21}^0 as desired. For this network, the only way to ensure that the dynamics between w_1 and w_2 are equal to G_{21}^0 is to consider the (proper) mapping from $\{w_1, w_3, w_4, v, r\} \to \{w_2\}$ or $\{w_1, w_3, w_5, v, r\} \to \{w_2\}$ □

It will be shown that one internal variable w_k from every independent path $w_i \to w_j$ must be included as a variable in the proper mapping to w_j to ensure that the dynamical relationship between w_i and w_j is G_{ji}^0, as desired.

Example 2: Consider the network described by:

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 0 & G_{12}^0 & 0 \\ G_{21}^0 & 0 & G_{23}^0 \\ 0 & G_{32}^0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} + \begin{bmatrix} v_1 + r_1 \\ v_2 \\ v_3 \end{bmatrix}$$

Suppose the objective is to obtain an estimate of G_{21}^0. Expressing w_2 in terms of w_1 plus a disturbance results in:

$$w_2 = \frac{1}{1 - G_{23}^0 G_{32}^0} \Big(G_{21}^0 w_1 + G_{23}^0 v_3 + v_2 \Big).$$

The relation between w_1 and w_2 is not G_{21}^0 as desired. The reason the factor $\frac{1}{1 - G_{23}^0 G_{32}^0}$ appears is due to the loop from $w_2 \to w_2$. In the proper mapping from $\{w_1, w_3, v, r\} \to \{w_2\}$ the relation between w_1 and w_2 is G_{21}^0 as desired. □

It will be shown that one internal variable w_k from every independent loop $w_j \to w_j$ must be included as a variable in the proper mapping to w_j to ensure that the dynamical relationship between w_i and w_j is G_{ji}^0, as desired.

The second guiding principle deals with the relation between the unmodeled component of w_j and the chosen predictor inputs w_k, $k \in \mathcal{D}_j$. Again it is first illustrated by example and then proved in the main proposition.

Example 3: Recall the network of Example 1. Again, suppose that the objective is to obtain an estimate of G_{21}^0. Choose to project the predictor inputs onto both r_1 and r_5 ($\mathcal{R}_2^s = \{1, 5\}$). Choose to express w_2 in terms of w_1, w_3, and w_4 ($\mathcal{D}_2 = \{1, 3, 4\}$):

$$w_2 = G_{21}^0 w_1 + G_{23}^0 w_3 + G_{25}^0 G_{54}^0 w_4 + G_{25}^0 v_5 + G_{25}^0 r_5 + v_2.$$

The unmodeled term (or noise) of w_2 is $G_{25}^0 v_5 + G_{25}^0 r_5 + v_2$ which is a function of r_5. Note also that the predictor inputs are also a function of r_5 since the projection of w_1, w_3 and w_4 onto r_5 is non-zero. The point is that the noise term is correlated to the predictor inputs. Therefore, biased estimates would result using the choices $\mathcal{D}_2 = \{1, 3, 4\}$, $\mathcal{R}_2^s = \{1, 5\}$ and $\mathcal{R}_2^p = \emptyset$.

However, choosing $\mathcal{R}_2^p = \{5\}$ would mean that the transfer between r_5 and w_2 is modeled, and consequently, the unmodeled component of w_2 is only a function of v_5 and v_2 which are uncorrelated to the predictor inputs. This is the idea behind the second guiding principle. □

These ideas are now be formalized and proved.

Proposition 1: Consider a dynamic network as defined in Section II-A that satisfies Assumption 1. Let $\{r_m\}$, $m \in \mathcal{R}_j^s$ be the external input(s) onto which will be projected. Let $\{w_k\}$, $k \in \mathcal{D}_j$ and $\{r_k\}$, $k \in \mathcal{R}_j^p$ be the sets of internal and external variables respectively that are included as inputs to the predictor (6). Consistent estimates of G_{ji}^0 are obtained using Algorithm 1 if the following conditions hold:

(a) The external excitation signals r_m, $m \in \mathcal{R}_j^s$ are uncorrelated to all disturbance signals.

(b) The set $\mathcal{D}_j$ satisfies the following conditions:
   1. $i \in \mathcal{D}_j$, $j \notin \mathcal{D}_j$,
   2. every path from $w_i \to w_j$ excluding the path G_{ji}^0 passes through a node w_k, $k \in \mathcal{D}_j$,
   3. every path from $w_j \to w_j$ passes through a node w_k, $k \in \mathcal{D}_j$,

(c) The set $\mathcal{R}_j^p$ satisfies the following condition:
   1. $k \in \mathcal{R}_j^p$ if there is a path from $r_m \to w_j$, $m \in \mathcal{R}_j^s$ that does not pass through a node w_ℓ, $\ell \in \mathcal{D}_j$.

(d) Power spectral density of $[w_{k_1}^{(\mathcal{R}_j^s)} \cdots w_{k_n}^{(\mathcal{R}_j^s)}\; r_{m_1} \cdots r_{m_n}]^T$, $k_* \in \mathcal{D}_j$, $m_* \in \mathcal{R}_j^p$, is positive definite for $\omega \in [-\pi, \pi]$.

(e) The parameterization is chosen flexible enough such that there exists a parameter θ^0 which can exactly describe the dynamics $w_k \to w_j$, $k \in \mathcal{D}_j$ and $r_k \to w_j$, $k \in \mathcal{R}_j^p$.

Before proceeding to the proof of the Proposition, consider an example which highlights the usefulness of the result.

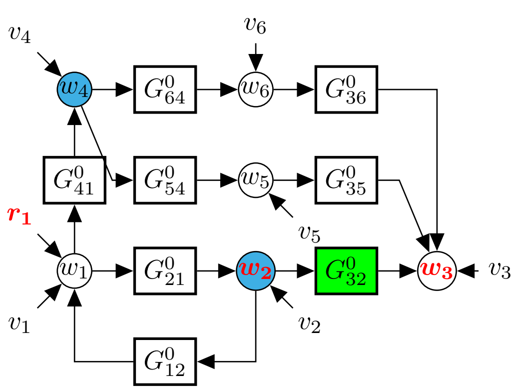

Fig. 1. Example of an interconnected network. For notational convenience labels of the w_i's have been placed inside each summation, which denotes that the output of the sum is the measured variable w_i. The module of interest is denoted in green.

Example 4: Consider the dynamic network in Fig. 1. Suppose the objective is to obtain consistent estimates of G_{32}^0 (marked in green) and r_1 is chosen as the external excitation signal to project onto ($\mathcal{R}_3^s = \{1\}$). From Proposition 1 the set of predictor inputs should satisfy Condition (b). Firstly, there are two paths from w_2 to w_3 (excluding G_{32}^0):

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

Therefore one internal variable from each of the two paths must be included as an input in the predictor. There are no paths from $w_3 \to w_3$ so Condition (b3) is automatically satisfied. Consequently, choosing $\mathcal{D}_3 = \{1, 2\}$ or $\{2, 4\}$ satisfies Condition (b) (denoted in blue in Fig.1).

There are two paths from $r_1 \to w_3$, both of which pass through $\mathcal{D}_3$, so choosing $\mathcal{R}_3^p = \emptyset$ satisfies Condition (c).

If sensors are to be placed in the network in order to obtain consistent estimates of G_{32}^0 then the fewest number of required sensors is 3, and they should be placed so that they measure the predicted variable w_3 (which is required in Step 5 of Algorithm 1 to calculate the prediction error) and and the predictor inputs w_2 and either w_1 or w_4.

By the reasoning in [1] w_2, w_5 and w_6 must be included as inputs in the predictor (6) i.e. $\mathcal{D}_3 = \mathcal{N}_3$. Clearly, the results of Proposition 1 allow for a strictly smaller set of predictor inputs for this example. Moreover, note the set of chosen predictor inputs ($\{w_2, w_4\}$) is not a subset of $\mathcal{N}_3$. $\square$

Note that in order for Condition (d) to hold, there must be a path from at least one r_m, $m \in \mathcal{R}_j^s$ to w_i. If not the $w_i^{(\mathcal{R}_j^s)} = 0$ and the power spectral density of Condition (d) will not be positive definite.

Another interesting point about Proposition 1 is that $\mathcal{R}_j^p \subseteq \mathcal{R}_j^s \setminus D_j$ (i.e. the set of external inputs that must be included in the predictor is a subset of the set of external inputs that the node variables w_k, $k \in \mathcal{D}_j$ are being projected onto). This makes sense in light of Guiding Principle 2: the predictor inputs are functions of r_m, $m \in \mathcal{R}_j^s$, and so the only unmodeled terms of w_j that could cause a problem (i.e. a correlation to the predictor inputs) are those that are functions r_k, $k \in \mathcal{R}_j^s$.

The following lemma is used to prove Proposition 1.

Lemma 1: Consider a dynamic network with transfer matrix G_0 that satisfies all conditions of Assumption 1. Let $\mathtt{G}_{ab}^0$ be the (a, b)th entry of $(I - G_0)^{-1}$. If there is no path from b to a then $\mathtt{G}_{ab}^0 = 0$. $\square$

The result follows directly from Mason's rule [11]. Next Proposition 1 is proved.

Proof: The proof will proceed as follows:

1. Show that w_j can be expressed in terms of w_k, $k \in \mathcal{D}_j$:
$$w_j = \sum_{k \in \mathcal{D}_j} \tilde{G}_{jk}^0 w_k + \sum_{k \in \mathcal{R}_j^p} \tilde{F}_{jk}^0 r_k + \sum_{k \in \mathcal{R}_j \setminus \mathcal{R}_j^s} \tilde{F}_{jk}^0 r_k + \sum_{k \in \mathcal{V}_j} \tilde{H}_{jk}^0 v_k \quad (8)$$
where the dynamics between $w_i \to w_j$ are G_{ji}^0, and $\mathcal{V}_j$ and $\mathcal{R}_j$ denote the sets of indices of all noise sources and external variables respectively that have a path to w_j.
2. Show that at the global minimum of $\bar{V}_j(\theta)$ it must be that $G_{ji}(q, \theta) = G_{ji}^0(q)$.

Step 1. With an abuse of notation, let $\bar{\mathcal{D}}_j$ denote the set of indices k, such that $k \notin \{j, \mathcal{D}_j\}$. Let $w_\mathcal{D}$ denote the vector $[w_{k_1}\; w_{k_2}\; \cdots]^T$, $k_* \in \mathcal{D}_j$, and let $G_{j\mathcal{D}}^0$ denote the vector $[G_{jk_1}^0\; G_{jk_2}^0\; \cdots]$, $k_* \in \mathcal{D}_j$. Let $v_\mathcal{D}$ and $r_\mathcal{D}$ denote vectors of size card$(\mathcal{D}_j)$ where the ith entry is replaced by a 0 if v_i or r_i respectively are not present in the data generating system. Using this notation, the network equations (2) are

$$\begin{bmatrix} w_j \\ w_\mathcal{D} \\ w_{\bar{\mathcal{D}}} \end{bmatrix} = \begin{bmatrix} 0 & G_{j\mathcal{D}} & G_{j\bar{\mathcal{D}}} \\ G_{\mathcal{D}j} & G_{\mathcal{D}\mathcal{D}} & G_{\mathcal{D}\bar{\mathcal{D}}} \\ G_{\bar{\mathcal{D}}j} & G_{\bar{\mathcal{D}}\mathcal{D}} & G_{\bar{\mathcal{D}}\bar{\mathcal{D}}} \end{bmatrix} \begin{bmatrix} w_j \\ w_\mathcal{D} \\ w_{\bar{\mathcal{D}}} \end{bmatrix} + \begin{bmatrix} v_j \\ v_\mathcal{D} \\ v_{\bar{\mathcal{D}}} \end{bmatrix} + \begin{bmatrix} r_j \\ r_\mathcal{D} \\ r_{\bar{\mathcal{D}}} \end{bmatrix}.$$

The variables $w_{\bar{\mathcal{D}}}$ can be eliminated from the equations:

$$\begin{bmatrix} w_j \\ w_{\mathcal{D}} \end{bmatrix} = \begin{bmatrix} 0 & G_{j\mathcal{D}} \\ G_{\mathcal{D}j} & G_{\mathcal{D}\mathcal{D}} \end{bmatrix} \begin{bmatrix} w_j \\ w_{\mathcal{D}} \end{bmatrix} + \begin{bmatrix} G_{j\bar{\mathcal{D}}} \\ G_{\mathcal{D}\bar{\mathcal{D}}} \end{bmatrix} (I - G_{\bar{\mathcal{D}}\bar{\mathcal{D}}})^{-1} \begin{bmatrix} G_{\bar{\mathcal{D}}j} & G_{\bar{\mathcal{D}}\mathcal{D}} \end{bmatrix} \begin{bmatrix} w_j \\ w_{\mathcal{D}} \end{bmatrix}$$
$$+ \begin{bmatrix} G_{j\bar{\mathcal{D}}} \\ G_{\mathcal{D}\bar{\mathcal{D}}} \end{bmatrix} (I - G_{\bar{\mathcal{D}}\bar{\mathcal{D}}})^{-1} (v_{\bar{\mathcal{D}}} + r_{\bar{\mathcal{D}}}) + \begin{bmatrix} v_j + r_j \\ v_{\mathcal{D}} + r_{\mathcal{D}} \end{bmatrix}$$
$$= \begin{bmatrix} \tilde{G}_{jj} & \tilde{G}_{j\mathcal{D}} \\ \tilde{G}_{\bar{\mathcal{D}}j} & \tilde{G}_{\mathcal{D}\mathcal{D}} \end{bmatrix} \begin{bmatrix} w_j \\ w_{\mathcal{D}} \end{bmatrix} + \begin{bmatrix} I & 0 & \tilde{F}_{j\bar{\mathcal{D}}} \\ 0 & I & \tilde{F}_{\mathcal{D}\bar{\mathcal{D}}} \end{bmatrix} \begin{bmatrix} v_j + r_j \\ v_{\mathcal{D}} + r_{\mathcal{D}} \\ v_{\bar{\mathcal{D}}} + r_{\bar{\mathcal{D}}} \end{bmatrix}. \quad (9)$$

Note that by Assumption 1a the inverse $(I - G_{\bar{\mathcal{D}}\bar{\mathcal{D}}})^{-1}$ exists, and that all transfers in the last line are proper.[2]

The following reasoning will show how Condition (b) ensures that $\tilde{G}^0_{jj} = 0$ and $\tilde{G}^0_{ji} = G^0_{ji}$. Consider the term $\tilde{G}^0_{jj}$:

$$\tilde{G}^0_{jj} = G^0_{j\bar{\mathcal{D}}} (I - G^0_{\bar{\mathcal{D}}\bar{\mathcal{D}}})^{-1} G^0_{\bar{\mathcal{D}}j} = \sum_{k_1,k_2 \in \bar{\mathcal{D}}} G^0_{jk_1} \mathtt{G}^0_{k_1k_2} G^0_{k_2j} \quad (10)$$

where $\mathtt{G}^0_{k_1k_2}$ is the (k_1, k_2) entry of $(I - G^0_{\bar{\mathcal{D}}\bar{\mathcal{D}}})^{-1}$. By Lemma 1 if there is no path from $w_{k_2} \to w_{k_1}$, then the transfer $\mathtt{G}^0_{k_1k_2}$ is zero. By Condition (b3) there is no path from $w_j \to w_j$ that does not pass through at least one node in $\mathcal{D}_j$. Therefore, by (10) $\tilde{G}^0_{jj} = 0$.

Secondly, by (9)

$$\tilde{G}_{ji} = G^0_{ji} + G_{j\bar{\mathcal{D}}} (I - G_{\bar{\mathcal{D}}\bar{\mathcal{D}}})^{-1} G_{\bar{\mathcal{D}}i} = G^0_{ji} + \sum_{k_1,k_2 \in \bar{\mathcal{D}}} G^0_{jk_1} \mathtt{G}^0_{k_1k_2} G^0_{k_2i}.$$

Again, by Lemma 1 and Condition (b2) the second term of the above equation is zero. Consequently, the dynamics between $w_i \to w_j$ are G^0_{ji}.

Finally, in order to show that (8) holds, it must be shown that w_j is only a function of r_k, $k \in \mathcal{R}^p_j$ and $k \in \mathcal{R}_j \setminus \mathcal{R}^s_j$. From (9) the transfer between $r_k \to w_j$, $k \in \bar{\mathcal{D}}_j$ is

$$\tilde{G}_{jk} = \big[[G_{j\bar{\mathcal{D}}} (I - G_{\bar{\mathcal{D}}\bar{\mathcal{D}}})^{-1}]\big]_k = \sum_{\ell \in \bar{\mathcal{D}}} G_{j\ell} \mathtt{G}_{\ell k} \quad (11)$$

where the $\big[[\ \cdot\ ]\big]_k$ denotes the kth entry of the vector in square brackets. By Lemma 1, if there is no path from r_k, $k \in \bar{\mathcal{D}}_j$ to w_j, that only passes through nodes in $\bar{\mathcal{D}}_j$ then (11) is zero. If there is such a path, then by Condition (c) m is in $\mathcal{R}^p_j$ if m is in $\mathcal{R}^s_j$ and otherwise m is in $\mathcal{R}_j \setminus \mathcal{R}^s_j$.

From (9) and by combining the last three results, it follows that w_j can be expressed as in (8).

Step 2. In order to determine if consistent estimates of G^0_{ji} will result by minimizing (5) it must be determined if the minimizer of

$$\bar{V}_j(\theta) = \lim_{N \to \infty} \mathbb{E}[V_N(\theta)] = \bar{\mathbb{E}}[\varepsilon^2_j(t, \theta)] \quad (12)$$

is equal to G^0_{ji}. Substitute the predictor (6) into (12):

$$\bar{V}_j(\theta) = \bar{\mathbb{E}}\Big[\Big(w_j - \sum_{k \in \mathcal{D}_j} G_{jk}(\theta) w_k^{(\mathcal{R}^s_j)} - \sum_{k \in \mathcal{R}^p_j} F_{jk}(\theta) r_k\Big)^2\Big]$$
$$= \bar{\mathbb{E}}\Big[\Big(\sum_{k \in \mathcal{D}_j} \Delta\tilde{G}_{jk}(\theta) w_k^{(\mathcal{R}^s_j)} + \sum_{k \in \mathcal{R}^p_j} \Delta\tilde{F}_{jk}(\theta) r_k + s_j\Big)^2\Big]$$

[2] From control theory $G(\infty) \neq 0$ if and only if both G and G^{-1} are proper.

where $\Delta\tilde{G}_{jk}(q, \theta) = \tilde{G}^0_{jk}(q) - G_{jk}(q, \theta)$, and $\Delta\tilde{F}_{jk}(q, \theta) = \tilde{G}^0_{jk}(q) - F_{jk}(q, \theta)$ and

$$s_j = \sum_{k \in \mathcal{D}_j} \tilde{G}_{jk} w_k^{(\perp\mathcal{R}^s_j)} + \sum_{k \in \mathcal{R}_j \setminus \mathcal{R}^s_j} \tilde{F}_{jk} r_k + \sum_{k \in \mathcal{V}^p_j} \tilde{H}_{jk} v_k$$

where $w_k^{(\perp\mathcal{R}^s_j)} = w_k - w_k^{(\mathcal{R}^s_j)}$ and $\mathcal{V}^p_j$ is the set of indices of all noise terms v_k that are present in the expression of w_j for the particular choice of $\mathcal{D}_j$. All terms that make up s_j are uncorrelated to r_m, $m \in \mathcal{R}^s_j$, which results in the following simplification:

$$\bar{V}_j(\theta) = \bar{\mathbb{E}}\Big[\Big(\sum_{k \in \mathcal{D}_j} \Delta G_{jk}(\theta) w_k^{(\mathcal{R}^s_j)} + \sum_{k \in \mathcal{R}^p_j} \Delta\tilde{F}_{jk}(\theta) r_k\Big)^2\Big] + \bar{\mathbb{E}}\big[(s_j)^2\big]$$

Let

$$\Delta X^T(e^{j\omega}, \theta) = \big[\Delta G_{jn_1}(e^{j\omega}, \theta) \ \cdots \ \Delta G_{jn_{\mathtt{n}}}(e^{j\omega}, \theta)\big]$$

then, by Parseval's theorem, $\bar{V}_j(\theta)$ is equal to

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \Delta X^T(e^{j\omega}, \theta) \Phi_{\mathcal{D}}^{(\mathcal{R}^s_j)}(\omega) \Delta X(e^{-j\omega}, \theta) + |S_j(e^{j\omega})|^2 \mathrm{d}\omega.$$

By Condition (d) $\Phi_{\mathcal{D}}^{(\mathcal{R}^s_j)}(\omega)$ is positive definite for all $\omega \in [\pi, \pi]$, therefore both terms in the sum are positive for all ω. Consequently, $\bar{V}_j(\theta)$ will be a minimum when $\Delta X(\theta) = 0$.

By Condition (e) $\Delta X(\theta) = 0$ when $G_{jk}(q, \theta) = \tilde{G}^0_{jk}(q, \mathcal{D}_j)$ and $F_{jk}(q, \theta) = \tilde{G}^0_{jk}(q, \mathcal{D}_j)$. Consequently, it follows that $G_{ji}(q, \theta) = G^0_{ji}(q)$ ∎

IV. Algorithms

Two algorithms will be presented. The first one provides a way checking Condition (b), or of finding a set $\mathcal{D}_j$ that satisfies Condition (b). The second one is an efficient parameterization to obtain the estimates.

Condition (b) can be reformulated using the notion of separating sets defined in Definition 2. Let the node w_j be split into two nodes: w_j^+ to which all paths coming into w_j are connected and w_j^- to which all paths leaving w_j are connected. The node w_j^+ is connected to w_j^- with the path $G_{j^+j^-} = 1$. The set $\mathcal{D}_j$ satisfies the following conditions:

1. $\mathcal{D}_j \setminus \{i\}$ is a $\{w_i\}$-$\{w_j\}$ separating set for the network with path G^0_{ji} removed,
2. $\mathcal{D}_j$ is a $\{w_j^-\}$-$\{w_j^+\}$ separating set.

These two conditions can be formulated as a single condition. Let w_i^+ and w_i^- be defined similar to w_j^+ and w_j^-. Then the set $\mathcal{D}_j$ is a $\{w_i^-, w_j^-\}$-$\{w_j^+\}$ separating set for the network with edge G^0_{ji} removed.

The advantage of reformulating the conditions in terms of separating sets is that there exist tools from graph theory to check if a given set is a separating set or to find (the smallest possible) separating sets [10], [12]. The following example will illustrate the use of separating sets.

Example 5: Consider the network shown in Fig. 2. Suppose that the objective is to obtain consistent estimates of G^0_{21} (denoted in green). Both w_1 and w_2 have been split into two nodes as described above. Choose to project onto all external variables that are present ($\mathcal{R}^s_2 = \{4, 5, 8\}$).

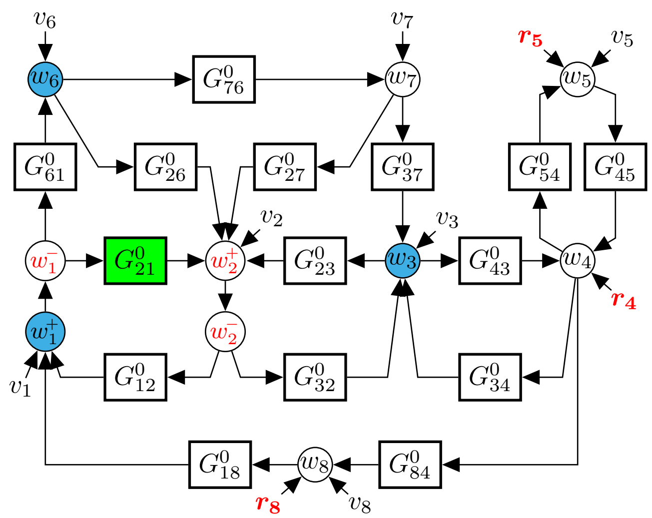

Fig. 2. Example of an interconnected network. For notational convenience labels of the w_i's have been placed inside each summation, which denotes that the output of the sum is the measured variable w_i.

By the above reasoning using separating sets, Conditions (b2)-(b3) are satisfied for the given network if $\mathcal{D}_2$ is a $\{w_1^-, w_2^-\}$-$\{w_2^+\}$ separating set for the network with G^0_{21} removed. These variables are marked in red in the figure.

There are many possible choices of $\mathcal{D}_2$, but the smallest choice is $\{w_1, w_6, w_3\}$ denoted in blue. Again, this is fewer predictor inputs than required in [1]. □

The second algorithm that is presented provides an efficient method to obtain the estimates. A main theme of this paper is that it is often necessary to include more than just w_i as an input to the predictor model. However, the dynamics that are estimated in addition to G^0_{ji} are not really of interest. A complicating factor is that the parameterization of all transfers $\tilde{G}_{jk}(\theta)$, $k \in \mathcal{D}_j$ must be chosen such that the data generating system can be represented by a parameter vector θ^0 in the model set (Condition (e) of Proposition 1).

Often (particularly in control applications) it is desirable to parameterize G^0_{ji} as

$$G_{ji}(q,\theta) = \frac{b_0^{ji} + b_1^{ji} q^{-1} + \cdots + b_{n_b}^{ji} q^{-n_b}}{1 + a_1^{ji} q^{-1} + \cdots + a_{n_a}^{ji} q^{-n_a}}. \qquad (13)$$

where n_a and n_b must be chosen a priori. However, each transfer $\tilde{G}_{jk}(q,\theta)$, $k \in \mathcal{D}_j \setminus \{i\}$ could be parameterized as high order FIR models or using basis functions:

$$G_{jk}(q,\theta) = \sum_{\ell=0}^{n} c_\ell^{jk} F_\ell(q)$$

where $F_\ell(q)$ are basis functions. Using these types of models does not require much a priori knowledge, and since the models of these transfer functions are not the primary objective of the identification it does not matter that the models are not in the form (13).

An advantage of this parameterization is that the parameters c_i^{jk} appear quadratically in the objective function (5) which means a separable least squares algorithm [13] can be used to minimize the objective function.

Remark 1: As a (rough) rule of thumb, the fewer the parameters to estimate, the lower the variance of the parameters [7] as long as the unmodeled disturbance term on w_j stays the same by removing the parameters. By this reasoning, using fewer predictor inputs may result in estimated parameters with a lower variance.

Consider the network in Example 4. Choose $\mathcal{D}_3 = \mathcal{N}_3$. Suppose that 100 tap FIR models are used to parameterize G^0_{35} and G^0_{36}. Then $200 + n_a + n_b + 1$ parameters need to be estimated.

Next choose $\mathcal{D}_3 = \{w_4, w_2\}$ as suggested in Example 4, then only $100 + n_a + n_b + 1$ parameters need to be estimated. The disturbance term (denoted s_3 in the proof of the proposition) affecting w_3 has not changed suggesting that the estimates in the second case will have lower variance. □

V. Conclusion

Conditions on the predictor inputs have been presented such that it is possible to obtain consistent estimates of a module embedded in a dynamic network using the Two Stage method of identification. This enables the user to design the least expensive sensor placement scheme for instance. Moreover, large networks do not pose a problem since there exist efficient algorithms to find separating sets.

VI. Acknowledgment

Thanks to Yasin Gunes for noticing Mason's rules are an easy proof of Lemma 1.

References

[1] P. M. J. Van den Hof, A. Dankers, P. S. C. Heuberger, and X. Bombois, "Identification of dynamic models in complex networks with prediction error methods - basic methods for consistent module estimates," *Automatica*, 2012, submitted.

[2] ——, "Identification in dynamic networks with known interconnection topology," in *Proceedings of 51 IEEE Conference on Decision and Control*, 2012, pp. 895–900.

[3] D. Materassi, M. Salapaka, and L. Giarré, "Relations between structure and estimators in networks of dynamical systems," in *Proceedings of 50th IEEE Conference on Decision and Control*, Orlando, USA, 2011.

[4] B. Sanandaji, T. Vincent, and M. Wakin, "Exact topology identification of large-scale interconnected dynamical systems from compressive observations," in *Proceedings of American Control Conference*, San Francisco, CA, USA, 2011, pp. 649–656.

[5] Y. Yuan, G.-B. Stan, S. Warnick, and J. Gonçalves, "Robust dynamical network reconstruction," in *Proceedings of 49th IEEE Conference on Decision and Control*, Atlanta, USA, 2010, pp. 810–815.

[6] M. Araki and M. Saeki, "Quantitative condition for the well-posedness of interconnected dynamical systems," *IEEE Transactions on Automatic Control*, vol. 28, no. 5, pp. 625–637, May 1983.

[7] L. Ljung, *System Identification. Theory for the User*, 2nd ed. Prentice Hall, 1999.

[8] U. Forssell and L. Ljung, "Closed-loop identification revisited," *Automatica*, vol. 35, pp. 1215–1241, 1999.

[9] P. M. J. Van den Hof and R. Schrama, "An indirect method for transfer function estimation from closed loop data," *Automatica*, vol. 29, no. 6, pp. 1523–1527, 1993.

[10] R. Diestel, *Graph Theory*, ser. Graduate Texts in Mathematics. Springer-Verlag, 1997, no. 173.

[11] S. J. Mason, "Feedback theory - further properties of signal flow graphs," *Proceedings of The IRE*, pp. 920–926, Jul. 1956.

[12] A. Kanevsky, "Finding all minimum-size separating vertex sets in a graph," *Networks*, vol. 23, pp. 533–541, 1993.

[13] A. Ruhe and P. A. Wedin, "Algorithms for separable nonlinear least squares problems," *Society for Industrial and Applied Mathematics Review*, vol. 22, pp. 318–337, Jul. 1980.