51st IEEE Conference on Decision and Control
December 10-13, 2012. Maui, Hawaii, USA

Dynamic Network Identification Using the Direct Prediction-Error Method

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

Abstract— **The problem of identifying dynamical models on the basis of measurement data is usually considered in a classical open-loop or closed-loop setting. In this paper this problem is generalized to linear dynamical systems that operate in a complex interconnection structure and the dynamical relationships between the measured variables signals need to be identified. It is shown that the classical Direct Method of closed-loop identification in the prediction-error context can be generalized to provide consistent model estimates, under specified experimental circumstances.**

I. Introduction

In the control field there is a trend towards distributed and networked control. Models are required to use model based control for interconnected systems. The goal of the work presented here is to obtain good models from data generated by a network.

The following terminology will be used [1]. The dynamic blocks of the network will be referred to as *modules*. The modules are *embedded* into a network by connecting the terminals of the modules to the terminals of other modules. The result is an *interconnection structure*.

In the companion paper [2], the opportunities to generalize closed-loop methods to dynamic networks are investigated. In this paper, attention will be focused on generalizing the Direct Method of closed loop identification ([3], [4]) to more complex interconnection structures. In the bulk of the paper it will be assumed that the interconnection structure of the network is known. It will be shown that the conditions such that the generalized Direct Method leads to consistent estimates of the module dynamics are very similar to those of the closed loop Direct Method. Moreover, it will be shown that some of these conditions are not just sufficient, but also generically necessary for consistent estimates.

Under particular assumptions the generalized Direct Method can also be used to identify the dynamics of a network with unknown interconnection structure. Importantly, the conditions presented in this paper are less restrictive than the conditions that are often imposed in the network identification literature [5], [6].

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

One of the mechanisms of the Direct Method that ensures consistent estimates is exact noise modeling. This mechanism has also been investigated in [7]. However, it could be questioned if it is practically feasible to exactly model all noise sources. In the companion paper, [2], methods of obtaining consistent estimates without exact noise modeling are investigated.

The paper is organized as follows: in Section II a dynamic network is defined and the Direct Method is presented; in Section III conditions under which the Direct Method leads to consistent estimates of the module dynamics are presented; in Section IV necessity of one of the conditions is proved; and in Section V it is shown how the Direct Method can be used when the interconnection structure is unknown.

II. Problem Setup

In this section, a dynamic network will be defined, and the Direct Method of identification will be presented.

A. Dynamic Networks

Throughout the paper it will be assumed that there are L measured variables $\{w_1, \dots, w_L\}$. As a generalization of the way measurements are taken in a closed loop, it will be assumed that measurements are taken such that each measured variable w_j can be written as:

$$w_j(t) = \sum_{i \in \mathcal{N}_j} G^0_{ji}(q) w_i(t) + G^0_{jr}(q) r_j(t) + v_j(t) \tag{1}$$

where G^0_{ji} represents the module dynamics of the data generating system, v_j represents a noise source, r_j represents an external excitation source which can be manipulated by the user, $\mathcal{N}_j$ denotes the set of indices i for which $G^0_{ji} \neq 0$, (i.e. the set of measured variables with direct causal connections to w_j) and q^{-1} is the delay operator (i.e. $q^{-1}u(t) = u(t-1)$). The noise and external excitation sources may or may not be present. The noise sources are considered to be stationary zero mean stochastic processes. It will be assumed that v_j can be represented as filtered white noise: $v_j(t) = H^0_j(q) e_j(t)$, where e is white noise. All the measured variables can be written in a single matrix equation as:

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} = \begin{bmatrix} 0 & G^0_{12} & \cdots & G^0_{1L} \\ G^0_{21} & 0 & \ddots & G^0_{2L} \\ \vdots & \ddots & \ddots & \vdots \\ G^0_{L1} & G^0_{L2} & \cdots & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix} + \begin{bmatrix} G_{1r} r_1 \\ G_{2r} r_2 \\ \vdots \\ G_{Lr} r_L \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_L \end{bmatrix}$$

$$= G^0 w + G^0_r r + H^0 e \tag{2a}$$

$$= (I - G^0)^{-1}(G^0_r r + H^0 e) \tag{2b}$$

The network is completely characterized by (G^0, G_r^0, H^0).

Often, it is natural to talk about 'paths' and 'loops' in a network. What is meant by a 'path' from i to j is that there are transfer functions such that the transfer function $G_{jn_1}G_{n_1n_2}\cdots G_{n_ki}$ is nonzero. A path without loops has $j, n_1, \ldots, n_k, i$ all unique. The length of the path is the number of transfer functions, n_k+1. A loop is a path from j to j. In the above example, if $i=j$, a loop results.

A network will be well posed if all module transfer functions are proper (causal) and if $\det(I-G^0)\neq 0$.

B. The Direct Method

In the closed-loop system identification literature the so called Direct Method can be used to consistently identify the plant dynamics if (1) there is noise present in the loop, (2) an external input is connected to the loop, and (3) there is a delay present in the loop [3]. This method will be extended to obtain consistent estimates of all the module dynamics embedded in the network.

The generalized Direct Method is presented in the following text. The transfer function modules and noise filters are modeled as transfer functions with unknown coefficients:

$$G_{ji}(q,\theta) = \frac{b_0^{ji} + b_1^{ji}q^{-1} + \cdots + b_{n_b}^{ji}q^{-n_b}}{1 + a_1^{ji}q^{-1} + \cdots + a_{n_a}^{ji}q^{-n_a}} \quad (3)$$

$$H_j(q,\theta) = \frac{1 + c_1^{j}q^{-1} + \cdots + c_{n_c}^{j}q^{-n_c}}{1 + d_1^{j}q^{-1} + \cdots + d_{n_d}^{j}q^{-n_d}} \quad (4)$$

where $i\in\mathcal{N}_j$ and $G_{jr}(q,\theta)$ is parameterized in the same way (ie. $i=r$ in the parameterization of $G_{ji}(q,\theta)$). Note that it is not necessary that n_a, n_b, n_c, n_d are the same for all $G_{ji}(\theta)$, however, for notational clarity they will be denoted the same. The one-step-ahead predictor for w_j is [4]:

$$\hat{w}_j(t,\theta) = H_j^{-1}(q,\theta)\Big(\sum_{i\in\mathcal{N}_j} G_{ji}(q,\theta)w_i(t) + G_{jr}(q,\theta)r_j(t)\Big) + \big(1-H_j^{-1}(q,\theta)\big)w_j(t) \quad (5)$$

which results in the following prediction error:

$$\begin{aligned}\varepsilon_j(t,\theta) &= w_j(t) - \hat{w}_j(t,\theta)\\ &= H_j^{-1}(q,\theta)\Big(w_j(t) - \sum_{i\in\mathcal{N}_j} G_{ji}(q,\theta)w_i(t) - G_{jr}(q,\theta)r_j(t)\Big)\\ &= H_j^{-1}(q,\theta)\Big(\sum_{i\in\mathcal{N}_j}\big(G_{ji}^0(q) - G_{ji}(q,\theta)\big)w_i(t)\\ &\qquad + \big(G_{jr}^0(q) - G_{jr}(q,\theta)\big)r_j(t) + v_j(t)\Big). \end{aligned} \quad (6)$$

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

$$V_j(\theta) = \frac{1}{N}\sum_{t=0}^{N-1}\varepsilon_j^2(t,\theta). \quad (7)$$

The question which will be investigated in this paper is under which conditions minimizing (7) results in consistent estimates of G_{ji}^0 and G_{jr}^0. In other words, if there exists a θ_0 such that $G_{ji}(\theta_0) = G_{ji}^0$ and $H_j(\theta_0) = H_j^0$, does minimizing (7) lead to a consistent estimate of θ_0? This question will be analyzed using the function:

$$\bar{V}_j(\theta) = \lim_{N\to\infty}\mathbb{E}[V_j(\theta)] = \bar{\mathbb{E}}[\varepsilon_j^2(t,\theta)] \quad (8)$$

Under standard assumptions of the Prediction Error identification literature, and that the external inputs $\{r_i(t)\}$ are bounded, then minimizing $V_j(\theta)$ is equivalent to minimizing $\bar{V}_j(\theta)$. For a formal analysis see Lemma 8.2 in [4].

III. Conditions for Consistent Estimates

The Direct Method for identification of the dynamics of a network is presented as a "local" method in the sense that the transfers G_{ji}^0, $i\in\mathcal{N}_j$ can be estimated without estimating the other transfer functions in the network. Moreover the conditions do not depend on global properties of the network, only on properties related to variable w_j. Before the statement and proof of the main theorem, a lemma will be introduced.

Lemma 1: Let $\mathcal{G}^0 = (I-G^0)^{-1}$, where G^0 is defined in (2b). The (i,j)th entry of $\mathcal{G}^0$, denoted $\mathcal{G}_{ij}^0$, has a delay if every path from j to i has a delay. Similarly, $\mathcal{G}_{ij}^0 = 0$ if there is no path from j to i.

The proof of the lemma is presented in the Appendix. The following theorem will present conditions under which the transfer functions G_{ji}^0, $i\in\mathcal{N}_j$ can be consistently estimated. The following notation will be used in the theorem: let $\mathcal{V}_j$ denote the set of indices of noise sources for which there are paths to j. Similarly, let $\mathcal{R}_j$ denote the set of indices of external inputs for which there are paths to j. Let $\mathfrak{n}_j$ be the number of elements in $\mathcal{N}_j$.

Theorem 1: Consider a dynamic network (G^0, G_r^0, H^0). Assume that the interconnection architecture is known. Assume the following conditions hold:

(a) The network is well posed.
(b) The network is stable in the sense that all signals are bounded.
(c) A noise source v_j is present that is uncorrelated to the external inputs r_k, $k\in\mathcal{R}_j$ and the noise sources v_k, $k\in\mathcal{V}_j\setminus j$
(d) The spectrum of $[w_j\ \ r_j\ \ w_{n_1}\ \cdots\ w_{n_{\mathfrak{n}}}]^T$, $n_*\in\mathcal{N}_j$, $\Phi_{\{j,\mathcal{N}_j\}}(e^{j\omega})$ is positive definite for $\omega\in[-\pi,\pi]$.
(e) Every path from j to j has a delay.
(f) It is known which transfer functions G_{ji}^0 have a delay.
(g) The data generating system (2b) is in the set of possible models (i.e. there exists a θ_0 such that $G_{ji}(\theta_0) = G_{ji}^0$, $i\in\mathcal{N}_j$, $G_{jr}(\theta) = G_{jr}^0$ and $H_j(\theta_0) = H_j^0$).

Then all the module transfer functions $G_{ji}^0, i\in\mathcal{N}_j$ and G_{jr}^0 can be estimated consistently using the Direct Method.

Proof: The proof will proceed as follows:

1) Show that the lower bound of the objective function $\bar{V}_j(\theta)$ (defined in (8)) is $\sigma_{e_j}^2$, the variance of e_j.
2) Show that $\bar{V}_j(\theta) = \sigma_{e_j}^2$ implies that $\theta=\theta_0$ (i.e the global minimum is attainable and unique).

In this proof, it will be assumed that $G_{jr}^0 = 0$. This is to save space, and increase the clarity of the proof. Every step of the proof would be exactly the same if $G_{jr}^0 \neq 0$ only an extra term would be included in every summation.

Step 1. Throughout the proof, it will be useful to expand the measured variable w_i in terms of all noise sources and external inputs that affect w_i. From (2b) and using the notation from Lemma 1 we have:

$$w_i = \sum_{k=1}^{L} \mathcal{G}^0_{ik}(v_k + r_k) = \sum_{k\in\mathcal{V}_i} \mathcal{G}^0_{ik} v_k + \sum_{k\in\mathcal{R}_i} \mathcal{G}^0_{ik} r_k \qquad (9)$$

where the second equality holds by Lemma 1 and the definitions of $\mathcal{V}_i$ and $\mathcal{R}_i$.

Now, (9) will be used to express the objective function in terms of only noise sources and external inputs. Substituting (6) into (8), $\bar{V}_j(\theta)$ can be written as:

$$\begin{aligned}
\bar{V}_j(\theta) &= \bar{\mathbb{E}}\Big[\Big(H_j^{-1}(\theta)\Big(v_j + \sum_{i\in\mathcal{N}_j}(G^0_{ji} - G_{ji}(\theta))w_i\Big)\Big)^2\Big] \\
&= \bar{\mathbb{E}}\Big[\Big(H_j^{-1}(\theta)\Big(v_j + \sum_{i\in\mathcal{N}_j}\Delta G_{ji}(\theta)\Big(\sum_{k\in\mathcal{V}_i}\mathcal{G}^0_{ik}v_k + \sum_{k\in\mathcal{R}_i}\mathcal{G}^0_{ik}r_k\Big)\Big)\Big)^2\Big] \\
&= \bar{\mathbb{E}}\Big[\Big(\Delta H_j(\theta)v_j + H_j^{-1}(\theta)\sum_{i\in\mathcal{N}_j}\sum_{k\in\mathcal{V}_i}\Delta G_{ji}(\theta)\mathcal{G}^0_{ik}v_k \\
&\qquad + H_j^{-1}(\theta)\sum_{i\in\mathcal{N}_j}\sum_{k\in\mathcal{R}_i}\Delta G_{ji}(\theta)\mathcal{G}^0_{ik}r_k + e_j\Big)^2\Big] \qquad (10)
\end{aligned}$$

where $\Delta G_{ji}(\theta) = G^0_{ji} - G_{ji}(\theta)$, and $\Delta H_j(\theta) = H_j^{-1}(\theta) - H_j^{0^{-1}}$. Next Conditions (f) and (e) will be used to simplify this expression.

By (f) if G^0_{ji} has a delay, then $G_{ji}(\theta)$ will be parameterized with a delay (i.e. $\Delta G_{ji}(\theta)$ has a delay if G^0_{ji} has a delay). Moreover, by Lemma 1 the term $G^0_{ji}\mathcal{G}_{ij}$ has a delay if all paths from j to j have a delay. By Condition (e), every path from j to j has a delay, therefore, $\Delta G_{ji}(\theta)\mathcal{G}^0_{ij}$ has a delay for all i.

Consequently every term in (10) is uncorrelated to e_j:

- since $H_j(\theta)$ and H^0_j are both monic, $\Delta H_j(\theta)v_j$ is a function of $v_j(t-k)$, $k > 1$;
- as described above, $\Delta G_{ji}(\theta)\mathcal{G}^0_{ij}v_j$ is also a function of $v_j(t-k)$ $k > 1$;
- by Condition (c) any term involving v_k, $k \in \mathcal{V}_j$, $k \neq j$ is uncorrelated to e_j;
- by Condition (c), e_j is uncorrelated to r_k for all k.

Using this reasoning to simplify (10) results in:

$$\begin{aligned}
\bar{V}_j(\theta) &= \bar{\mathbb{E}}\Big[\Big(\Delta H_j(\theta)v_j + H_j^{-1}(\theta)\sum_{i\in\mathcal{N}_j}\sum_{k\in\mathcal{V}_i}\Delta G_{ji}(\theta)\mathcal{G}^0_{ik}v_k \\
&\qquad + H_j^{-1}(\theta)\sum_{i\in\mathcal{N}_j}\sum_{k\in\mathcal{R}_i}\Delta G_{ji}(\theta)\mathcal{G}^0_{ik}r_k\Big)^2\Big] + \sigma^2_{e_j} \\
&= \bar{\mathbb{E}}\Big[\Big(\Delta H_j(\theta)v_j + H_j^{-1}(\theta)\sum_{i\in\mathcal{N}_j}\Delta G_{ji}(\theta)w_i\Big)^2\Big] + \sigma^2_{e_j} \qquad (11)
\end{aligned}$$

where $\sigma^2_{e_j}$ is the variance of e_j. From (11), it is clear that $\bar{V}_j(\theta) \geq \sigma^2_{e_j}$. This concludes the first step.

Step 2. Next it must be shown that the global minimum of $\bar{V}_j(\theta)$ is attainable and unique. This will be done by showing

$$\bar{V}_j(\theta) = \sigma^2_{e_j} \;\Rightarrow\; \theta = \theta_0.$$

Using (11), $\bar{V}_j(\theta) = \sigma^2_{e_j}$ can be written as

$$\bar{\mathbb{E}}\Big[\sum_{i\in\mathcal{N}_j}\frac{\Delta G_{ji}(\theta)}{H_j(\theta)}w_i + \Delta H_j(\theta)v_j\Big)^2\Big] + \sigma^2_{e_j} = \sigma^2_{e_j}$$

$$\bar{\mathbb{E}}\left[\left(\begin{bmatrix}\Delta H_j(\theta) & \frac{\Delta G_{jn_1}(\theta)}{H_j(\theta)} & \cdots & \frac{\Delta G_{jn_n}(\theta)}{H_j(\theta)}\end{bmatrix}\begin{bmatrix}v_j\\ w_{n_1}\\ \vdots\\ w_{n_n}\end{bmatrix}\right)^2\right] = 0$$

$$\bar{\mathbb{E}}\left[\left(\Delta x(\theta)\begin{bmatrix}1 & -G^0_{jn_1} & \cdots & -G^0_{jn_n}\\ & 1 & & \\ & & \ddots & \\ & & & 1\end{bmatrix}\begin{bmatrix}w_j\\ w_{n_1}\\ \vdots\\ w_{n_n}\end{bmatrix}\right)^2\right] = 0$$

$$\bar{\mathbb{E}}\Big[\Big(\Delta x(\theta)^T J w_{\{j,\mathcal{N}_j\}}\Big)^2\Big] = 0 \qquad (12)$$

where

$$\Delta x(\theta)^T = \Big[\Delta H_j(\theta) \;\; \frac{\Delta G_{jn_1}(\theta)}{H_j(\theta)} \;\cdots\; \frac{\Delta G_{jn_n}(\theta)}{H_j(\theta)}\Big],$$

$$w^T_{\{i,\mathcal{N}_j\}} = [w_j \;\; w_{n_1} \;\cdots\; w_{n_n}],\; n_k \in \mathcal{N}_j.$$

Using Parseval's Theorem results in:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\Delta x(e^{j\omega},\theta)^T J\Phi_{\{j,\mathcal{N}_j\}}(\omega)J^*\Delta x(e^{-j\omega},\theta)\mathrm{d}\omega = 0 \qquad (13)$$

for $\omega \in [-\pi,\pi)$, where J^* denotes the conjugate transpose of J. By Condition (d), $\Phi_{\{j,\mathcal{N}_j\}}(\omega)$ is positive definite. Moreover, $J(e^{j\omega})$ is full rank for all ω. Thus the only way (13) can equal zero is if each entry of $[\Delta H_j \; \Delta G_{jn_1} \cdots \Delta G_{jn_{n_1}}]$ is equal to zero for all ω. Therefore, by Condition (g) and if the parameterization of $G_{ji}(\theta)$ is such that the only way that $G^0_{ji} - G_{ji}(\theta)$ is equal to zero is when $G_{ji}(\theta) = G^0_{ji}$, the global minimum of $\bar{V}_j(\theta)$ is unique. ■

Remark 1: There exists an alternative proof of this theorem. From the closed loop identification literature it is known that if there are no algebraic loops in the system, then the direct method can be used to obtain consistent estimates of the plant [8]. Therefore, (1) show that it is possible to transform any dynamic network into an equivalent dynamic network which has an interconnection structure consisting of a single closed loop where w_j is the only variable being fed back. Then, (2), Lemma 1 can be used to show that if all paths in the network that loop through w_j have a delay then there is no algebraic loop. Thus the conditions of [8] are satisfied, and consequently, consistent estimates are possible.

It is interesting that the Direct Method is local in the sense that it is possible to consistently estimate the dynamics G_{ji}, $i \in \mathcal{N}_j$ without estimating any other dynamics G_{ki}, $k \neq j$. Moreover all the conditions, except Condition (a) are also local to variable w_j. Next, it will be shown that Condition (e) is not only sufficient, but also generically necessary.

IV. Necessity of Condition (e)

Theorem 2: Consider a dynamic network as described in Theorem 1. Let $G_{ji}(\theta)$, $i \in \mathcal{N}_j$, $G_{jr}(\theta)$ and $H_j(\theta)$ be independently parameterized. Under the conditions of Theorem 1 consistent estimates of the transfer functions G^0_{ji}, $i \in \mathcal{N}_j$ and

G^0_{jr} are generically possible using the Direct Method only if every loop that passes through w_j has a delay.

Proof: The proof will proceed as follows: by condition (g), there exists a θ_0 such that $G_{ji}(\theta_0) = G^0_{ji}$, $i \in \mathcal{N}_j$, $G_{jr}(\theta) = G^0_{jr}$ and $H_j(\theta_0) = H^0_j$. It will be shown that the derivative of $\bar{V}_j(\theta)$ can only equal zero at θ_0 if every loop that passes through j has at least one delay.

First the derivative of $\bar{V}_j(\theta)$ must be calculated. Let θ be partitioned as:

$$\theta^T = \begin{bmatrix} b^T_{jn_1} & a^T_{jn_1} & \cdots & b^T_{jn_\mathfrak{n}} & a^T_{jn_\mathfrak{n}} & c^T_j & d^T_j \end{bmatrix},$$
$$b^T_{ji} = [b^{ji}_0 \ \cdots \ b^{ji}_{n_b}], \qquad a^T_{ji} = [a^{ji}_1 \ \cdots \ a^{ji}_{n_a}],$$
$$c^T_j = [c^j_1 \ \cdots \ c^j_{n_c}], \qquad d^T_j = [d^j_1 \ \cdots \ d^j_{n_d}],$$

where $a^{ji}_k, b^{ji}_k, c^j_k, d^j_k$ are the parameters of the model transfer functions $G_{ji}(\theta)$, $i \in \mathcal{N}_j$, $H_j(\theta)$ as defined in (3), $n_1, \ldots, n_\mathfrak{n} \in \mathcal{N}_j$. The derivative of $\bar{V}_j(\theta)$ is:

$$\frac{\mathrm{d}\bar{V}_j(\theta)}{\mathrm{d}\theta} = 2\bar{\mathbb{E}}\Big[\varepsilon_j(\theta)\frac{\mathrm{d}\varepsilon_j(\theta)}{\mathrm{d}\theta}\Big]$$
$$= -2\bar{\mathbb{E}}\Big[\varepsilon_j(\theta)H_j^{-1}(\theta)\Big(\frac{\mathrm{d}H_j(\theta)}{\mathrm{d}\theta}\varepsilon_j(\theta) + \sum_{i\in\mathcal{N}_j}\frac{\mathrm{d}G_{ji}(\theta)}{\mathrm{d}\theta}w_i\Big)\Big].$$

It is possible to split the derivative of $\bar{V}_j(\theta)$ into $\mathfrak{n}+1$ separate components: the derivative w.r.t. $\left[\begin{smallmatrix} b_{ji} \\ a_{ji}\end{smallmatrix}\right]$, $i \in \mathcal{N}_j$ denoted $\frac{\mathrm{d}\bar{V}_j(\theta)}{\mathrm{d}\theta_{a_{ji},b_{ji}}}$ and the derivative w.r.t. $\left[\begin{smallmatrix} c_j \\ d_j\end{smallmatrix}\right]$, denoted $\frac{\mathrm{d}\bar{V}_j(\theta)}{\mathrm{d}\theta_{c_j,d_j}}$. This is due to the fact that all the transfer functions $G_{ji}(\theta)$, $i \in \mathcal{N}_j$ and $H_j(\theta)$ are all independently parameterized. The resulting $\mathfrak{n}+1$ equations are:

$$\frac{\mathrm{d}\bar{V}_j(\theta)}{\mathrm{d}\theta_{c_j,d_j}} = \bar{\mathbb{E}}\Big[\varepsilon_j(\theta)H_j^{-1}(\theta)\frac{\mathrm{d}H_j(\theta)}{\mathrm{d}\theta_{c_j,d_j}}\varepsilon_j(\theta)\Big], \tag{14a}$$
$$\frac{\mathrm{d}\bar{V}_j(\theta)}{\mathrm{d}\theta_{b_{ji},a_{ji}}} = \bar{\mathbb{E}}\Big[\varepsilon_j(\theta)H_j^{-1}(\theta)\frac{\mathrm{d}G_{ji}(\theta)}{\mathrm{d}\theta_{b_{ji},a_{ji}}}w_i\Big], \ i \in \mathcal{N}_j \tag{14b}$$

where,

$$\frac{\mathrm{d}H_j(\theta)^T}{\mathrm{d}\theta_{c_j,d_j}} = \Big[\frac{q^{-1}}{D_j(\theta)} \cdots \frac{q^{-n_c}}{D_j(\theta)} \ \frac{-H_j(\theta)q^{-1}}{D_j(\theta)} \cdots \frac{-H_j(\theta)q^{-n_d}}{D_j(\theta)}\Big] \tag{15a}$$
$$\frac{\mathrm{d}G_{ji}(\theta)^T}{\mathrm{d}\theta_{b_{ji},a_{ji}}} = \Big[\frac{1}{A_{ji}(\theta)} \cdots \frac{q^{-n_b}}{A_{ji}(\theta)} \ \frac{-G_{ji}(\theta)q^{-1}}{A_{ji}(\theta)} \cdots \frac{-G_{ji}(\theta)q^{-n_a}}{A_{ji}(\theta)}\Big] \tag{15b}$$

where $D_j(\theta) = 1 + d^j_1 q^{-1} + \cdots + d^j_{n_d} q^{-n_d}$, and $A_{ji}(\theta) = 1 + a^{ji}_1 q^{-1} + \cdots + a^{ji}_{n_a} q^{-n_a}$.

Next it will be shown that $\frac{\mathrm{d}\bar{V}_j(\theta)}{\mathrm{d}\theta}$ can equal 0 at $\theta = \theta_0$ only if every loop that passes through w_j has a delay. First notice that $\varepsilon_j(\theta_0) = e_j$. This means that (14a) reduces to:

$$\frac{\mathrm{d}\bar{V}_j(\theta_0)}{\mathrm{d}\theta_{c_j,d_j}} = \bar{\mathbb{E}}\Big[e_j H_j^{-1}(\theta_0)\frac{\mathrm{d}H_j(\theta_0)}{\mathrm{d}\theta_{c_j,d_j}}e_j\Big] = 0$$

where the expression equals zero due to the fact that each entry in $\frac{\mathrm{d}H_j(\theta)}{\mathrm{d}\theta_{c_j,d_j}}$ has a delay (see (15a)), and $\mathbb{E}[e_j(t)e_j(t-k)] = 0$, $k = 1, 2, \ldots$. Note that this equation holds irrespective of the delay structure of the network.

Next, consider what happens to (14b) at $\theta = \theta_0$:

$$\frac{\mathrm{d}\bar{V}_j(\theta_0)}{\mathrm{d}\theta_{b_{ji},a_{ji}}} = \bar{\mathbb{E}}\Big[e_j H_j^{-1}(\theta_0)\frac{\mathrm{d}G_{ji}(\theta_0)}{\mathrm{d}\theta_{b_{ji},a_{ji}}}w_i\Big]$$
$$= \bar{\mathbb{E}}\Big[e_j H_j^{-1}(\theta_0)\frac{\mathrm{d}G_{ji}(\theta_0)}{\mathrm{d}\theta_{b_{ji},a_{ji}}}\Big(\sum_{k\in\mathcal{V}_i}\mathcal{G}^0_{ik}v_k + \sum_{k\in\mathcal{R}_i}\mathcal{G}^0_{ik}r_k\Big)\Big]$$
$$= \bar{\mathbb{E}}\Big[e_j H_j^{-1}(\theta_0)\frac{\mathrm{d}G_{ji}(\theta_0)}{\mathrm{d}\theta_{b_{ji},a_{ji}}}\mathcal{G}^0_{ij}v_j\Big] \tag{16}$$

where the second equality holds by Lemma 1 and the definitions of $\mathcal{V}_i$ and $\mathcal{R}_i$ (as shown in (9)), and the third equality holds since e_j is uncorrelated to v_k, $k \in \mathcal{V}_j \setminus j$, and r_k, $k \in \mathcal{R}_j$ by Condition (c).

Consider the following useful facts about (16):

- The only way $\bar{\mathbb{E}}[e_j(t)e_j(t-k)]$ can equal zero is if $k \neq 0$.
- By Condition (f) it is known which transfer functions G^0_{ji} have delays. If G^0_{ji} has a delay, then $G_{ji}(\theta)$ will be parameterized with a delay, which means by (15b) that $\frac{\mathrm{d}G_{ji}(\theta_0)}{\mathrm{d}\theta_{b_{ji},a_{ji}}}$ will have a delay.
- Generically $\mathcal{G}^0_{ij}$ will equal 0 only if there is no path from i to j (generically the paths from i to j will not cancel each other out such that the sum of all paths from i to j is equal to zero). Similarly, generically, $\mathcal{G}^0_{ji}$ will have a delay only if all paths from i to j have a delay.

By the first point, the only way that (16) can equal zero is if $\frac{\mathrm{d}G_{ji}(\theta_0)}{\mathrm{d}\theta_{b_{ji},a_{ji}}}\mathcal{G}^0_{ij}$ is equal to zero, or has a delay. Generically, it will not equal zero or have a delay, unless every path from j to j contains a loop. ■

V. Dynamic Networks with Unknown Interconnection Structure

In Theorem 1 it was shown that the transfer functions G_{ji}, $i \in \mathcal{N}_j$ can be consistently estimated under certain conditions. However, the transfer functions G_{ji} $i \in \{1, \ldots, L\} \setminus \{j\}$ can also be consistently estimated (some of the estimated transfer functions are 0). In other words, the direct method can be used for a network with unknown interconnection structure.

Corollary 1: If the setup described in Theorem 1 is extended in order to estimate the transfers G^0_{ji}, $i \in \{1, \ldots, L\} \setminus \{j\}$, then the results remain basically invariant, except

(c) v_j must be uncorrelated to all external excitation sources and noise sources entering the network.

(d) The spectrum of $z := [r_j \ w_1 \ \ldots \ w_L]$, $\Phi_z(\omega)$ must be positive definite for $\omega \in [-\pi, \pi]$.

Some of the identified transfer functions will be 0. □

Note that the possibility of estimating a 0 transfer function leads to an identifiability issue since the parameters $a^{ji}_1, \ldots, a^{ji}_{n_a}$ in (3) become undetermined in such a case. However, the b^{ji}_k parameters are not undetermined, and will be equal to zero, resulting in a zero transfer no matter what the a^{ji}_k parameters are.

Corollary 1 is a generalization of current methods in the literature: currently, when identifying in an unknown interconnection structure the assumptions are that every transfer function has a delay, is stable and is driven only by a white or colored noise noise (see [5] for example). In Corollary 1, the delay condition has been relaxed and unstable transfers are no problem (using the method of [9]).

Remark 2: If the conditions of Corollary 1 are applied to each measured variable w_j $j \in \{1,\dots,L\}$ the result is that all noise sources must be uncorrelated to each other (i.e. H^0 must be diagonal).

Suppose that each module transfer is identified using the method presented in this paper (and the conditions of Corollary 1 are satisfied for every w_j, $j \in \{1,\dots,L\}$). By Remark 2, H^0 must be diagonal. For any dynamic network, there is only one representation such that H^0 is diagonal. However, there do exist other representations with non-diagonal noise models. See [10] for more in depth analysis. The following example illustrates the point.

Example 1: Consider a dynamic network described by the following equations:

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 0 & G^0_{12} & G^0_{13} \\ G^0_{21} & 0 & G^0_{23} \\ 0 & G^0_{32} & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} + \begin{bmatrix} H^0_1 & & \\ & H^0_2 & \\ & & H^0_3 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}.$$

The network can be represented graphically as shown in Fig. 1a. Note that H^0 is diagonal in this case. The following equations also describe the exact same dynamics between the measured variables (substitute the expression for w_3 into the expression for w_1):

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 0 & \tilde{G}^0_{12} & 0 \\ G^0_{21} & 0 & G^0_{23} \\ 0 & G^0_{32} & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} + \begin{bmatrix} H^0_1 & & G^0_{13}H^0_3 \\ & H^0_2 & \\ & & H^0_3 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}$$

which can be represented graphically as shown in Fig. 1b. In the figure, $\tilde{G}^0_{12} = G^0_{12} + G^0_{13}G^0_{32}$ and $\tilde{v}_1 = G^0_{13}v_3 + v_1$. It seems that the two networks have "different" interconnection structures. However, the conditions of Corollary 1 are not satisfied in the second network when attempting to identify $\tilde{G}^0_{12}$: $\tilde{v}_1$ is not uncorrelated to v_3. Therefore, this network must be transformed into the form of the first network, where all noise sources are uncorrelated, and then Corollary 1 applies to all measured variables in the network.

In conclusion, by Corollary 1 only the interconnection structure which corresponds to a diagonal H^0 matrix can be estimated (i.e. only one interconnection structure can be identified using the method presented in this paper).

VI. Conclusion

In this paper the classical Direct Method of closed loop identification has been generalized to include more complex network structures. Conditions have been presented under which consistent estimates can be obtained. It has also been shown that the method can be applied to a network with unknown interconnection architecture. This idea will be further investigated in the future.

References

[1] J. Willems, "Modelling interconnected systems," in *ISCCSP*, Malta, Mar. 2008, pp. 421–424.

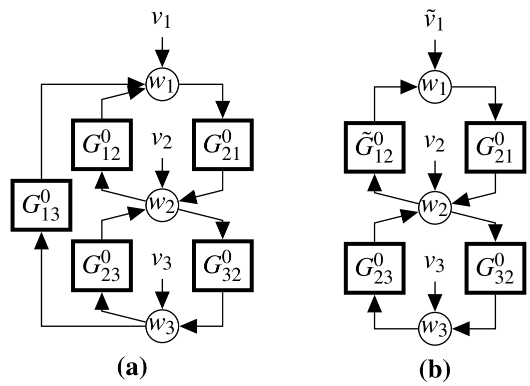

Fig. 1. Graphical representation of the two systems considered in Example 1. For notational convenience labels of the w_i's have been placed inside each summation, which denotes that the output of the sum is the measured variable w_i.

[2] P. M. J. Van den Hof, A. Dankers, P. S. C. Heuberger, and X. Bombois, "Identification in dynamic networks with known interconnection topology," in *Proceedings of 51 IEEE Conference on Decision and Control*, 2012, to appear.

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

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

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

[6] D. Materassi and G. Innocenti, "Topological identification in networks of dynamical systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1860–1871, 2010.

[7] A. Dankers, P. M. J. Van den Hof, P. S. C. Heuberger, and X. Bombois, "Dynamic network structure identification with prediction error methods - basic examples," in *Proceedings of 16th IFAC Symposium on System Identification*, Brussels, Belgium, Jul. 2012.

[8] P. M. J. Van den Hof, D. de Vries, and P. Schoen, "Delay structure conditions for identifiablility of closed loop systems," *Automatica*, vol. 28, no. 5, pp. 1047–1050, 1992.

[9] U. Forssell and L. Ljung, "Identification of unstable systems using output error and box-jenkins model structures," *IEEE Transactions on Automatic Control*, no. 1, 2000.

[10] J. Gonçalves and S. Warnick, "Necessary and sufficient conditions for dynamical structure reconstruction of lti networks," *IEEE Transactions on Automatic Control*, vol. 53, no. 7, pp. 1670–1674, Aug. 2008.

[11] M. Bona, *Combinatorics of Permutations*, ser. Discrete Mathematics and Its Applications. Chapman and Hall, 2004.

[12] G. Strang, *Linear Algebra and its Applications*, 4th ed. Thomson Brooks/Cole, 2006.

Appendix
Proof of Lemma 1

Before proving Lemma 1, some preliminary material on permutations is presented. Often permutations are thought of as ways of listing n objects so that each object is listed exactly once. However, they can also be thought of as functions [11].

For example, consider a permutation of 6 objects: 251634. The function f could be defined as $f(1) = 2$, $f(2) = 5$, $f(3) = 1$, $f(4) = 6$, $f(5) = 3$, and $f(6) = 4$. This leads to the so called 2 line notation of a permutation [11]:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 1 & 6 & 3 & 4 \end{pmatrix}$$

which emphasizes that f maps 1 to 2, 2 to 5, 3 to 1, etc.

Another interesting fact is revealed using this notation: no matter how many times f is applied, $f^m(4) = 4$ or $f^m(4) = 6$, and $f^m(4)$ will never have any other values. In other words, f cyclically permutes 4 and 6 and f cyclically permutes $1,2,5,3$. This leads to another notation for the permutation f; the so-called cycle notation. For the example: $f = (1253)(46)$. This idea will be used in the proof of Lemma 1.

Proof of Lemma 1: The proof will proceed as follows:

1) the (i,j)th entry of $(I-G)^{-1}$ is the jith cofactor of $(I-G)$,
2) the jith cofactor is a sum of products of permutations of the transfer functions G^0_{nm},
3) each product has a factor equal to a path from i to j,
4) moreover the factor multiplied by the path is monic (never has a delay)

Therefore this reasoning shows that the (i,j)th entry has a delay if and only if each path from i to j has a delay. By the same reasoning it follows that the (i,j)th entry is zero if and only if each path from i to j is zero.

The proof will focus on the first column of $(I-G)^{-1}$, but the reasoning applies to any column. For notational clarity the superscript 0 has been dropped on G and all G^0_{ji}.

The inverse of a nonsingular matrix is equal to:

$$(I-G)^{-1} = \frac{1}{\det(I-G)}\begin{bmatrix} C_{11} & \cdots & C_{L1} \\ \vdots & & \vdots \\ C_{1L} & \cdots & C_{LL}\end{bmatrix},$$

where C_{ij} is the ijth cofactor of $(I-G)$ [12]. Therefore, the $(i,1)$th entry, $i \neq 1$ of $(I-G)^{-1}$ can be permuted by row operations into the form:

$$\pm\begin{vmatrix} G_{i,1} & G_{i,2} & \cdots & G_{i,i-1} & G_{i,i+1} & \cdots & G_{i,L} \\ G_{21} & -1 & & G_{2,i-1} & G_{2,i+1} & \cdots & G_{2L} \\ \vdots & & \ddots & \vdots & \vdots & & \vdots \\ G_{i-1,1} & G_{i-1,2} & \cdots & -1 & G_{i-1,i+1} & \cdots & G_{i-1,L} \\ G_{i+1,1} & G_{i+1,2} & \cdots & G_{i+1,i-1} & -1 & & G_{i+1,L} \\ \vdots & \vdots & & \vdots & & \ddots & \\ G_{L1} & G_{L2} & \cdots & G_{L,i-1} & G_{L,i+1} & & -1 \end{vmatrix} \quad (17)$$

The $\pm$ signs are not important, since they will be absorbed into the factor which multiplies the path. To find an expression for C_{1i} the above determinant will be evaluated.

For notational convenience, denote (17) as $|A|$, and let $[[A]]_{ij} = a_{ij}$. The Leibniz formula for $|A|$ is [12]:

$$\det A = \sum_{\sigma\in S_{L-1}} \operatorname{sgn}(\sigma) a_{1\sigma_1} a_{2\sigma_2} \cdots a_{L-1,\sigma_{L-1}}, \quad (18)$$

where σ is a permutation of $\{1,\ldots,L-1\}$, S_{L-1} is the set of all permutations of $\{1,\ldots,L-1\}$, σ_i denotes the ith element of the permutation σ, and $\operatorname{sgn}(\sigma)$ is -1 if the permutation is odd, and 1 if the permutation is even. As mentioned, the sign is not important.

In the following it will be shown that each product in (18) has a factor which is equal to a path from i to j.

Recall the two line notation of a permutation:

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & L-1 \\ \sigma_1 & \sigma_2 & \sigma_3 & \cdots & \sigma_{L-1}\end{pmatrix} \quad (19)$$

Note each column in (19) corresponds to the subscripts of the term a_{k,σ_k} in the product in (18). Moreover, (19) can be written in cycle notation as $(1n_1n_2\cdots n_k)(*\cdots*)$ where attention is focused on the cycle that contains 1. Rewriting in two line notation:

$$\begin{pmatrix} 1 & n_1 & n_2 & \cdots & n_{\ell-1} & m_1 & \cdots & m_{L-\ell} \\ n_1 & n_2 & n_3 & \cdots & 1 & * & \cdots & * \end{pmatrix} \quad (20)$$

where $\{1,n_1,\ldots,n_{\ell-1}\}$ and $\{m_1,\ldots,m_{L-\ell}\}$ are mutually exclusive subsets of $\{1,\ldots,L-1\}$ such that $\{1,n_1,\ldots,n_{\ell-1}\}\cup\{m_1,\ldots,m_{L-\ell}\} = \{1,\ldots,L-1\}$ and each element in $\{1,n_1,\ldots,n_{\ell-1}\}$ is unique. Note (20) is the same as (19) except that columns have been switched, and labeling has changed. Using the new labels the product in (18) is:

$$a_{1n_1}a_{n_1n_2}\cdots a_{n_{\ell-2}n_{\ell-1}}a_{n_{\ell-1}1}\cdot\mathcal{A}(m_1,\ldots,m_{L-\ell}) \quad (21)$$

where $\mathcal{A}(m_1,\ldots,m_{L-\ell})$ denotes the product of the remaining $a_{m_*m_*}$ that are not part of the cycle containing 1. Now the terms a_{ij} must be converted back into G_{**} notation.

Recall that a_{11} denotes the $(1,1)$ entry of C_{1i}, which is equal to G_{i1} (see (17). From the expression of C_{1i}:

$$a_{1n_1} = \begin{cases} G_{in_1} & \text{if } n_1 < i \\ G_{i,n_1+1} & \text{if } n_1 \geq i\end{cases} \quad (22a)$$

$$a_{n_an_b} = \begin{cases} G_{n_an_b} & \text{if } n_a < i, n_b < i \\ G_{n_a,n_b+1} & \text{if } n_a < i, n_b \geq i \\ G_{n_a+1,n_b} & \text{if } n_a \geq i, n_b < i \\ G_{n_a+1,n_b+1} & \text{if } n_a \geq i, n_b \geq i\end{cases} \quad (22b)$$

Re-label the set $\{n_1,\ldots,n_k\}$ as $\{\delta_1,\ldots,\delta_k\}$, where $\delta_\ell = n_\ell$ if $n_\ell < i$, and $\delta_\ell = n_\ell + 1$ if $n_\ell \geq i$.

Using the δ labels in (21) and using (22a) - (22b) to replace the a_{ij} notation results in:

$$\begin{aligned} &a_{1n_1}a_{n_1n_2}\cdots a_{n_{\ell-1}1}\cdot\mathcal{A}(m_1,\ldots,m_{L-\ell-1}) \\ &\qquad = G_{i\delta_1}G_{\delta_1\delta_2}\cdots G_{\delta_k1}\cdot\mathcal{A}(m_1,\ldots,m_{L-\ell-1}). \quad (23)\end{aligned}$$

where $G_{i\delta_1}\cdots G_{\delta_k1}$ is a path from i to 1. Thus far it has been shown that every product in the sum (18) has the form (23). Subsequently, the $(i,1)$th entry of $(I-G)^{-1}$ is a sum of terms of the form

$$G_{i\delta_1}G_{\delta_1\delta_2}\cdots G_{\delta_k1}\frac{\mathcal{A}(m_1,\ldots,m_{L-\ell-1})}{\det(I-G)}.$$

All terms containing the same first factor can be grouped together. Then it must be shown that

$$\sum_{\sigma\in S_{L-\ell-1}}\frac{\mathcal{A}(\sigma_1,\ldots,\sigma_{L-\ell-1})}{\det(I-G)} \quad (24)$$

is always monic, where $S_{L-\ell-1}$ is the set of all permutations of $\{m_1,\ldots,m_{L-\ell-1}\}$. Note that for $\sigma = m,\ldots,L-\ell-1$, it holds that $\mathcal{A} = a_{m_1m_1}\cdots a_{m_{L-\ell}m_{L-\ell}} = 1$. Since $(I-G)$ has ones on the diagonal, by (18) the determinant will always have a term equal to 1. In conclusion (24) is monic. By the reasoning at the start of the proof, the Lemma is proved. □