Local identification in dynamic networks using a multi-step least squares method

Stefanie J.M. Fonken, Karthik R. Ramaswamy and Paul M.J. Van den Hof

***Abstract*— For identification of a single module in a linear dynamic network with correlated disturbances different methods are available in a prediction error setting. While indirect methods fully rely on the presence of a sufficient number of external excitation signals for achieving data-informativity, the local direct method with a MIMO predictor model can exploit also non-measured disturbance signals for data-informativity. However a simple two-node example shows that this local direct method can also be conservative in terms of the number of external excitation signals that is required. Inspired by a recently introduced multi-step method for full network identification, we present a multi-step least squares method for single module identification. In a first indirect step a model is estimated that is used to reconstruct the innovation on a set of output signals, which in a second step is used to directly estimate the module dynamics with a MISO predictor model. The resulting path-based conditions for data-informativity show that the multi-step method requires a smaller number of excitation signals for data-informativity than the local direct method.**

I. INTRODUCTION

Modern day dynamic systems consist of interconnected subsystems that can be represented as large-scale networks. In the domain of system identification, data-driven modeling suitable for these dynamic networks has been the subject of many studies, in among others [1], [2], [3], [4], [5], [6].

In the problem of local identification the objective is to consistently identify a single module using local measured data within a network of which the interconnection structure is given. Current available local identification methods in a prediction error setting are typically categorized under indirect methods and direct methods.

Indirect methods [7], [3], [8] typically estimate the tranfers from measured (user manipulated) external signals to the internal variables which are the nodes w. These methods therefore rely on the presences of a sufficient number of measured excitation signals to achieve data-informativity. Additionally, indirect methods require post processing to obtain the estimates of the target module. Local direct methods [9], [10] directly estimate the transfers between the node signals, where the node signals can receive excitation from either measured (user manipulated) or unmeasured (disturbances) external signals to achieve data-informativity. The local direct method therefore achieves a consistent estimate with maximum likelihood properties under the condition that an appropriate disturbance model is estimated. In the presence of correlated disturbances the local direct method might require the use of a multiple-input-multiple-output (MIMO) predictor model to consistently estimate with maximum likelihood properties. This is opposed to the indirect methods that achieve consistency without maximum likelihood properties, but do not require accurate modeling of the disturbances.

Consider the 2-node example shown in Figure 1, with target module G_{21} that represents the interconnecting dynamics from node signal w_1 to w_2, and where the disturbances are represented as filtered white noise signals e_1 and e_2. Moreover, the external excitation signals u_1 and u_2 are measured and might be present or not. For the situation there are no correlated disturbances ($H_{12} = H_{21} = 0$), the example is a classic closed-loop problem for which we can use classic identification methods [11]. In the presence of correlated disturbances the single module identifiability check [12], which is a method independent check, indicates that target module G_{21} is identifiable for the situation either u_1 or u_2 is present. For the 2-node example, the indirect method estimates target module G_{21} by identifying the mapping from $u_1 \rightarrow w_1$ and $u_1 \rightarrow w_2$ and taking the quotient. The indirect method therefore relies on the presence of u_1, whether there are correlated disturbances present or not. The local direct method [9] requires a MIMO predictor model with both w_1 and w_2 as outputs to appropriately model the disturbances. However, as shown in [13], [14], [15] this predictor model requires both u_1 and u_2 to be present to achieve data-informativity. Apparently in this example the data-informativity conditions for the local direct method are conservative when compared to the single module identifiability results and the indirect method.

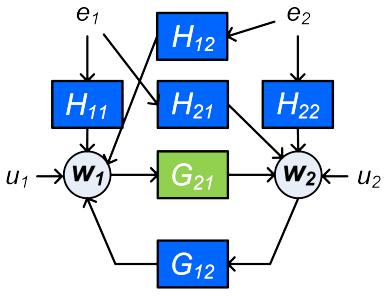

Fig. 1. 2-node example for local identification of target module G_{21}

This project has received funding from the European Research Council (ERC), Advanced Research Grant SYSDYNET, under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 694504).

The authors are with the Control Systems Group, Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands. Email: {s.j.m.fonken,p.m.j.vandenhof}@tue.nl, k.r.ramaswamy89@gmail.com.

Manuscript 562 submitted to 2023 62nd IEEE Conference on Decision and Control (CDC) . Received March 31, 2023.

In this paper, our aim is to develop a local identification method that has the key advantages of the local direct method but requires less external excitation signals u compared to the local direct method.

The multi-step least squares method [16], developed for full network identification, relies on accurate modeling of the disturbances and directly estimates the target module, similar to the local direct method. Yet, the multi-step least squares methods can identify the full 2-node network in the example while requiring the presence of either u_1 or u_2 to achieve data-informativity. Nevertheless, there is no local identification version of this multi-step method available. In this paper we will develop such a local identification method, and show that it indeed allows for a smaller number of excitation signals compared to the local direct method.

Section 2 presents the network setup and the results of [9], [15] and [16]. Section 3 describes the layout of the algorithm for local identification, where the 3 main steps of the algorithm are elaborated in Sections 4 to 6. Section 7 provides the data-informativity conditions of the local identification algorithm. The 2-node example in Figure 1 is then worked out in Section 8, and the results are concluded in Section 9.

II. DYNAMIC NETWORK SETUP AND CURRENT METHODS

A. Dynamic network setup

We follow the module framework setup in [2], and consider dynamic networks represented by

$$w(t) = G(q)w(t) + H(q)e(t) + u(t), \tag{1}$$

where q^{-1} is the delay operator, i.e. $q^{-1}w(t) = w(t-1)$, w is a L dimensional vector, $G(q)$ a hollow and rational transfer function matrix and $H(q)$ is a rational disturbance model. Moreover, $e(t)$ is $p \leq L$ dimensional vector containing white noise processes. The inputs signals $u(t) = R(q)r(t)$ are driven by the measured external excitation signals $r(t)$, with $R(q)$ a known rational transfer function matrix. We assume that the network is well-posed, and for simplicity we assume all elements in $G(q)$ strictly proper rational transfer functions. The set of node indexes $\{1 \cdots L\}$ is denoted by $\mathcal{L}$.

B. Local direct method

The local direct method for single module identification [9] is a prediction error method that is built on a representation of a subnetwork through a predictor model

$$w_{\mathcal{Y}} = \bar{G}(q)w_{\mathcal{D}} + \bar{H}(q)\xi_{\mathcal{Y}} + u_{\mathcal{Y}} \tag{2}$$

focusing on identifying the target module G_{ji}, with $i \in \mathcal{D}$ and $j \in \mathcal{Y}$. The sets $\mathcal{Y}$ and $\mathcal{D}$ are chosen in such a way that the target module remains invariant in the predictor model, i.e. $\bar{G}_{ji}(q) = G_{ji}(q)$, while confounding variables[1] are treated in such a way that $\bar{G}_{ji}(q)$ can be estimated consistently. For a proper treatment of confounding variables, a possibly multivariate noise model $\bar{H}(q)$ is exploited through which correlations between disturbances on nodes in $w_{\mathcal{Y}}$ and $w_{\mathcal{D}}$ can be explicitly modelled. Note that it is allowed that $\mathcal{Y} \cap \mathcal{D} \neq \emptyset$.

While $u_{\mathcal{Y}}$ can be decomposed as $u_{\mathcal{Y}} = \bar{J}(q)u_{\mathcal{K}} + \bar{S}u_{\mathcal{P}}$ with $\bar{J}(q)$ dynamic and $\bar{S}$ a binary matrix, data informativity is achieved if the vector signal $\kappa := [w_{\mathcal{D}}^T \ \xi_{\mathcal{Y}}^T \ u_{\mathcal{K}}^T]^T$ is persistently exciting, typically achieved by requiring that $\Phi_\kappa(\omega) \succ 0$ for almost all ω. The data-informativity condition can be satisfied generically by satisfying path-based conditions on the graph of the network [15].

[1] A confounding variable is an unmeasured variable that has paths to both the input and the output of an estimation problem [17].

C. Multi-step least squares

The full network identification method in [16] first models the mapping $u \to w$ on the basis of a high-order ARX model. This ARX model is then used to reconstruct the innovation signals in vector e. Next the reconstructed innovation $\hat{e}$ can be treated as a measured input according to

$$w = Gw + (H - I)\hat{e} + Ie + u, \tag{3}$$

for a monic H, that now represents a network with a noise model equal to I, implying that there are no correlated disturbances present. Identification of the full network can therefore be done using MISO predictors for each row $j \in \mathcal{L}$ of the network. These MISO predictor models directly estimate the inneighboring modules of output node w_j and the disturbances that have a direct path to w_j. In contrast to the local direct method, which models the disturbance using unmeasured innovation signals, the multi-step method treats the reconstructed innovation signals as measured signals in the modeling procedure.

III. LAYOUT OF THE ALGORITHM

We will introduce the following stepwise algorithm that is actually a local version of the network identification algorithm presented in [16].

Step 1:

a. Select a subset of nodes $\mathcal{S} \subset \mathcal{L}$, with $i, j \in \mathcal{S}$, on the basis of which we are going to estimate the dynamics of the single module G_{ji};
b. Remove all non-selected nodes from the network by immersion [18], i.e. by adjusting the network representation such that the node signals $w_{\mathcal{S}}$ remain invariant.
c. Decompose the set $\mathcal{S}$ in $\mathcal{Y}$ and $\mathcal{U}$, with $j \in \mathcal{Y}$, in such a way that there are no confounding variables for the estimation problem $w_{\mathcal{U}} \to w_{\mathcal{Y}}$ [19], [9]. This can always be done, if necessary by choosing $\mathcal{U} = \emptyset$.

Step 2:

- In the resulting subnetwork description

$$\underbrace{\begin{bmatrix} w_{\mathcal{Y}} \\ w_{\mathcal{U}} \end{bmatrix}}_{w_{\mathcal{S}}} = \bar{G}_{\mathcal{S}}(q) \begin{bmatrix} w_{\mathcal{Y}} \\ w_{\mathcal{U}} \end{bmatrix} + \begin{bmatrix} \bar{H}(q) & 0 \\ 0 & \bar{H}_{\mathcal{U}}(q) \end{bmatrix} \begin{bmatrix} \xi_{\mathcal{Y}} \\ \xi_{\mathcal{U}} \end{bmatrix} + u_{\mathcal{S}}$$

estimate a high order predictor model for the node signals $w_{\mathcal{Y}}$ in the mapping $u_{\mathcal{S}} \to w_{\mathcal{S}}$, and use that model to estimate the innovation signal $\xi_{\mathcal{Y}}$.

Step 3:

- Estimate a MISO model for the node signal w_j, using the estimated innovation signal $\hat{\xi}_{\mathcal{Y}}$ as a measured input, next to the relevant input signals that are present in $w_{\mathcal{S}}$.

In the following Sections we will present and analyse the separate steps in the Algorithm in more detail, and pay particular attention to the data-informativity conditions.

IV. STEP 1: NETWORK EQUATIONS FOR A SUBNETWORK

A. Subnetwork selection and immersion

In the problem of local identification we can use the topology information to decide which node signals $\mathcal{S} \subset \mathcal{L}$ to select for the identification procedure. The subnetwork with selected nodes is then obtained by a dynamic network operation that appropriately removes the unselected nodes while keeping the selected node signals invariant, referred to as immersion [18]. The subnetwork is then represented by the immersed network according to

$$w_{\mathcal{S}} = \bar{G}_{\mathcal{S}} w_{\mathcal{S}} + \bar{H}_{\mathcal{S}} \xi_{\mathcal{S}} + u_{\mathcal{S}}, \tag{4}$$

where $w_{\mathcal{S}}$ is invariant, $\bar{G}_{\mathcal{S}}$ represents the immersed module dynamics, $u_{\mathcal{S}}$ contains the immersed external excitation signals according $u_{\mathcal{S}} = \bar{R}(q)u_{\mathcal{L}}$ with $\bar{R}(q) \in \mathbb{R}(z)^{dim(\mathcal{S})\times L}$, and where the disturbance process $\bar{H}_{\mathcal{S}}\xi_{\mathcal{S}}$ is the result of a spectral factorization applied to the immersed disturbance model, with $\xi_{\mathcal{S}}$ a vector of white noise processes. It is assumed that the disturbance process $\bar{H}_{\mathcal{S}}\xi_{\mathcal{S}}$ is full rank.

The use of an immersed network allows local identification methods to use local network measurements, where the selected node signals should satisfy the following condition.

Condition 1. Consider a target module G_{ji}. A subset of node signals $\mathcal{S} \subset \mathcal{L}$ with $\{w_i, w_j\} \in w_{\mathcal{S}}$ is said to satisfy the Parallel Path and Loop (PPL) condition, if every path from w_i to w_j that does not pass through G_{ji}, and every loop around w_j, pass through a node in $\mathcal{S}$.

The condition stems from [18], where it has been shown that it guarantees that after immersion of the nodes in $\mathcal{L}\backslash\mathcal{S}$, leading to the system representation (4) the module dynamics of G_{ji} remains invariant in the representation of the immersed network.

Any subnetwork $\mathcal{S}$ that satisfies Condition 1 will allow for the identification of a consistent estimate of the target module. The particular choice of $\mathcal{S}$ will however affect the conditions on data-informativity, that are intrinsically present in the consistency results. For the continuation of the algorithm we assume that the subset $\mathcal{S}$ is selected by the user.

B. Decomposition of the subnetwork

Inspired by the work in [19] and [9] we decompose set $\mathcal{S}$ in subsets $\mathcal{Y}$ and $\mathcal{U}$, for which there is no innovation signal $\xi_k, k \in \mathcal{S}$ that has direct paths to both $\mathcal{Y}$ and $\mathcal{U}$, meaning there are no confounding variable for the estimation problem $u_{\mathcal{U}} \to w_{\mathcal{Y}}$. Then the disturbance model of the subnetwork adheres to a block diagonal structure written as

$$\underbrace{\begin{bmatrix} w_{\mathcal{Y}} \\ u_{\mathcal{U}} \end{bmatrix}}_{w_{\mathcal{S}}} = \underbrace{\begin{bmatrix} \bar{G}_{\mathcal{Y}\mathcal{Y}} & \bar{G}_{\mathcal{Y}\mathcal{U}} \\ \bar{G}_{\mathcal{U}\mathcal{Y}} & \bar{G}_{\mathcal{U}\mathcal{U}} \end{bmatrix}}_{\bar{G}_{\mathcal{S}}} \underbrace{\begin{bmatrix} w_{\mathcal{Y}} \\ u_{\mathcal{U}} \end{bmatrix}}_{w_{\mathcal{S}}} + \underbrace{\begin{bmatrix} \bar{H} & 0 \\ 0 & \bar{H}_{\mathcal{U}} \end{bmatrix}}_{\bar{H}_{\mathcal{S}}} \underbrace{\begin{bmatrix} \xi_{\mathcal{Y}} \\ \xi_{\mathcal{U}} \end{bmatrix}}_{\xi_{\mathcal{S}}} + \underbrace{\begin{bmatrix} u_{\mathcal{Y}} \\ u_{\mathcal{U}} \end{bmatrix}}_{u_{\mathcal{S}}}, \tag{5}$$

with $\xi_{\mathcal{S}}$ a vector of white noise processes, $\bar{H}$ is monic, stable and stably invertible, and $\bar{G}_{\mathcal{Y}\mathcal{Y}}$ is hollow. The target module G_{ji} remains invariant during the decomposition, as formulated in the following result.

Proposition 1. For every set $\mathcal{S}$ that satisfies the PPL condition in Condition 1 there exists a decomposition in sets $\mathcal{Y}$ and $\mathcal{U}$, with $j \in \mathcal{Y}$, for which there exists a structured form (5) that has no confounding variables for the estimation problem $u_{\mathcal{U}} \to w_{\mathcal{Y}}$ and $\bar{G}_{ji} = G_{ji}$.

Proof: The Proposition is a special case of Theorem 1 in [9], for the situation $\mathcal{Q} = \mathcal{Y}$, $\mathcal{O} = \mathcal{B} = \emptyset$, where the work in [9] shows that a decomposition in a structured form (5) can always be found for a selected subnetwork. □

As an immediate result of Proposition 1 the expression for nodes $w_{\mathcal{Y}}$ equals the first block row of (5), which includes the target output w_j. The external excitation signals $u_{\mathcal{Y}}$ that enter nodes $w_{\mathcal{Y}}$ require further specification.

Depending on the selected nodes in the subnetwork $w_{\mathcal{S}}$, the immersion operation leads to unknown dynamics appearing, for example for the path $u_k, k \in \mathcal{L}\backslash S$ to $w_{\mathcal{Y}}$. Then the external excitation signals enter the nodes $w_{\mathcal{Y}}$ with either an unknown dynamic term or a known constant term, as formulated next.

Proposition 2. The signals $u_{\mathcal{Y}}$ in (5), can be written as

$$u_{\mathcal{Y}} = \bar{J}(q)u_{\mathcal{K}} + \bar{S}u_{\mathcal{P}}, \tag{6}$$

with $\bar{J}(q)$ a dynamic transfer function and $\bar{S}$ a binary constant matrix and where $u_{\mathcal{K}}$ and $u_{\mathcal{P}}$ are defined as follows.

1) $\mathcal{P} \subset \mathcal{Y}$ and for $\ell \in \mathcal{Y}$, $u_\ell \in u_{\mathcal{P}}$ if all loops around w_ℓ pass through a node in $\mathcal{Y} \cup \mathcal{U}$.
2) The set of excitation signals $u_{\mathcal{K}}$ is composed of
   - all signals $u_{\mathcal{Y}}$ that are not in $u_{\mathcal{P}}$, and
   - all signals $u_{\mathcal{L}\backslash\mathcal{S}}$ of which the effect on $w_{\mathcal{Y}}$ is not covered by the inputs $u_{\mathcal{U}}$.

Proof: The Proposition is a special case of Proposition 3 in [15], for the situation $\mathcal{Q} = \mathcal{Y}$, $\mathcal{O} = \mathcal{B} = \emptyset$. □

Using the results in Proposition 1 and 2 we arrive at an expression for the nodes $w_{\mathcal{Y}}$

$$w_{\mathcal{Y}} = \bar{G}(q) \begin{bmatrix} w_{\mathcal{Y}} \\ u_{\mathcal{U}} \end{bmatrix} + \bar{H}(q)\xi_{\mathcal{Y}} + \bar{J}(q)u_{\mathcal{K}} + \bar{S}u_{\mathcal{P}}, \tag{7}$$

where in the algorithm's next step we show how we use predictor models of node signals $w_{\mathcal{Y}}$ to reconstruct the innovation signals $\xi_{\mathcal{Y}}$.

V. Step 2: High order ARX model to reconstruct the innovation signal $\xi_{\mathcal{Y}}$

For obtaining a reconstruction of the innovation signals we model the dynamics present in the mapping $u_{\mathcal{S}} \to w_{\mathcal{S}}$ and use this model to reconstruct the innovation signal $\xi_{\mathcal{Y}}$.

Analogous to the full network identification method in Section 2, we first consider the closed-loop network equations that represent the mapping from $u_{\mathcal{S}} \to w_{\mathcal{S}}$ based on the subnetwork equation (4), omitting (q) for convenience

$$w_{\mathcal{S}} = \Big(I - \bar{H}_{\mathcal{S}}^{-1}(q)\big(I - \bar{G}_{\mathcal{S}}(q)\big)\Big)w_{\mathcal{S}} + \bar{H}_{\mathcal{S}}^{-1}(q)u_{\mathcal{S}} + \xi_{\mathcal{S}}. \quad (8)$$

Based on the mapping (8) we construct an expression for the one step ahead predictor

$$\hat{w}_{\mathcal{S}}(t|t-1) = \Big(I - \bar{H}_{\mathcal{S}}^{-1}(q)\big(I - \bar{G}_{\mathcal{S}}(q)\big)\Big)w_{\mathcal{S}} + \bar{H}_{\mathcal{S}}^{-1}(q)u_{\mathcal{S}}. \quad (9)$$

Then due to the block-diagonal structure in $\bar{H}_{\mathcal{S}}$ (5), we can isolate the $\hat{w}_{\mathcal{Y}}(t|t-1)$ part of the predictor (9), omitting q for convenience

$$\hat{w}_{\mathcal{Y}}(t|t-1) = \big(I - \underbrace{\begin{bmatrix}\bar{H}^{-1} & 0\end{bmatrix}(I - \bar{G}_{\mathcal{S}})}_{A(q)}\big)w_{\mathcal{S}} + \underbrace{\bar{H}^{-1}\begin{bmatrix}\bar{S} & \bar{J}\end{bmatrix}}_{B(q)}\begin{bmatrix}u_{\mathcal{P}}\\ u_{\mathcal{K}}\end{bmatrix}, \quad (10)$$

where the decomposition of $u_{\mathcal{Y}}$ according to (6) has been substituted. We use a high order ARX model, where the parametrization of the filters $A(q)$ and $B(q)$ is chosen according

$$\begin{aligned} A(q,\zeta) &= I + A_1 q^{-1} + \cdots + A_n q^{-n},\\ B(q,\zeta) &= B_0 + B_1 q^{-1} + \cdots + B_{n-1} q^{-(n-1)}, \end{aligned} \quad (11)$$

which means that the predictor model (10) accurately approximates the rational filters $A(q)$ and $B(q)$ with polynomial functions (11) if the model order n is chosen sufficiently high [20]. The ARX model is estimated according to $\hat{\zeta}_N^n = \text{argmin}_\zeta \frac{1}{N}\sum_{t-1}^{N} \varepsilon_{\mathcal{Y}}^\top(t,\zeta)\varepsilon_{\mathcal{Y}}(t,\zeta)$, with data length N and

$$\varepsilon_{\mathcal{Y}}(t,\zeta) = w_{\mathcal{Y}}(t) - \hat{w}_{\mathcal{Y}}(t|t-1,\zeta) = w_{\mathcal{Y}}(t) - \varphi(t)\zeta \quad (12)$$

and $\varphi(t)$ composed of the appropriate terms in $w_{\mathcal{S}}, u_{\mathcal{P}}$ and $u_{\mathcal{K}}$, which leads to the analytical solution

$$\hat{\zeta}_N^n = \Big[\frac{1}{N}\sum_{t-1}^{N}\varphi(t)\varphi^\top(t)\Big]^{-1}\frac{1}{N}\sum_{t-1}^{N}\varphi(t)w_{\mathcal{Y}}(t). \quad (13)$$

We will show that estimates $\hat{\zeta}_N^n$ are consistent.

Proposition 3. Consider an subnetwork (5) that satisfies Condition 1, and consider predictor model (10) Then, with analytical solution (13), the estimates $A(\hat{\zeta}_N^n)$ and $B(\hat{\zeta}_N^n)$ are consistent if the following conditions hold

1) The external excitation $u_{\mathcal{P}\cup\mathcal{K}}$ is uncorrelated to the noise $\xi_{\mathcal{Y}}$
2) The spectral density $\Phi_\kappa(\omega) \succ 0$ for almost all ω, with $\kappa = \begin{bmatrix} w_{\mathcal{S}} \\ u_{\mathcal{P}\cup\mathcal{K}}\end{bmatrix}$ (data-informativity condition)
3) $A(q,\zeta)$ and $B(q,\zeta)$ are of high-order, such that $n \to \infty$
4) There exists a parameter ζ^0 such that $\big(A(\zeta^0), B(\zeta^0)\big)$ represents the dynamics in the closed-loop mapping $u_{\mathcal{S}} \to w_{\mathcal{S}}$ in (8).

Proof: The Proposition is special case of Proposition 1 in [16], for the situation that $w_{\mathcal{S}}$ has full rank disturbances, for which we only need predictor models of node signals $w_{\mathcal{Y}}$ based on signals $w_{\mathcal{S}}$ and $u_{\mathcal{K}\cup\mathcal{P}}$. □

The consistency property of estimate $\hat{\zeta}_N^n$ implies that

$$\varepsilon_{\mathcal{Y}}(t,\hat{\zeta}_N^n) \to \xi_{\mathcal{Y}}(t) \quad \text{w.p. 1 as } N \to \infty \ \forall\, t. \quad (14)$$

Therefore, the reconstructed innovation $\varepsilon_{\mathcal{Y}}(t,\hat{\zeta}_N^n) = w_{\mathcal{Y}} - \hat{w}_{\mathcal{Y}}(t|t-1,\hat{\zeta}_N^n)$ is a consistent estimate of $\xi_{\mathcal{Y}}(t)$.

VI. Step 3: Parametric target module estimate

By using the reconstructed innovation signal $\varepsilon_{\mathcal{Y}}(\hat{\zeta}_N^n)$ as a measured input, we rewrite the term $\bar{H}(q)\xi_{\mathcal{Y}}$ in (7) as

$$\bar{H}(q)\xi_{\mathcal{Y}} = (\bar{H}(q) - I)\hat{\varepsilon}_{\mathcal{Y}} + I\xi_{\mathcal{Y}},$$

with $\hat{\varepsilon}_{\mathcal{Y}} := \varepsilon_{\mathcal{Y}}(\hat{\zeta}_N^n)$, similar to the full network situation in Section 2.C. Therefore, for the identification of the target module we only consider the MISO predictor model for the $j-th$ row of (7)

$$\begin{aligned} \hat{w}_j(t|t-1) = &\sum_{k\in\mathcal{N}_j^-}\bar{G}_{jk}(\eta)w_k + \sum_{\ell\in\mathcal{Y}}\check{H}_{j\ell}(\eta)\hat{\varepsilon}_\ell \\ &+ \sum_{\gamma\in\mathcal{K}_j}\bar{J}_{j\gamma}(\eta)u_\gamma + \sum_{\beta\in\mathcal{P}_j}\bar{S}_{j\beta}u_\beta, \end{aligned} \quad (15)$$

with $\check{H}_{j\ell}$ an element of $\check{H} = \bar{H} - I$, $\mathcal{N}_j^-$ are the inneighbors of w_j where $\mathcal{N}_j^-$ is a subset of $\mathcal{U}\cup\mathcal{Y}\backslash j$, and $\mathcal{K}_j$ and $\mathcal{P}_j$ are subsets of $\mathcal{K}$ and $\mathcal{P}$ respectively.

The model is estimated according to

$$\hat{\eta}_N^n = \text{argmin}_\eta \frac{1}{N}\sum_{t-1}^{N}\varepsilon_j^\top(t,\eta)\varepsilon_j(t,\eta), \quad (16)$$

with $\varepsilon_j(t,\eta) = w_j(t) - \hat{w}_j(t|t-1,\eta)$. The conditions for consistency of estimates $\hat{\eta}_N^n$ are formulated next.

Proposition 4. Consider a dynamic network that satisfies Condition 1 and the one-step ahead predictor (15), for which a consistent estimate $\hat{\varepsilon}_{\mathcal{Y}}$ is available and the external excitation signals are defined according to Proposition 2. Then, with criterion (16) the estimates $\bar{G}_{jk}(\hat{\eta}_N^n)$, $\bar{H}_{j\ell}(\hat{\eta}_N^n)$ and $\bar{J}_{j\gamma}(\hat{\eta}_N^n)$ are consistent if the following conditions hold

1) The spectral density $\Phi_{\bar{\kappa}}(\omega) \succ 0$ for almost all ω, with $\bar{\kappa} = \begin{bmatrix} w_{\mathcal{N}_j^-} \\ u_{\mathcal{K}_j} \\ \xi_{\mathcal{Y}}\end{bmatrix}$ (data-informativity condition).
2) The data generating system is in the model set, i.e. there exists a η^0 such that $\bar{G}_{jk}(q,\eta^0) = \bar{G}_{jk}(q)$, $\check{H}_{j\ell}(q,\eta^0) = \bar{H}_{j\ell}(q) - I_{jj}$ and $\bar{J}_{j\gamma}(q,\eta^0) = \bar{J}_{j\gamma}(q)$.

Proof: The Proposition is a special case of Proposition 2 in [16], for the situation that $w_{\mathcal{S}}$ has full rank disturbances. The proof is structured similarly to the proof in [16] with some slight adaptations as given in the Appendix.

Note that Proposition 4 holds for all model structures that satisfy the system in the model set condition.

It appears that the local multi-step least squares method results in a MISO predictor model of the $j-th$ row to consistently estimate the target module, given that the required reconstructed innovation signals are consistently estimated in the previous step.

In order to continue with a purely least squares based estimation technique, we use the Weighted Null Space Fitting method [21], [22] to estimate the rational transfer functions in (15) as was done in [16]. Since the result in Proposition 4 is not limited to a least squares type of estimator, one could alternatively use a kernel-based identification method [23] to identify predictor model (15). Both methods have convex optimization properties that have a low computational complexity.

In the next Section we will further analyse the spectral data-informativity conditions in Propositions 3 and 4.

VII. PATH-BASED CONDITIONS FOR DATA-INFORMATIVITY

The spectral conditions in Proposition 3 and 4 that reflect data-informativity, can generically be satisfied by verifying path-based conditions on the graph of the network. This has been shown in [14], [15], based on the graph-based results in [3]. The resulting path-based conditions for the two Propositions are formulated next, for the mapping from persistently excited external signals to vectors κ and $\bar{\kappa}$.

Proposition 5. The spectrum condition in Proposition 3 is generically satisfied if there are $dim(w_{\mathcal{S}})$ vertex-disjoint paths from $(\xi_{\mathcal{S}}, u_{\mathcal{L}\setminus\mathcal{P}\cup\mathcal{K}}) \to w_{\mathcal{S}}$, and the external signals on the starting nodes of the vertex disjoint paths are persistently exciting.

Proof: See Appendix.

Proposition 6. Let $\bar{e}_{\mathcal{U}\setminus\mathcal{W}_j^-}$ be defined as all signals e that have a direct or unmeasured path to $u_{\mathcal{U}\setminus\mathcal{W}_j^-}$. Then the spectrum condition in Proposition 4 is generically satisfied if there are $dim(w_{\mathcal{N}_j^-\setminus\mathcal{U}})$ vertex disjoint paths from $(u_{\mathcal{L}\setminus\mathcal{K}_j}, \bar{e}_{\mathcal{U}\setminus\mathcal{W}_j^-}) \to w_{\mathcal{N}_j^-\setminus\mathcal{U}}$ that do not pass through $w_{\mathcal{N}_j^-\cap\mathcal{U}}$, and the external signals on the starting nodes of the vertex disjoint paths are persistently exciting.

Proof: See Appendix.

The results in Propositions 3 and 5 show that due to full rank disturbances on $w_{\mathcal{S}}$, there will always be $dim(w_{\mathcal{S}})$ vertex disjoint paths from $\xi_{\mathcal{S}}$ to $w_{\mathcal{S}}$, and so the data-informativity conditions for Step 2 of the algorithm are always generically satisfied. The question whether we need additional external excitation signals $u_{\mathcal{L}\setminus\mathcal{K}_j}$ therefore depends on whether these are needed to satisfy the data-informativity conditions in Proposition 4 and 6.

Comparing the spectral condition in Proposition 4 to the spectral condition in the local direct method [15], given by

$$\Phi_\kappa(\omega) \succ 0, \kappa = \begin{bmatrix} w_{\mathcal{D}} \\ \xi_{\mathcal{Y}} \\ u_{\mathcal{K}} \end{bmatrix}, \tag{17}$$

we note that the MISO predictor model (15) only requires data-informativity for a subset of the signals $w_{\mathcal{S}}$ and $u_{\mathcal{K}}$. The result in Proposition 6 then implies that we have more external excitation signals available for achieving data-informativity, where additionally the set of node signals for which data-informativity is required is smaller compared to the current direct methods. Moreover, note that for the MISO predictor (15) $j \notin \mathcal{N}_j^-$, while it is still allowed in the local direct method for node signals to appear as both predictor input and output $\mathcal{Y}\cap\mathcal{D} \neq \emptyset$, for which we require $dim(\mathcal{Y}\cap\mathcal{D})$ vertex disjoint paths from independent excitation signals $u_{\mathcal{P}}$ to achieve data-informativity [15].

The results show that the local identification version of the multi-step least squares method indeed allows for a smaller number of measured external excitation signals to achieve generic data-informativity compared to the local direct method.

VIII. 2-NODE EXAMPLE

We consider the 2-node example in Figure 1 as our selected subnetwork for the local identification method. Due to the correlated disturbances on w_1 and w_2, with $j = 2$, we select $\mathcal{Y} = \mathcal{S} = \mathcal{L} = \{1, 2\}$ and $\mathcal{U} = \emptyset$. We reconstruct both the innovation signals e_1 and e_2, and use them as measured inputs in the MISO predictor model (15) which results in

$$\hat{w}_2(t|t-1,\eta) = G_{21}(\eta)w_1 + H_{21}(\eta)\hat{\varepsilon}_1 + \check{H}_{22}(\eta)\hat{\varepsilon}_2 + u_2 \tag{18}$$

where $\check{H}_{22} = H_{22} - 1$, and $\hat{\varepsilon} = \varepsilon(\hat{\zeta}_N^n)$. Since u_2 has a known path to the output node j, $u_2 \in u_{\mathcal{P}_j}$. The effect of external excitation u_1 is incorporated in input $w_1 = w_{\mathcal{N}_j^-}$ and therefore does not appear in (18), which implies that set $\mathcal{K}_j = \emptyset$. Considering the data-informativity condition in Proposition 6 this implies that there should be one path from either u_1 or u_2 to w_1, and this is true for either u_1 or u_2. As a result the 2-node example can be solved by having either u_1 or u_2 present, whereas the local direct method [9], [14], [15] requires both u_1 and u_2 to be present to obtain a consistent estimation of the target module G_{21}. This mechanism will extend to situations of larger networks, where subsets $\mathcal{S} \subset \mathcal{L}$ need to be chosen.

IX. CONCLUSIONS

We have introduced a multi-step least squares method for single module identification that uses a MISO predictor model to directly and consistently estimate the target module, in the presence of correlated disturbances. We have shown that, when compared to current local direct methods, the data-informativity conditions for the local multi-step least squares method are relaxed while keeping key advantages of current local direct methods.

Manuscript 562 submitted to 2023 62nd IEEE Conference on Decision and Control (CDC) . Received March 31, 2023.

APPENDIX

Proof of Proposition 4 From Proposition 3 we know that $\hat{\zeta}_N^n$ is consistent and therefore $\varepsilon_{\mathcal{V}}(\hat{\zeta}_N^n)$ is a consistent estimate of $\xi_{\mathcal{V}}$. The prediction error is then written as

$$\begin{aligned}\varepsilon_j(t,\eta) &= w_j - \hat{w}_j(t|t-1,\eta)\\ &= \sum_{k\in\mathcal{N}_j^-} \Delta\bar{G}_{jk}(\eta)w_k + \sum_{\ell\in\mathcal{V}} \Delta\check{H}_{j\ell}(\eta)\xi_\ell + \xi_j \\ &+ \sum_{\gamma\in\mathcal{K}_j} \Delta\bar{J}_{j\gamma}(\eta)u_\gamma,\end{aligned} \tag{19}$$

with $\Delta\bar{G}_{jk} = \bar{G}^0_{jk} - \bar{G}_{jk}(\eta)$, $\Delta\check{H}_{j\ell}(\eta) = (\bar{H}^0_{jk} - I_{jj}) - \check{H}_{j\ell}(\eta)$, $\Delta\bar{J}_{j\gamma} = \bar{J}^0_{j\gamma} - \bar{J}_{j\gamma}(\eta)$. From here on we can follow the proof of Proposition 2 in [16]. □

Proof of Proposition 5 The generic data-informativity conditions are satisfied if there are $dim(\kappa)$ vertex disjoint paths in the mapping

$$(\xi_{\mathcal{S}}, u_{\mathcal{L}}) \to (w_{\mathcal{S}}, u_{\mathcal{P}\cup\mathcal{K}}).$$

We can remove the external excitation signals that have the same components on the right and left hand side, which results in

$$(\xi_{\mathcal{S}}, u_{\mathcal{L}\setminus\mathcal{P}\cup\mathcal{K}}) \to w_{\mathcal{S}}.$$

□

Proof of Proposition 6 The generic data-informativity conditions are satisfied if there are $dim(\bar{\kappa})$ vertex disjoint paths $(u_{\mathcal{L}}, \xi_{\mathcal{S}}) \to (w_{\mathcal{N}_j^-}, \xi_{\mathcal{V}}, u_{\mathcal{K}_j})$, or equivalently

$$(u_{\mathcal{L}}, \xi_{\mathcal{V}}, \xi_{\mathcal{U}}) \to (w_{\mathcal{N}_j^-}, \xi_{\mathcal{V}}, u_{\mathcal{K}_j}).$$

The external signals that have similar component on both the left hand side and right hand side of the mapping can be removed, leading to $(u_{\mathcal{L}\setminus\mathcal{K}_j}, \xi_{\mathcal{U}}) \to w_{\mathcal{N}_j^-}$, or equivalently

$$(u_{\mathcal{L}\setminus\mathcal{K}_j}, \xi_{\mathcal{U}\setminus\mathcal{N}_j^-}, \xi_{\mathcal{N}_j^-\cap\mathcal{U}}) \to (w_{\mathcal{N}_j^-\setminus\mathcal{U}}, w_{\mathcal{N}_j^-\cap\mathcal{U}}).$$

Under the full rank noise assumption we know that $\xi_{\mathcal{N}_j^-\cap\mathcal{U}}$ has $dim(w_{\mathcal{N}_j^-\cap\mathcal{U}})$ vertex disjoint paths to $w_{\mathcal{N}_j^-\cap\mathcal{U}}$. Therefore we can remove $w_{\mathcal{N}_j^-\cap\mathcal{U}}$ and $\xi_{\mathcal{N}_j^-\cap\mathcal{U}}$ under the condition that the paths do not pass through $w_{\mathcal{N}_j^-\cap\mathcal{U}}$. This result in

$$(u_{\mathcal{L}\setminus\mathcal{K}_j}, \xi_{\mathcal{U}\setminus\mathcal{N}_j^-}) \to w_{\mathcal{N}_j^-\setminus\mathcal{U}},$$

where the vertex disjoint paths are not allowed to pass through $w_{\mathcal{N}_j^-\cap\mathcal{U}}$. Relating the signals $\xi_{\mathcal{U}\setminus\mathcal{N}_j^-}$ to the noise signals in the original network then leads to the result. □

REFERENCES

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

[2] P. M. J. Van den Hof, A. G. 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*, vol. 49, no. 10, pp. 2994–3006, 2013.

[3] J. Hendrickx, M. Gevers, and A. Bazanella, "Identifiability of dynamical networks with partial node measurements," *IEEE Trans. Autom. Control*, vol. 64, no. 6, pp. 2240–2253, 2019.

[4] N. Everitt, G. Bottegal, and H. Hjalmarsson, "An empirical Bayes approach to identification of modules in dynamic networks," *Automatica*, vol. 91, pp. 144–151, May 2018.

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

[6] M. Verhaegen, C. Yu, and B. Sinquin, *Data-Driven Identification of Networks of Dynamic Systems*. Cambridge University Press, 2022.

[7] M. Gevers, A. Bazanella, and G. Vian da Silva, "A practical method for the consistent identification of a module in a dynamical network," *IFAC-PapersOnLine*, vol. 51-15, pp. 862–867, 2018, proc. 18th IFAC Symp. System Identif. (SYSID2018).

[8] A. S. Bazanella, M. Gevers, and J. M. Hendrickx, "Network identification with partial excitation and measurement," in *Proc. 58th IEEE Conference on Decision and Control (CDC)*, 2019, pp. 5500–5506.

[9] K. R. Ramaswamy and P. M. J. Van den Hof, "A local direct method for module identification in dynamic networks with correlated noise," *IEEE Trans. Automatic Control*, vol. 66, pp. 5237–5252, November 2021.

[10] S. Jahandari and D. Materassi, "Optimal selection of observations for identification of multiple modules in dynamic networks," *IEEE Trans. Automatic Control*, vol. 67, no. 9, pp. 4703–4716, 2022.

[11] L. Ljung, *System Identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[12] S. Shi, X. Cheng, and P. M. J. Van den Hof, "Single module identifiability in linear dynamic networks with partial excitation and measurement," *IEEE Trans. Automatic Control*, vol. 68, pp. 285–300, January 2023.

[13] P. M. J. Van den Hof, A. G. Dankers, and H. H. M. Weerts, "From closed-loop identification to dynamic networks: generalization of the direct method," in *Proc. 56nd IEEE Conf. on Decision and Control (CDC)*. Melbourne, Australia: IEEE, 2017, pp. 5845–5850.

[14] P. M. J. Van den Hof and K. R. Ramaswamy, "Path-based data-informativity conditions for single module identification in dynamic networks," in *Proc. 59th IEEE Conf. on Decision and Control (CDC)*. Jeju Island, Republic of Korea: IEEE, 2020, pp. 4354–4359.

[15] P. M. J. Van den Hof, K. R. Ramaswamy, and S. J. M. Fonken, "Integrating data-informativity conditions in predictor models for single module identification in dynamic networks," in *Proc. 22nd IFAC World Congress*, Yokohama, Japan, 2023, to appear.

[16] S. J. M. Fonken, K. R. Ramaswamy, and P. M. J. Van den Hof, "A scalable multi-step least squares method for network identification with unknown disturbance topology," *Automatica*, vol. 141, no. 110295, July 2022.

[17] J. Pearl, *Causality: Models, Reasoning, and Inference*. New York: Cambridge University Press, 2000.

[18] A. G. Dankers, P. M. J. Van den Hof, P. S. C. Heuberger, and X. Bombois, "Identification of dynamic models in complex networks with prediction error methods: Predictor input selection," *IEEE Trans. on Automatic Control*, vol. 61, no. 4, pp. 937–952, 2016.

[19] S. Shi, X. Cheng, B. De Schutter, and P. M. J. Van den Hof, "Signal selection for local module identification in linear dynamic networks: A graphical approach," in *Proc. 22nd IFAC World Congress*, Yokohama, Japan, 2023, to appear.

[20] L. Ljung and B. Wahlberg, "Asymptotic properties of the least-squares method for estimating transfer functions and disturbance spectra," *Advances in Applied Probability*, vol. 24, no. 2, pp. 412–440, 1992.

[21] M. Galrinho, C. R. Rojas, and H. Hjalmarsson, "Parametric identification using weighted null-space fitting," *IEEE Trans. Automatic Control*, vol. 64, no. 7, pp. 2798–2813, 2019.

[22] S. J. M. Fonken, M. Ferizbegovic, and H. Hjalmarsson, "Consistent identification of dynamic networks subject to white noise using weighted null-space fitting," *IFAC-PapersOnLine*, vol. 53-2, pp. 46–51, 2020, proc. 21st IFAC World Congress, Berlin, Germany.

[23] K. R. Ramaswamy, G. Bottegal, and P. M. J. Van den Hof, "Learning linear models in a dynamic network using regularized kernel-based methods," *Automatica*, vol. 129, no. 109591, July 2021.