Single module identification in dynamic networks - the current status ⋆

Paul M.J. Van den Hof and Karthik R. Ramaswamy

Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands
(e-mail: p.m.j.vandenhof@tue.nl; k.r.ramaswamy@tue.nl).

Abstract: Over the last decade, the problem of data-driven modeling in linear dynamic networks has been introduced in the literature, and has shown to contain many different challenging research questions, that go far beyond the classical problems in open-loop and closed-loop identification. The structural and topological properties of networks become a central ingredient in the related identification setting, as well as the selection of locations for signals to be sensed and for excitation signals to be added. In this seminar we will present an overview of recent results that are obtained for the problem of identification of a single link/module in a dynamic network of which the topology is given. The surveyed methods include extensions of the direct and indirect methods of closed-loop identification, as well as Wiener filter approaches and Bayesian kernel-based methods. Particular attention will be given to the selection of signals that need to be available for measurement/excitation, and accuracy properties of the estimated models in terms of consistency and minimum variance properties.

Keywords: System identification, identifiability, dynamic networks, kernel-based methods, systems over graphs.

1. INTRODUCTION

Linear dynamic networks are structured systems that are composed of interconnected linear time-invariant systems. Typically a dynamic network induces a graph, with vertices and edges, that represents the topology of the network. Often a network is represented in a state-space form with states as node signals represented by the vertices in the graph, and the state transitions as links or edges in the graph. However in an identification setting, where not all states of a system are typically measured, it has appeared to be attractive to represent the network in a graph that has (measured) node signals as vertices, and dynamic transfer functions on the links/edges. The basic setting of *Dynamic Structure Functions* that was introduced in Gonçalves and Warnick (2008), was generalized to a stochastic estimation and identification setting in Van den Hof et al. (2013), and has been adopted by several different authors.

In this setting a dynamic network is built up out of L scalar *internal variables* or *nodes* w_j, $j = 1, \dots, L$, and K *external variables* r_k, $k = 1, \dots K$. Each internal variable is described as:

$$w_j(t) = \sum_{\substack{l=1 \\ l \neq j}}^{L} G_{jl}(q) w_l(t) + u_j(t) + v_j(t) \qquad (1)$$

where q^{-1} is the delay operator, i.e. $q^{-1}w_j(t) = w_j(t-1)$;

- G_{jl} are proper rational transfer functions, referred to as *modules*.
- u_j is an input signal, $u_j(t) = \sum_{k=1}^{K} R_{jk}(q) r_k(t)$ with r_k *external variables* that can directly be manipulated by the user.
- v_j is *process noise*, where the vector process $v = [v_1 \cdots v_L]^T$ is modelled as a stationary stochastic process with rational spectral density $\Phi_v(\omega)$, such that there exists a white noise process $e := [e_1 \cdots e_L]^T$, with covariance matrix $\Lambda > 0$ such that $v(t) = H(q)e(t)$, where H is square, stable, monic and minimum-phase.

When combining the L node signals we arrive at the full network expression

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

which results in the matrix equation:

$$w = G(q)w + R(q)r + H(q)e, \qquad (2)$$

where by construction the matrix G is hollow, i.e. it has diagonal entries 0, while it encodes the topology of the network, i.e. $G_{j\ell}(q) \neq 0$ if and only if there is a connection from node w_ℓ to node w_j in the network.

The single module identification problem to be considered is the problem of identifying one particular module $G_{ji}(q)$ on the basis of measured time-series of a subset of variables in w, and possibly r, for the situation that the network topology is known. This is illustrated in the network depicted in Figure 1.

⋆ 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).

It may be clear that simply measuring the input and output of the target module and estimating a model on the basis of these signals, will generally not lead to accurate results, because of the signal correlations that are induced by the remaining part of the network.

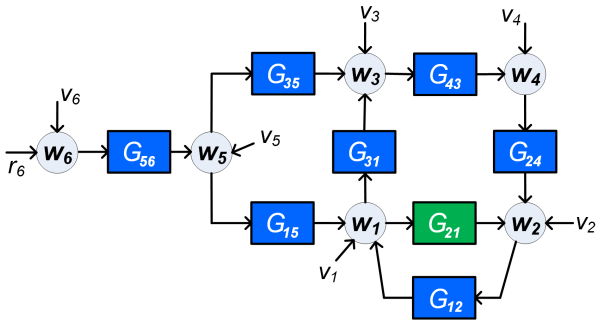

Fig. 1. Example network with the green module G_{21} being the target module for identification (Ramaswamy and Van den Hof (2019)).

Non-uniqueness

The network representation (2) will in general be non-unique. E.g. for the situation $R(q) = 0$, the dynamic properties of the network are reflected by the spectral density $\Phi_w(\omega)$, while this spectrum can be generated by different combinations of G, H and Λ. In this situation uniqueness is achieved e.g. if the noise model H is restricted to be diagonal (Bottegal et al. (2018)). However, in situations that the network and its topology result from structured first principle modelling, it is still relevant to consider situations of non-diagonal H, as disturbances in different locations of the network can very well be correlated. For the general case of $R(q) \neq 0$, the freedom of transforming the network to equivalent representations is analyzed in Weerts et al. (2020).

2. MAIN APPROACHES

We can distinguish several different approaches for addressing the single module identification problem, where the target module is indicated by G_{ji}.

(1) A *direct method*, that is based on selecting a particular set of predictor input signals w_k, $k \in \mathcal{D}$, and a set of predicted output signals w_ℓ, $\ell \in \mathcal{Y}$, with $i \in \mathcal{D}$, $j \in \mathcal{Y}$, and estimating a dynamic model based on a prediction error:

$$\varepsilon(t,\theta) = \bar{H}(q,\theta)^{-1}[w_{\mathcal{Y}}(t) - \bar{G}(q,\theta)w_{\mathcal{D}}(t)], \quad (3)$$

where $\bar{G}(q,\theta)$ and $\bar{H}(q,\theta)$ are parametrized transfer function matrices. The target module is then embedded in the model $\bar{G}(q,\theta)$, and the objective is to estimate the target module consistently and possibly with minimum variance.

(2) An *indirect method*, that is based on selecting a particular set of predictor input signals r_k, $k \in \mathcal{D}$, and a set of predicted outputs w_ℓ, $\ell \in \mathcal{Y}$, that are used in a predictor model, leading to

$$\varepsilon(t,\theta) = w_{\mathcal{Y}}(t) - \bar{T}(q,\theta)r_{\mathcal{D}}(t). \quad (4)$$

The matrix $\bar{T}$ refers to a submatrix of the network transfer matrix $T := (I-G)^{-1}R$, which maps external signals r into internal node signals w. In order to extract the dynamics of a particular module G_{ji} from an estimated $\bar{T}$, a postprocessing step is necessary. Consistency of the target module estimate is the typical objective. Different variations of indirect methods exist, including two-stage and instrumental variable (IV) methods.

(3) Non-parametric approaches, where relations between particularly estimated spectral densities of internal signals are used as a basis for module estimation.

(4) Subspace methods, that can handle non-measured interacting signals between the modules, see Yu and Verhaegen (2018).

(5) Recently a generalization of the direct and indirect method was introduced based on a predictor model with prediction error

$$\varepsilon(t,\theta) = \bar{H}(q,\theta)^{-1}[w_{\mathcal{Y}}(t) - \bar{G}(q,\theta)w_{\mathcal{D}_w}(t) - \bar{T}(q,\theta)r_{\mathcal{D}_r}(t)],$$

and where the target module estimate is obtained after post-processing the estimated $\bar{G}$ and $\bar{T}$. This method allows for more flexibility in selecting the node signals to be measured for identification, see Ramaswamy et al. (2019).

The direct method (3) has node signals $w_{\mathcal{D}}$ as predictor inputs, and therefore utilizes both external signals r and e for creating data-informativity. On the other hand, indirect methods rely on external excitation signals r only for data informativity, and therefore will typically require more "expensive" external excitations. The direct method provides asymptotically efficient estimates (i.e. consistency and minimum variance for the identification setup) at the cost of the need to include noise models $\bar{H}(q,\theta)$. The indirect method and its variations provides consistent estimates but not with minimum variance. When the node signals are measured with sensor noise (errors-in-variables (EIV) situation), the direct method becomes biased and the indirect method provides consistent estimates of the target module.

3. DIRECT METHOD FOR UNCORRELATED DISTURBANCES

In the situation that it is known that the process disturbances are uncorrelated, i.e. Φ_v is diagonal, the direct identification method can typically be reduced to a MISO problem, i.e. in (3) $w_{\mathcal{D}}$ is a vector signal, and $w_{\mathcal{Y}} = w_j$ is scalar. The simplest situation is when all w-in-neighbors of w_j (denoted by $w_{\mathcal{N}_j^-}$) are included, i.e. $w_{\mathcal{D}} = w_{\mathcal{N}_j^-}$.

The target module can then directly be parametrized and estimated as part of a MISO model, see Van den Hof et al. (2013). For consistency, conditions on informativity of the data have to be satisfied, implying that sufficient excitation should be present in the predictor inputs. A typical, but conservative, condition is that $\Phi_{w_m}(\omega)$ is positive definite for a sufficient number of frequencies where w_m is the vector of stacked predicted output and predictor input signals. For a particular situation, a less conservative condition is formulated in Gevers and Bazanella (2015).

In general it is not necessary to use the full set $w_{\mathcal{N}_j^-}$ as predictor input. It is sufficient to select a subset $w_{\mathcal{D}}$ that satisfies the property that upon removal (*immersion*) of the remaining unmeasured nodes from the network, the target module remains invariant. This is achieved if the *parallel path and loop condition* is satisfied (Dankers et al.

(2016)): all parallel paths from w_i to output w_j that do not pass through G_{ji}, and all loops around w_j need to pass through a measured node signal in $w_{\mathcal{D}}$. *Confounding variables* [1] can occur during the process of immersion and leads to the loss of consistency. This can be resolved by including additional node signals in $w_{\mathcal{D}}$, see Dankers et al. (2017). By using abstractions as an alternative to immersion, the parallel path and loop condition can further be generalized, introducing more flexibility in the signal selection, see Weerts et al. (2020).

4. DIRECT METHOD FOR CORRELATED DISTURBANCES

In the situation that it is known that the process noise is correlated and the disturbance correlation structure, i.e. the non-zero entries of Φ_v, is known the problem of dealing with confounding variables becomes even more pronounced. Next to the *indirect confounding variables* that can occur due to non-measured input signals, now also *direct confounding variables* appear as a result of disturbance correlations, leading to biased estimates when not properly being taken care of. The direct confounding variables can only be handled by enlarging the output vector $w_{\mathcal{Y}}$ to include those node signals that have disturbances that are correlated with the disturbance on $w_{\mathcal{Y}}$, and including an appropriate multivariate noise model. Building on the results in the previous section this leads to a triple set of conditions for arriving at consistent (and efficient) estimates of the target module:

(1) Selecting predictor $w_{\mathcal{D}}$ to satisfy the parallel path and loop condition;
(2) Handling of indirect and direct confounding variables, by adding predictor inputs and/or adding predicted outputs; and
(3) Persistency of excitation conditions, typically formulated in the form of a positive definite signal spectrum of a specified signal vector, and thus requiring sufficient (external) excitation, either through excitation or through disturbance signals.

Path-based conditions for most of these aspects have been formulated in Ramaswamy and Van den Hof (2019). The resulting MIMO identification setup, indicated by (3), is sketched in Figure 2, where $w_{\mathcal{Q}}$ denote the node signals that appear both as input and as output in the predictor model. The results provide multiple options for the user to choose from, in terms of how many and which node signals to include in the estimation setup, dependent on the availability of measured node signals.

5. INDIRECT METHOD AND ITS VARIANTS

The network model (2) can be rewritten as $w = Tr + \bar{v}$ where $\bar{v} = (I - G)^{-1}He$. A consistent estimate $\hat{T}(q)$ of $T(q)$ can be obtained using open loop MIMO identification method as in (4). On the basis of $\hat{T}(q)$, a consistent estimate $\hat{G}$ of G can be obtained by solving $(I - \hat{G})\hat{T}(q) = R$. By identifiying only a submatrix of T and solving only a subset of the above equations, a target module embedded

[1] Confounding variables are unmeasured variables that affect both the inputs and outputs of an estimation problem.

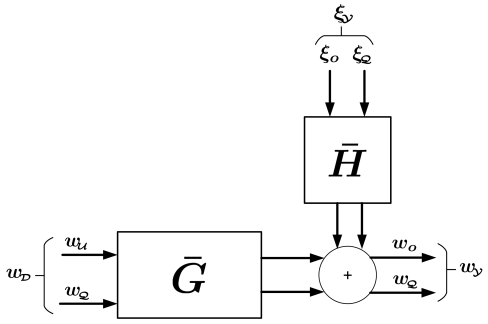

Fig. 2. Identification setup and classification signals in the input and output of the identification problem, Ramaswamy and Van den Hof (2019).

in the dynamic network can be identified, see Gevers et al. (2018); Hendrickx et al. (2019). For the situation of having excitation signals on all node signals, it has been analyzed which node signals to measure for consistent identification of the target module. This is further relaxed in Bazanella et al. (2019). Related indirect methods, such as the *two-stage method* and the *Instrumental Variable* (IV) method have been presented in Van den Hof et al. (2013) and Dankers et al. (2015) respectively. A semi-parametric approach has been introduced in Galrinho et al. (2017) where a parametric model of the target module is consistently identified using a multi-step approach, while avoiding non-convex optimization. All the indirect methods can handle an EIV setting as well as networks with correlated process noise.

6. NON-PARAMETRIC APPROACH

Frequency domain non-parametric approaches are provided in Dankers and Van den Hof (2015) and Materassi and Salapaka (2015). In these methods the spectral densities of different signals in the network are used to consistently identify the target module. Flexibility to identify both proper and non-proper modules using a Wiener filter based approach has been provided in Materassi and Salapaka (2015). This approach uses the d-separation principle from the probabilistic graphical model theory to select the node signals for the identification in dynamic networks, see Materassi and Salapaka (2019).

7. KERNEL-BASED METHODS

By incorporating kernel-based method, the impulse response(s) of the modules are modeled as zero-mean Gaussian processes whose covariance(s) are described by a kernel that ensures smoothness and stability of the model. A probabilistic description of the model is obtained and the coefficients of the impulse response(s) are obtained by estimating the hyperparameters of the kernel by maximizing the marginal likelihood of the data. In this way, the impulse response of each module is obtained through estimating only a few hyperparameters (eg. 2 hyperparamters per module for stable spline kernel).
In Chiuso and Pillonetto (2012), this has been applied in a time domain non-parametric approach for estimating models in a MISO setup with white output noise. Following a semi-parametric approach to a dynamic network with only sensor noise (no process noise), the increase in variance due

to high order modeling in a two-stage method is reduced in Everitt et al. (2018) by applying a kernel-based method. Similarly, the direct method demands a model order selection step for all modules in the MISO setup, which increases the complexity and estimation of large number of nuisance parameters. In Ramaswamy et al. (2018), a kernel-based method has been used to tackle these problems by modeling the target module as a parametric model and the remaining modules in the MISO setup as Gaussian processes, thus avoiding the model order selection step and decreasing the number of parameters. This offers a substantially reduced variance of the target module estimate.

8. IDENTIFIABILITY

When considering network identifiability (Weerts et al. (2018a)), conditions can be formulated for identifiability of a single module in a network model set. This typically leads to rank conditions on particular transfer functions from external signals to particular node signals. In a generic sense, this can be translated to path-based conditions on the graph of the network model set (Hendrickx et al. (2019) and followed up by Weerts et al. (2018b)). A synthesis procedure to assign and locate the minimum number of external excitation signals for guaranteeing local module generic identifiability, is provided in Shi et al. (2020). Note that these results are independent of the particular identification method considered.

REFERENCES

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

Bottegal, G., Chiuso, A., and Van den Hof, P.M.J. (2018). On dynamic network modeling of stationary multivariate processes. *IFAC-PapersOnLine*, 51-15, 850–861.

Chiuso, A. and Pillonetto, G. (2012). A Bayesian approach to sparse dynamic network identification. *Automatica*, 48(8), 1553–1565.

Dankers, A.G. and Van den Hof, P.M.J. (2015). Non-parametric identification in dynamic networks. In *Proc. 54th IEEE Conf. on Decision and Control (CDC)*, 3487–3492. Osaka, Japan.

Dankers, A.G., Van den Hof, P.M.J., Bombois, X., and Heuberger, P.S.C. (2015). Errors-in-variables identification in dynamic networks – consistency results for an instrumental variable approach. *Automatica*, 62, 39–50.

Dankers, A.G., Van den Hof, P.M.J., Heuberger, P.S.C., and Bombois, X. (2016). Identification of dynamic models in complex networks with prediction error methods: Predictor input selection. *IEEE Trans. on Automatic Control*, 61(4), 937–952.

Dankers, A.G., Van den Hof, P.M.J., Materassi, D., and Weerts, H.H.M. (2017). Relaxed predictor input selection rules for handling confounding variables in dynamic networks. *IFAC-PapersOnLine*, 50-1, 3983–3988.

Everitt, N., Bottegal, G., and Hjalmarsson, H. (2018). An empirical Bayes approach to identification of modules in dynamic networks. *Automatica*, 91, 144–151.

Galrinho, M., Everitt, N., and Hjalmarsson, H. (2017). Incorporating noise modeling in dynamic networks using non-parametric models. *IFAC-PapersOnLine*, 50(1), 10568–10573.

Gevers, M. and Bazanella, A.S. (2015). Identification in dynamic networks: identifiability and experiment design. In *Proc. 2015 IEEE 54th Conf. Decision and Control, Osaka, Japan*, 4006–4011.

Gevers, M., Bazanella, A., and Vian da Silva, G. (2018). A practical method for the consistent identification of a module in a dynamical network. *IFAC-PapersOnLine*, 51-15, 862–867.

Gonçalves, J. and Warnick, S. (2008). Necessary and sufficient conditions for dynamical structure reconstruction of LTI networks. *IEEE Trans. Autom. Control*, 53(7), 1670–1674.

Hendrickx, J., Gevers, M., and Bazanella, A. (2019). Identifiability of dynamical networks with partial node measurements. *IEEE Trans. Autom. Control*, 64(6), 2240–2253.

Materassi, D. and Salapaka, M. (2015). Identification of network components in presence of unobserved nodes. In *Proc. 2015 IEEE 54th Conf. Decision and Control, Osaka, Japan*, 1563–1568.

Materassi, D. and Salapaka, M.V. (2019). Signal selection for estimation and identification in networks of dynamic systems: a graphical model approach. *ArXiv preprint arXiv:1905.12132*.

Ramaswamy, K.R., Bottegal, G., and Van den Hof, P.M.J. (2018). Local module identification in dynamic networks using regularized kernel-based methods. In *Proc. 57th IEEE Conf. on Decision and Control (CDC)*, 4713–4718. Miami Beach, FL.

Ramaswamy, K.R. and Van den Hof, P.M.J. (2019). A local direct method for module identification in dynamic networks with correlated noise. *ArXiv:1908.00976*. Provisionally accepted by IEEE Trans. Automatic Control.

Ramaswamy, K.R., Van den Hof, P.M.J., and G., D.A. (2019). Generalized sensing and actuation schemes for local module identification in dynamic networks. In *Proc. 58th IEEE Conf. on Decision and Control (CDC)*, 5519–5524. Nice, France.

Shi, S., Cheng, X., and Van den Hof, P.M.J. (2020). Excitation allocation for generic identifiability of a single module in dynamic networks: A graphic approach. In *Proc. 21st IFAC World Congress*. Berlin, Germany.

Van den Hof, P.M.J., Dankers, A.G., Heuberger, P.S.C., and Bombois, X. (2013). Identification of dynamic models in complex networks with prediction error methods - basic methods for consistent module estimates. *Automatica*, 49(10), 2994–3006.

Weerts, H.H.M., Linder, J., Enqvist, M., and Van den Hof, P.M.J. (2020). Abstractions of linear dynamic networks for input selection in local module identification. *Automatica*, 117.

Weerts, H.H.M., Van den Hof, P.M.J., and Dankers, A.G. (2018a). Identifiability of linear dynamic networks. *Automatica*, 89, 247–258.

Weerts, H.H.M., Van den Hof, P.M.J., and Dankers, A.G. (2018b). Single module identifiability in linear dynamic networks. In *Proc. 57th IEEE Conf. on Decision and Control (CDC)*, 4725–4730. IEEE, Miami Beach, FL.

Yu, C. and Verhaegen, M. (2018). Subspace identification of individual systems operating in a network (SI^2ON). *IEEE Trans. Automatic Control*, 63(4), 1120–1125.