Abdalmoaty, M ., Rojas, C R., Hjalmarsson, H. (2018)
Identication of a Class of Nonlinear Dynamical Networks
In:
IFAC-PapersOnLine

Preprints,18th IFAC Symposium on System Identification
July 9-11, 2018. Stockholm, Sweden

Identification of a Class of Nonlinear Dynamical Networks ⋆

Mohamed R.H. Abdalmoaty Cristian R. Rojas Håkan Hjalmarsson

Automatic Control Department and ACCESS Linnaeus Center, KTH Royal institute of Technology, Stockholm, Sweden (e-mail: {mohamed.abdalmoaty, cristian.rojas, hakan.hjalmarsson}@ee.kth.se)

Abstract: Identification of dynamic networks has attracted considerable interest recently. So far the main focus has been on linear time-invariant networks. Meanwhile, most real-life systems exhibit nonlinear behaviors; consider, for example, two *stochastic* linear time-invariant systems connected in series, each of which has a nonlinearity at its output. The estimation problem in this case is recognized to be challenging, due to the analytical intractability of both the likelihood function and the optimal one-step ahead predictors of the measured nodes. In this contribution, we introduce a relatively simple prediction error method that may be used for the estimation of nonlinear dynamical networks. The estimator is defined using a deterministic predictor that is nonlinear in the known signals. The estimation problem can be defined using closed-form analytical expressions in several non-trivial cases, and Monte Carlo approximations are not necessarily required. We show, that this is the case for some block-oriented networks with no feedback loops and where all the nonlinear modules are polynomials. Consequently, the proposed method can be applied in situations considered challenging by current approaches. The performance of the estimation method is illustrated on a numerical simulation example.

Keywords: System Identification, Dynamical Networks, Stochastic Systems, Block-Oriented Models, Prediction Error Method.

1. INTRODUCTION

In recent years, system identification of dynamical models in complex networks (dynamical networks) has gained an increasing attention, mainly due to the increased complexity of modern applications. Nowadays, dynamical networks can be found in many domains, for example, in systems biology, economic systems, and engineering applications such as smart power grids and transportation systems, to name a few. They play an important role in understanding complex real-life systems, but also as a modeling tool for the operation and design of complex technological systems; see Lamnabhi-Lagarrigue et al. [2017].

So far, the focus of the system identification community has been mostly on dynamical networks of Linear Time-Invariant (LTI) systems. The problem of identifying the interconnection structure (topology) of LTI networks has been considered under different assumptions in Materassi and Innocenti [2010], Sanandaji et al. [2011], Materassi and Salapaka [2012a,b], Materassi et al. [2013]. A framework for consistency-based identification of a single module in an LTI network with known interconnection structure has been introduced in Van den Hof et al. [2013]. In that framework, it is assumed that all the nodes are observed and classical closed-loop identification methods [Van den Hof et al. 1992, Forssell and Ljung 1999] were generalized to the LTI dynamical networks scenario. In a related contribution, Dankers et al. [2016] considered the selection problem of variables to be used as inputs to obtain a consistent estimate of a specific module. Everitt et al. [2016] introduced an alternative approach for the identification of a single module via nonparametric modeling of the rest of the network. The identification of dynamical networks with unobserved nodes has been considered in Linder and Enqvist [2017]. Moreover, conditions for identifiability of LTI dynamical networks have been studied in Weerts et al. [2015, 2016] and more recently in Gevers et al. [2017].

On the other hand, contributions on the identification of nonlinear dynamical networks are so far limited to very special cases. See, for example, Giri and Bai [2010] and Schoukens et al. [2014] for the identification of block-oriented nonlinear models. Apart from the very specific interconnection structure, these models have only two nodes (inputs and outputs) and they are meant to represent a single isolated system. Furthermore, in almost every case (additive) noise is only allowed at the output node (i.e., only deterministic systems are considered). The presence of unobserved stochastic processes at the input of a nonlinear module (block) renders the commonly used estimation methods analytically intractable. In recent years, there has been a growing interest in the problem of consistent estimation in these situations. Most of the available methods are based on Monte Carlo approximations of the Maximum Likelihood Estimator or the optimal Mean-Square

⋆ This work was supported by the Swedish Research Council under contracts 2015-05285, 2016-06079 and by the European Research Council under the advanced grant LEARN, contract 267381.

Error (MSE) predictor in a prediction error method. See, for example, Durbin and Koopman [1997], Abdalmoaty and Hjalmarsson [2016] for methods based on importance sampling, as well as Ninness et al. [2010], Schön et al. [2011], Wills et al. [2013], Lindsten [2013], Schön et al. [2015] for methods based on sequential Monte Carlo (SMC a.k.a. particle filters) algorithms and (particle) Markov Chain Monte Carlo ((p)MCMC) algorithms. Under some conditions, these methods are asymptotically efficient and some have been shown to provide interesting results on several benchmark examples. However, their application is so far limited to cases where fundamental difficulties, such as the sample degeneracy problem (see Doucet and Johansen [2009]), can be avoided. It is also known that their convergence can be slow; for example, when the latent processes have small variances.

In this contribution, we introduce a consistent prediction error method (PEM) based on a computationally attractive one-step ahead predictor [Abdalmoaty 2017]. We assume that the interconnection structure of the dynamic network (the network's topology) is given, and that only a subset of the network's nodes is measured. It is also assumed that all the nodes are disturbed by unobserved stationary stochastic disturbance, and that the modeled network has no feedback loops. The proposed predictor can then be seen as an extension of the Output-Error predictor of linear models [Ljung 1999] to a general class of nonlinear models where the measured outputs have a finite mean value. It only relies on the first moment of the measured nodes, and the computation of the likelihood function is not required; this is a great computational advantage. We show that in the case of block-oriented networks, where the modules are either linear time-invariant (LTI) or polynomial nonlinearities, the predictor and the estimation problem may be defined using closed-form expressions. Furthermore, when standard ergodicity assumptions on the data hold, consistency of the proposed PEM estimator can be established whenever the parameterization of the network is identifiable via the mean of the measured nodes. The price paid for by-passing the likelihood function computation is a loss of accuracy. A detailed account of the asymptotic analysis will be considered in a future contribution by the authors.

The paper's outline is as follows. In Section 2, we set the notations, the assumptions and formulate the main problem. In Section 3, we introduce the Output-Error-type predictor, which is then used in Section 4 to construct a consistent PEM. In Section 5, the performance of the estimation method is demonstrated on a numerical simulation example. Finally, the paper is concluded in Section 6.

2. PROBLEM FORMULATION

We extend the framework used in Van den Hof et al. [2013] as follows. We consider parametric dynamical networks consisting of $L \in \mathbb{N}$ internal variables, called nodes and denoted $\boldsymbol{z}_t^{(1)}, \dots, \boldsymbol{z}_t^{(L)}$. The dynamics of each node is defined by the relation

$$\boldsymbol{z}_t^{(i)} = \sum_{j \in \mathfrak{J}_i} G_j(q;\theta^\circ)\boldsymbol{z}_t^{(j)} + \sum_{k \in \mathfrak{K}_i} f_k(\boldsymbol{z}_t^{(k)};\theta^\circ) + u_t^{(i)} + \boldsymbol{w}_t^{(i)}$$

in which θ° is a finite-dimensional real vector of parameters, $G_j(q;\theta^\circ)$ are stable rational LTI systems, $f_k(\cdot;\theta^\circ)$ are static nonlinear functions, $u_t^{(i)}$ are external deterministic known signals, and $\boldsymbol{w}_t^{(i)}$ are unobserved zero mean strictly stationary process disturbances. Here, $\mathfrak{J}_i \subset \{1,\dots,L\}\backslash\{i\}$ is the set of nodes connected to the i^{th} node through LTI systems, and $\mathfrak{K}_i \subset \{1,\dots,L\} \setminus \{i\}$ is the set of nodes connected to the i^{th} node through static nonlinear functions.

In this paper, we will assume that the interconnection structure of the nodes is given, and that the network has no feedback loops. Furthermore, we only assume that a strict subset $\mathfrak{M}$ of the nodes are measured, and stack the measured variables in a column vector $\boldsymbol{y}_t$. Similarly, let us stack the external known signals in a column vector u_t, and assume that a data set

$$D_N := \{(y_t, u_t) : t = 1,\dots,N\}, \quad N \in \mathbb{N},$$

corresponding to a realization of the network's nodes, is given. Our objective is to formulate an identification method in a prediction error framework; in particular, we are interested in consistent estimators of θ°.

3. A PREDICTOR OF THE NETWORK'S MEASURED NODES

An essential component of any prediction error method is the one-step ahead predictor used to define the prediction errors. It is usually the case that the optimal mean-square error (MSE) predictor, given by the conditional mean, is used. However, as pointed out in Section 1, the optimal MSE predictor is analytically intractable in general; and, so far, approximate computations based on Monte Carlo methods can be computationally expensive or infeasible.

Fortunately, PEMs do not necessarily require an optimal predictor in order to construct consistent estimators; the used predictors do not have to be defined based on the exact full probabilistic structure of the data. A one-step predictor can be defined in several ways that may even include some ad hoc non-probabilistic arguments; see Section II.B in Ljung [1978] and Section 3.3 Ljung [1999].

Consider, for instance, a stochastic stable LTI rational model

$$\boldsymbol{y}_t = G(q;\theta)u_t + H(q)\boldsymbol{\varepsilon}_t$$

where $\boldsymbol{\varepsilon}_t$ is a zero mean stochastic process. Then, it is well-known that if the data is collected in open-loop and when standard conditions hold, the PEM estimator

$$\hat{\boldsymbol{\theta}}_N := \sum_{t=1}^{N}(\boldsymbol{y}_t - \hat{y}_t(\theta))^2$$

based on the Output-Error (OE) predictor

$$\hat{y}_t(\theta) = G(q;\theta)u_t$$

is consistent [Ljung 1999, Theorem 8.4], i.e.,

$$\hat{\boldsymbol{\theta}}_N \xrightarrow{\text{a.s.}} \theta^\circ \quad \text{as} \quad N \to \infty,$$

where the symbol $\xrightarrow{\text{a.s.}}$ denotes almost sure convergence of random variables. Observe that the knowledge of the exact noise model or the exact full distribution of $\boldsymbol{\varepsilon}_t$ were not required to obtain a consistent estimate; the only used information regarding the probabilistic structure of the data is the mean of the model's output.

It is therefore possible to generalize the above observation to a large class of dynamical models whose output posses a finite mean value; and we define what we call the *Output-Error-type (OE-type) predictors.*

***Definition* 1.** (OE-type Predictor [Abdalmoaty 2017]). The OE-type predictor of a measured process $\boldsymbol{y}_t$ is defined as the deterministic quantity

$$\hat{y}_t(\theta) := \mathbb{E}[\boldsymbol{y}_t | U_t; \theta] \tag{1}$$

in which $U_t := [u_1 \dots u_t]^\top$ is a vector containing the history of known signals influencing $\boldsymbol{y}_t$, and θ is a vector of parameters that may contain some nuisance parameters (e.g., variances/moments of latent disturbances).

The expectation in (1) is with respect to the common underlying probability space of the basic stochastic processes (all disturbances and measurement noise). A major advantage of the OE-type predictor is its simplicity; it is much easier to compute compared to the conditional mean, because no marginalization integrals have to be computed. The OE-type predictor may in fact be given in terms of a tractable or *closed-form expression* in several non-trivial cases.

To clarify this idea, consider the network shown in Figure 1. It is defined by interconnecting five nodes using three LTI systems and two static nonlinearities; there is only one external known scalar signal: u_t, and only one node is measured: $\boldsymbol{y}_t$. This network may, for instance, be a part of a bigger and more complicated network where it can be seen as a nonlinear module; the effect from other modules in the bigger network can be modeled using the disturbance signals $\boldsymbol{w}_t$. An alternative representation of

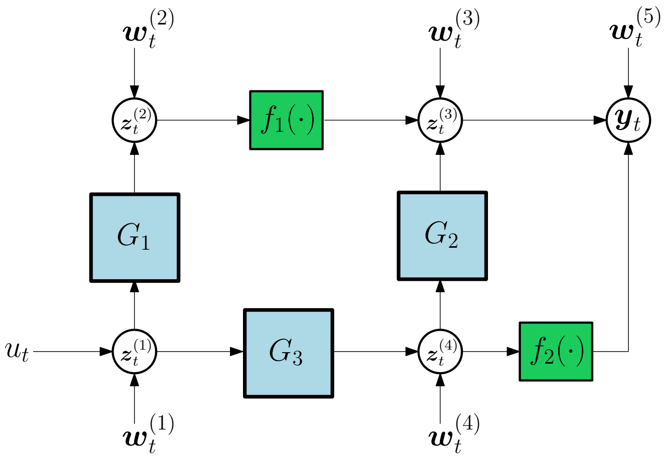

Fig. 1. An example of a block-oriented dynamical network with no loops. Only $\boldsymbol{y}_t$ is measured and all the other nodes are latent variables.

the same network is shown in Figure 2, where it is clear that the network is acyclic; i.e., has no feedback loops. However, due to the existence of the nonlinear links, the optimal predictor of $\boldsymbol{y}_t$ and the likelihood function are analytically intractable.

It is important to stress here that the computations of the OE-type predictor do not require the specification of the full distribution of $\boldsymbol{w}_t$ (it does not need to be Gaussian). However, it is assumed that the signals $\boldsymbol{w}_t$ are centered strictly stationary processes. Furthermore, we assume that the nonlinearities f_1 and f_2 are either polynomials or can be approximated well using polynomials. For all $x \in \mathbb{R}$, let us define

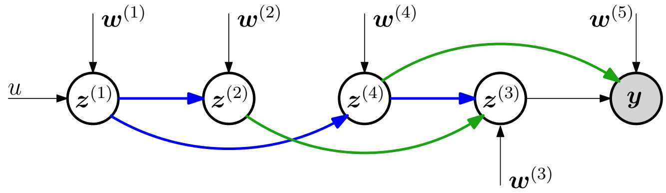

Fig. 2. Node-and-Link visualization of the network in Figure 1. Blue links denotes LTI modules, and green links represents nonlinear modules. Only the shaded node is measured

$$f_1(x) := \sum_{k=1}^{F_1} f_{1k} x^k,$$

$$f_2(x) := \sum_{k=1}^{F_2} f_{2k} x^k,$$

for some $F_1, F_2 \in \mathbb{N}$ and real coefficients $\{f_{1k}\}_{k=1}^{F_1}$ and $\{f_{2k}\}_{k=1}^{F_2}$. Then, it is straightforward to see that

$$\mathbb{E}[\boldsymbol{y}_t | U_t] = \mathbb{E}[f_1(\boldsymbol{z}_t^{(2)})] + \mathbb{E}[f_2(\boldsymbol{z}_t^{(4)})] + G_2 \mathbb{E}[\boldsymbol{z}_t^{(4)}]$$
$$= \sum_{k=1}^{F_1} f_{1k}\, m_k^{(2)} + \sum_{k=1}^{F_2} f_{2k}\, m_k^{(4)} + G_2 G_3 u_t,$$

where symbols $m_k^{(2)}$ and $m_k^{(4)}$ denote the k^{th} moments of the random variables

$$\boldsymbol{z}_t^{(2)} = G_1 u_t + G_1 \boldsymbol{w}_t^{(1)} + \boldsymbol{w}_t^{(2)}$$

and

$$\boldsymbol{z}_t^{(4)} = G_3 u_t + G_3 \boldsymbol{w}_t^{(1)} + \boldsymbol{w}_t^{(4)}$$

respectively. These are known functions of the known signals $\{u_t\}$, the parameter θ, and the moments of the disturbances $\boldsymbol{w}_t^{(1)}$, $\boldsymbol{w}_t^{(2)}$ and $\boldsymbol{w}_t^{(4)}$. Thus, the OE-type predictor of $\boldsymbol{y}_t$ is given in terms of a closed-form expression; it is parameterized by θ as well as any unknown moments of $\boldsymbol{w}_t$.

For general nonlinearities or more complicated topologies (e.g., those with feedback loops), Monte Carlo integration methods may be used to evaluate the expectation defining the OE-type predictor. Notice that even in these scenarios, the required computations are simpler than those of the optimal MSE predictor because there are no marginalization integrals to compute.

4. A CONSISTENT PREDICTION ERROR METHOD

Once the predictors of the measured nodes of the network are defined, a PEM estimator of the network's parameters can be defined. Here, we define a PEM based on the OE-type predictor as follows.

***Definition* 2.** (OE-PEM estimator). The OE-type PEM estimator of the network's parameters θ is defined as

$$\hat{\boldsymbol{\theta}}_N := \arg\min_\theta \sum_{t=1}^{N} \|\boldsymbol{y}_t - \hat{y}_t(\theta)\|^2$$

where $\hat{y}_t(\theta) = \mathbb{E}[\boldsymbol{y}_t | U_t; \theta]$ is the OE-type predictor defined in Definition 1.

Notice that the objective function of the OE-PEM problem is given in terms of closed-form expressions in all cases where the OE-type predictor has a closed-form expression.

However, similar to PEMs in a linear setting, the resulting optimization problem is in general nonlinear in θ, and iterative numerical minimization methods are needed to solve the problem.

Under standard ergodicity assumptions on the data, it may be shown that

$$\hat{\boldsymbol{\theta}}_N \xrightarrow{\text{a.s.}} \theta^* := \arg\min_{\theta} \left\{ \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \mathbb{E}[\|\boldsymbol{e}_t(\theta)\|^2] \right\}$$

as $N \to \infty$, in which $\boldsymbol{e}_t(\theta) = \boldsymbol{y}_t - \hat{y}_t(\theta)$ is the prediction error process. Furthermore, whenever the identifiability condition

$$\mathbb{E}[\boldsymbol{y}_t|U_t;\theta] = \mathbb{E}[\boldsymbol{y}_t|U_t;\theta^\circ] \iff \theta = \theta^\circ$$

holds for all sufficiently large $t \in \mathbb{Z}$, the OE-PEM estimator will be consistent and it holds that

$$\hat{\boldsymbol{\theta}}_N \xrightarrow{\text{a.s.}} \theta^\circ \quad \text{as} \quad N \to \infty.$$

The identifiability condition depends on the experimental circumstances and the parameterization of the network. A detailed asymptotic analysis will be given in a future contribution.

5. NUMERICAL EXAMPLE

In this section, we apply the OE-PEM estimator defined above to the nonlinear dynamical network shown in Figure 3. The network is defined by interconnecting three stochastic linear time-invariance models and two static nonlinear blocks. The linear modules of the network have the following representations:

$$G_1(q;\theta^\circ) = \frac{q^{-1} + b_{12}q^{-2}}{1 + f_{11}q^{-1} + f_{12}q^{-2}},$$
$$G_2(q;\theta^\circ) = \frac{q^{-1} + b_{22}q^{-2}}{1 + f_{21}q^{-1} + f_{22}q^{-2}},$$
$$G_3(q;\theta^\circ) = \frac{q^{-1}}{1 + f_{31}q^{-1}},$$

with

$$\theta^\circ = [b_{12}\ f_{11}\ f_{12}\ b_{22}\ f_{21}\ f_{22}\ f_{31}]^\top = [0.2\ -1\ 0.8125\ -0.2\ 1.6\ 0.8425\ -0.7]^\top .$$

To simplify the exposition, the static nonlinear blocks of the network are defined, for all $x \in \mathbb{R}$, by the functions

$$f_1(x) = x^2,$$
$$f_2(x) = \left(\frac{x}{20}\right)^3.$$

The basic stochastic processes $\boldsymbol{\varepsilon}_t^{(1)}$ and $\boldsymbol{\varepsilon}_t^{(2)}$ are assumed independent and mutually independent stationary Gaussian white noise with unit variance. The disturbance models are

$$H_1(q) = \frac{q^{-1} + 0.18q^{-2} - 0.484q^{-3} + 0.0096q^{-4}}{1 - 0.6q^{-1} + 0.09q^{-2} + 0.15q^{-3} - 0.085q^{-4}},$$
$$H_2(q) = \frac{0.9}{1 - 0.5q^{-1}}.$$

The resulting process disturbances $\boldsymbol{w}_t^{(1)}$ and $\boldsymbol{w}_t^{(2)}$ are therefore stationary Gaussian processes with some variances $\lambda_w^{(1)}$, $\lambda_w^{(2)}$ respectively and the spectra shown in Figure 4.

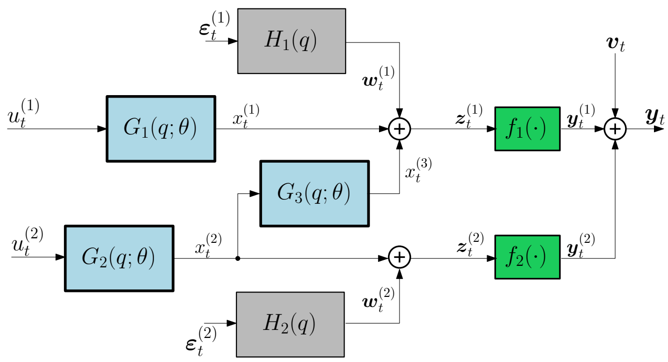

Fig. 3. A block diagram of the block-oriented dynamical network considered in Section 5. It has two known scalar inputs, $u_t^{(1)}$, $u_t^{(2)}$, and one measured scalar output $\boldsymbol{y}_t$.

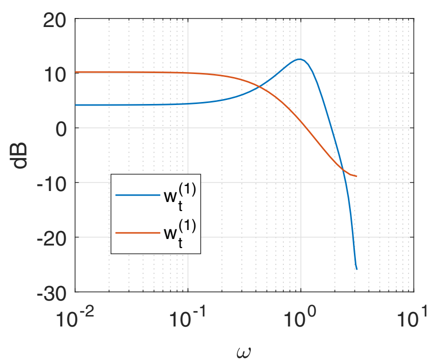

Fig. 4. The spectra of the process disturbances $\boldsymbol{w}_t^{(1)}$ (in blue) and $\boldsymbol{w}_t^{(2)}$ (in red) in the example of Section 5.

The signals $\{u_t^{(1)}\}$ and $\{u_t^{(2)}\}$ are given; they are independent realizations of

$$\boldsymbol{u}_t^{(1)} \sim \mathcal{N}(0, 0.5), \quad \boldsymbol{u}_t^{(2)} \sim \mathcal{N}(0, 4)$$

for all values of t, and the independent process

$$\boldsymbol{v}_t \sim \mathcal{N}(0,1), \quad \forall t$$

represents measurement noise. The output of the network for a given parameter θ can then be written as

$$\boldsymbol{y}_t(\theta) = \boldsymbol{y}_t^{(1)}(\theta) + \boldsymbol{y}_t^{(2)}(\theta) + \boldsymbol{v}_t, \quad t = 1,2,3,\ldots \tag{2}$$

where

$$\begin{aligned}
\boldsymbol{y}_t^{(1)}(\theta) &= \left(\boldsymbol{z}_t^{(1)}(\theta)\right)^2, \\
\boldsymbol{z}_t^{(1)}(\theta) &= x_t^{(1)}(\theta) + x_t^{(3)}(\theta) + \boldsymbol{w}_t^{(1)}, \\
x_t^{(1)}(\theta) &= G_1(q;\theta)u_t^{(1)}, \\
x_t^{(3)}(\theta) &= G_3(q;\theta)[G_2(q;\theta)u_t^{(2)}], \\
\boldsymbol{y}_t^{(2)}(\theta) &= \left(\frac{\boldsymbol{z}_t^{(2)}(\theta)}{20}\right)^3, \\
\boldsymbol{z}_t^{(2)}(\theta) &= x_t^{(2)}(\theta) + \boldsymbol{w}_t^{(2)}, \\
x_t^{(2)}(\theta) &= G_2(q;\theta)u_t^{(2)}.
\end{aligned} \tag{3}$$

Due to the assumptions on the basic stochastic processes $\boldsymbol{\varepsilon}_t^{(1)}$ and $\boldsymbol{\varepsilon}_t^{(2)}$, the measurement noise and the form of the static nonlinear functions, it is straightforward to compute the OE-type predictor of the network's output; it is given by the *closed-form expression*

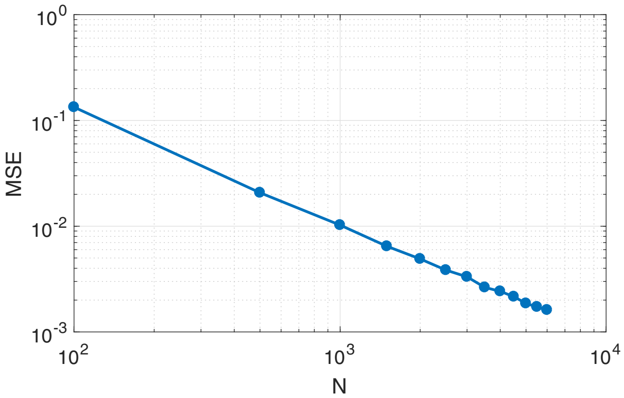

Fig. 5. Simulation results for the example of Section 5: The average MSE of the estimator is shown for different values of N (shown in log-log scale)

$$\hat{y}_t(\theta, \lambda_w^{(1)}, \lambda_w^{(2)}) = \mathbb{E}[\boldsymbol{y}_t^{(1)}; \theta] + \mathbb{E}[\boldsymbol{y}_t^{(2)}; \theta],$$

where

$$\mathbb{E}[\boldsymbol{y}_t^{(1)}; \theta, \lambda_w^{(1)}] = \left(x_t^{(1)}(\theta) + x_t^{(3)}(\theta)\right)^2 + 2\lambda_w^{(1)} \left(x_t^{(1)}(\theta) + x_t^{(3)}(\theta)\right),$$

$$\mathbb{E}[\boldsymbol{y}_t^{(2)}; \theta, \lambda_w^{(2)}] = \left(\frac{x_t^{(2)}(\theta)}{20}\right)^3 + 3\lambda_w^{(2)} \left(\frac{x_t^{(2)}(\theta)}{20}\right),$$

and the deterministic signals $x_t^{(1)}(\theta), x_t^{(2)}(\theta)$ and $x_t^{(3)}(\theta)$ are defined in (3). Observe that we used the assumption that the signals $\boldsymbol{w}_t$ are Gaussian, and therefore their moments are given in terms of their variances.

The OE-PEM estimator is then given by

$$\hat{\boldsymbol{\theta}}_N = \arg \min_{\theta, \lambda_w^{(1)}, \lambda_w^{(2)}} \sum_{t=1}^{N} (\boldsymbol{y}_t - \hat{y}_t(\theta, \lambda_w^{(1)}, \lambda_w^{(2)}))^2.$$

Observe that we also minimize (jointly) over the unknown variances of the disturbances $\boldsymbol{w}_t$, which can be seen as nuisance parameters.

To demonstrate the consistency of the resulting PEM estimator, we ran simulation studies for different values of N between 100 and 6000. For each value of N, the estimator is computed for 3000 independent realizations (over the inputs, the process disturbances and the measurement noise) using the Levenberg-Marquardt algorithm [1] . To help avoid possible local solutions, the true value θ° was used to initialize the algorithm in all cases.

The simulations results are summarized in Figures 6, 5 and Table 1. Figure 6 shows the average bias for different values of N; it is clear that the OE-PEM estimator is asymptotically unbiased. The average MSE of the estimator is shown using a log-log scale in Figure 5; the simulation results indicate the consistency of the OE-PEM estimator. Moreover, Table 1 shows the average and the standard deviation of the parameter estimates when $N = 600$.

6. CONCLUSIONS

In this contribution, we proposed a consistent PEM for the identification of dynamical networks involving nonlinear modules. The method is based on an OE-type predictor that is nonlinear in the known external signals. The major advantage of such a predictor is its simplicity; we showed that it is possible to obtain closed-form expressions in non-trivial cases. This is a great computational advantage. The performance of the proposed estimation method was illustrated on a numerical simulation example. The simulation results clearly indicate the consistency of the proposed PEM estimator. A detailed account of the convergence and consistency of the method will be considered by the authors in a future contribution.

[1] as implemented by the Matlab function `lsqnonlin`.

REFERENCES

Abdalmoaty, M.R. (2017). *Learning Stochastic Nonlinear Dynamical Systems Using Non-stationary Linear Predictors*. Licentiate thesis, KTH Royal Institute of Technology, Automatic Control Department.

Abdalmoaty, M.R. and Hjalmarsson, H. (2016). A simulated maximum likelihood method for estimation of stochastic Wiener systems. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, 3060–3065.

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

Doucet, A. and Johansen, A.M. (2009). A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of nonlinear filtering*, 12(656-704), 3.

Durbin, J. and Koopman, S.J. (1997). Monte carlo maximum likelihood estimation for non-gaussian state space models. *Biometrika*, 84(3), 669–684.

Everitt, N., Bottegal, G., Rojas, C.R., and Hjalmarsson, H. (2016). Identification of modules in dynamic networks: An empirical bayes approach. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, 4612–4617.

Forssell, U. and Ljung, L. (1999). Closed-loop identification revisited. *Automatica*, 35(7), 1215 – 1241.

Gevers, M., Bazanella, A.S., and Parraga, A. (2017). On the identifiability of dynamical networks. *IFAC-PapersOnLine*, 50(1), 10580 – 10585. 20th IFAC World Congress.

Giri, F. and Bai, E. (2010). *Block-oriented Nonlinear System Identification*. Springer.

Lamnabhi-Lagarrigue, F., Annaswamy, A., Engell, S., Isaksson, A., Khargonekar, P., Murray, R.M., Nijmeijer, H., Samad, T., Tilbury, D., and den Hof, P.V. (2017). Systems & control for the future of humanity, research agenda: Current and future roles, impact and grand challenges. *Annual Reviews in Control*, 43(Supplement C), 1 – 64.

Linder, J. and Enqvist, M. (2017). Identification and prediction in dynamic networks with unobservable nodes. *IFAC-PapersOnLine*, 50(1), 10574 – 10579. 20th IFAC World Congress.

Lindsten, F. (2013). An efficient stochastic approximation EM algorithm using conditional particle filters. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 6274–6278.

Ljung, L. (1978). Convergence analysis of parametric identification methods. *IEEE Transactions on Automatic Control*, 23(5), 770–783.

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

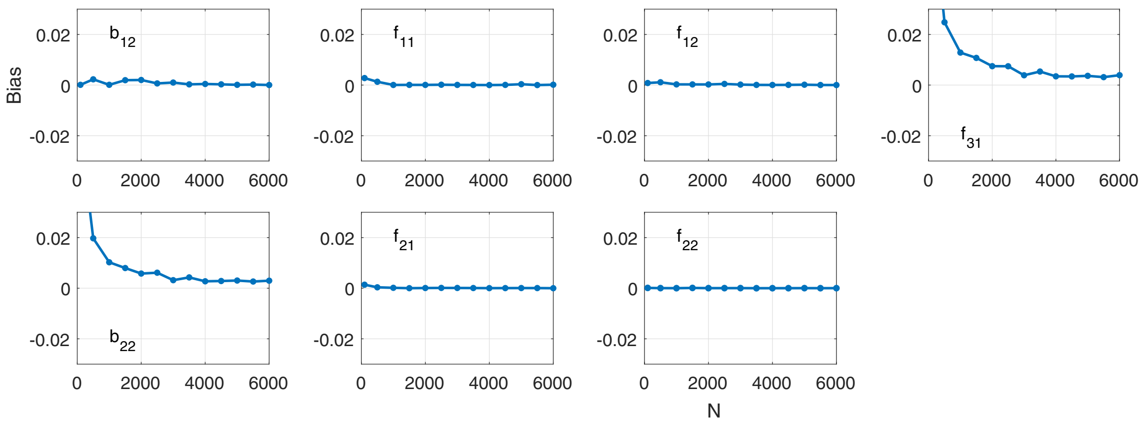

Fig. 6. Simulation results for the example of Section 5: the average bias of the estimator is shown for different values of N between 100 and 6000.

Table 1. The mean and the standard deviation of the estimated parameters when $N = 6000$. The values are approximated by averaging over 3000 independent MC realizations of the inputs, disturbances and measurement noise.

	b_{12}	f_{11}	f_{12}	b_{22}	f_{21}	f_{22}	f_{31}
true value	0.2	-1	0.8125	-0.2	1.6	0.8425	-0.7
estimates	0.2 ± 0.039	-1 ± 0.009	0.8125 ± 0.0076	-0.197 ± 0.023	1.6 ± 0.001	0.8425 ± 0.0011	-0.696 ± 0.0296

Materassi, D. and Innocenti, G. (2010). Topological identification in networks of dynamical systems. *IEEE Transactions on Automatic Control*, 55(8), 1860–1871.

Materassi, D., Innocenti, G., Giarr, L., and Salapaka, M. (2013). Model identification of a network as compressing sensing. *Systems & Control Letters*, 62(8), 664 – 672.

Materassi, D. and Salapaka, M.V. (2012a). Network reconstruction of dynamical polytrees with unobserved nodes. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, 4629–4634.

Materassi, D. and Salapaka, M.V. (2012b). On the problem of reconstructing an unknown topology via locality properties of the wiener filter. *IEEE Transactions on Automatic Control*, 57(7), 1765–1777.

Ninness, B., Wills, A., and Schön, T. (2010). Estimation of general nonlinear state-space systems. In *49th IEEE Conference on Decision and Control, Atlanta, Georgia, USA*, 1–6.

Sanandaji, B.M., Vincent, T.L., and Wakin, M.B. (2011). Exact topology identification of large-scale interconnected dynamical systems from compressive observations. In *Proceedings of the 2011 American Control Conference*, 649–656.

Schön, T.B., Lindsten, F., Dahlin, J., Wågberg, J., Naesseth, C.A., Svensson, A., and Dai, L. (2015). Sequential Monte Carlo methods for system identification. *IFAC-PapersOnLine*, 48(28), 775 – 786. 17th IFAC Symposium on System Identification SYSID 2015 Beijing, China, 1921 October 2015.

Schön, T.B., Wills, A., and Ninness, B. (2011). System identification of nonlinear state-space models. *Automatica*, 47(1), 39 – 49.

Schoukens, J., Marconato, A., Pintelon, R., Rolain, Y., Schoukens, M., Tiels, K., Vanbeylen, L., Vandersteen, G., and Mulders, A.V. (2014). System identification in a real world. In *2014 IEEE 13th International Workshop on Advanced Motion Control (AMC)*, 1–9.

Van den Hof, P.M.J., Schrama, R.J.P., and Bosgra, O.H. (1992). An indirect method for transfer function estimation from closed loop data. In *[1992] Proceedings of the 31st IEEE Conference on Decision and Control*, 1702–1706 vol.2.

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

Weerts, H.H., Dankers, A.G., and den Hof, P.M.V. (2015). Identifiability in dynamic network identification. *IFAC-PapersOnLine*, 48(28), 1409 – 1414. 17th IFAC Symposium on System Identification SYSID 2015.

Weerts, H.H., den Hof, P.M.V., and Dankers, A.G. (2016). Identifiability of dynamic networks with part of the nodes noise-free. *IFAC-PapersOnLine*, 49(13), 19 – 24. 12th IFAC Workshop on Adaptation and Learning in Control and Signal Processing ALCOSP 2016.

Wills, A., Schön, T.B., Ljung, L., and Ninness, B. (2013). Identification of Hammerstein-Wiener models. *Automatica*, 49(1), 70–81.