# On Identification of Networked Systems with Time-invariant Topology

**Mohsen Zamani** [*] **Brett Ninness** [**] **Juan C. Agüero** [****]

[*] *School of Electrical Engineering and Computer Science, The University of Newcastle, Callaghan, NSW 2308, Australia. (e-mail: mohsen.zamani@newcastle.edu.au).*
[**] *School of Electrical Engineering and Computer Science, The University of Newcastle, Callaghan, NSW 2308, Australia. (e-mail: brett.ninness@newcastle.edu.au).*
[***] *Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile. (email:juan.aguero@usm.cl).*
[****] *School of Electrical Engineering and Computer Science, The University of Newcastle, Callaghan, NSW 2308, Australia. (e-mail: juan.aguero@newcastle.edu.au).*

**Abstract:** This paper studies the problem of identification for networked systems. We consider both heterogeneous and homogeneous networks. It is assumed that the interconnection topology is time-invariant and known. We demonstrate how the parameter matrices of each individual subsystem can be identified from the input-output information obtained from the whole network. To this end, the well-known technique of Maximum-likelihood (ML) is exploited for obtaining estimations of system matrices.

*Keywords:* Identification, Networked systems, Maximum likelihood, Time-invariant topology

## 1. INTRODUCTION

Networked systems have become fundamental elements in humans life. Due to their complexity and profound applications, in different branches of science and technology researchers have studied these systems from different perspectives. In particular, within the field of systems and control researchers have significantly contributed in our understandings of these systems see for e.g. Hespanha et al. (2007), Ren et al. (2007), Zhang et al. (2001). In systems and control literature networked systems have been explored from different directions [1] such as distributed control Sinopoli et al. (2003), Olfati-Saber and Murray (2002), security Amin et al. (2009), Cárdenas et al. (2011), coordination control Anderson et al. (2008), Zamani and Lin (2009), Ji and Egerstedt (2007), dynamical analysis Zamani et al. (2013), Fax and Murray (2004), Fuhrmann and Helmke (2013), Zamani et al. (2014) and system identification Dankers (2014).

The current paper is particularly concerned with the identification problem for networked systems. It is worthwhile mentioning that problems associated with identification of networked systems go beyond common issues discussed in classical system identification literature. Firstly, it may not be feasible due to practical or security concerns to obtain measurements from one or several subsystems within a network. Secondly, there may exist a causality matter raising due to the interconnection topology. Thirdly, as networks are increasing both in number of subsystems and connecting links, any algorithm designed for identifying such systems must be able to handle this issue.

Despite these motivations, the problem of identification for networked systems has only recently attracted interest in systems and control community. We now briefly review some of them. The authors in Haber and Verhaegen (2013) proposed a distributed version of moving horizon estimation method for series networks. The same set of authors in Haber and Verhaegen (2014), developed a subspace-based algorithm for identification of parameters for the same type of networks. The reconstruction of a cascade system from a given unstructured system estimate was explored in Sandberg et al. (2014).

The authors of Dankers et al. (2013) extended direct method in classical closed-loop identification for tackling the identification problem for networked systems. The same problem was researched in Van den Hof et al. (2012) for networked systems with known topology. The work Sanandaji et al. (2011) explored sparse networks and used compressive sensing techniques to identify the network topology.

The authors of Fazlyaband and Preciado (2014) exploited adaptive control techniques to identify the network topology, when it is time-invariant, from input and state measurements. Moreover, the problem of topology identification was studied in Shahrampour and Preciado (2014) using spectral analysis approach. The authors of Hayden et al. (2014) examined the identification of networked systems where there exists an unknown intrinsic noise.

[1] These directions are not necessarily mutually exclusive.

In the networked systems literature researchers often start their analysis by considering the state space representation (usually linear and time-invariant) that they claim models the whole network see e.g. Fazlyaband and Preciado (2014), Shahrampour and Preciado (2014) and Hayden et al. (2014). The dynamics of each subsystem is usually assumed to be a simple integrator [2]. Furthermore, the state matrix usually is equivalent to the adjacency matrix. *In this paper, we propose a formulation that is effective and provides a good insight into the identification problem for networked systems. We assume that each subsystem is finite-dimensional but do not impose any restriction on its dimension size. Our proposed formulation, enables us to transform the problem of identification for networked systems into the well-known problem of identification of linear time-invariant (LTI) systems that there exist very well-developed tools to handle.*

In the current paper, we assume that the interconnection topology is known and time-invariant. Furthermore, the whole networked system is stable. Then our objective is to find estimates for parameter matrices of each individual subsystem using measurements obtained from one or more of them.

To this end, we use the well-known Maximum-likelihood (ML) technique. It is important to note that standard approaches to identify systems in state-space form such as Gibson and Ninness (2005), Agüero et al. (2012) and Shumway and Stoffer (1982) are not practical for the problem of interest in this paper. In fact, Expectation-maximization (EM) estimation algorithms to identify state-space systems typically utilize fully parametrized matrices. However, here there exist linear relations between entries of the network parameter matrices [3]. Hence, we use the gradient-based algorithm to find solutions for the ML problem.

The structure of this paper is as follows. In the next section, the problem formulation is introduced. Then in Section 3, we briefly review the ML approach. Section 4 introduces a gradient-based solution for the ML estimation problem. In Section, 5 we present a numerical example. Finally, Section 6 provides the concluding remarks and plans for future work.

## 2. PROBLEM FORMULATION

We consider networks of $M$ linear systems, coupled through constant interconnection parameters. Each subsystem is assumed to have the state-space representation as a linear discrete-time system

$$\begin{aligned} x_{t+1}^i &= A_i x_t^i + B_i v_t^i \\ w_t^i &= C_i x_t^i,\ i = 1, \ldots, M. \end{aligned} \tag{1}$$

Here, $A_i \in \mathbb{R}^{n_i \times n_i}$, $B_i \in \mathbb{R}^{n_i \times m_i}$ and $C_i \in \mathbb{R}^{p_i \times n_i}$ are the associated system matrices. We assume that each system is reachable and observable and that the subsystems are interconnected by static coupling laws

$$v_t^i = \sum_{j=1}^{M} L_{ij} w_t^j + R_i u_t + d_t^i \in \mathbb{R}^{m_i}$$

with $L_{ij} \in \mathbb{R}^{m_i \times p_j}$, $R_i \in \mathbb{R}^{m_i \times m}$ and $u_t \in \mathbb{R}^m$. The vectors $u_t$ and $d_t^i$ denote the external input to the whole network and the disturbance signal affecting each individual subsystem. Further, we assume that there is a $p$-dimensional interconnected output given by

$$y_t = \sum_{i=1}^{M} S_i w_t^i + D u_t + e_t \quad \text{with} \quad S_i \in \mathbb{R}^{p \times p_i},$$

where $i = 1, \ldots, M$ and $e_t$ is the measurement noise.

Define $\overline{m} = \sum_{i=1}^M m_i$, $\overline{p} = \sum_{i=1}^M p_i$, $\overline{n} = \sum_{i=1}^M n_i$ and the *coupling matrices*

$$\begin{aligned} L &= [L_{ij}]_{ij} \in \mathbb{R}^{\overline{m} \times \overline{p}} \\ R &= \begin{bmatrix} R_1^\top & \ldots & R_M^\top \end{bmatrix}^\top \in \mathbb{R}^{\overline{m} \times m} \\ S &= [S_1 \ \ldots \ S_M] \in \mathbb{R}^{p \times \overline{p}} \\ D &\in \mathbb{R}^{p \times m} \end{aligned}$$

as well as *node matrices*

$$\begin{aligned} A &= \mathrm{diag}\,(A_1, \ldots, A_M) \\ B &= \mathrm{diag}\,(B_1, \ldots, B_M) \\ C &= \mathrm{diag}\,(C_1, \ldots, C_M), \\ x_t &:= \begin{bmatrix} x_t^1 \\ \vdots \\ x_t^M \end{bmatrix} \in \mathbb{R}^{\overline{n}} \\ d_t &:= \begin{bmatrix} d_t^1 \\ \vdots \\ d_t^M \end{bmatrix} \in \mathbb{R}^{\overline{m}}. \end{aligned} \tag{2}$$

Then the networked system may be formulated as

$$\begin{aligned} x_{t+1} &= \mathbf{A} x_t + \mathbf{B} u_t + B d_t \\ y_t &= \mathbf{C} x_t + D u_t + e_t, \end{aligned} \tag{3}$$

with matrices

$$\mathbf{A} := A + BLC \quad \mathbf{B} := BR, \quad \mathbf{C} := SC. \tag{4}$$

*The formulation above is useful as it expresses a network as the LTI model (3) that contains both subsystem and topology parameters.*

Fig. 1 demonstrates how each subsystem is connected to others and command signal propagates through the network. It is worthwhile mentioning that from control design point of view, the matrix $L$ behaves as a form of output feedback operator for the whole network. In this paper, we assume that while the topology is known; the system matrices for each subsystem are not. It is desirable to find suitable estimates for these parameters using the available information; viz. $y_t$ and $u_t$.

We now state the assumptions used in this paper.

(1) The system (3) is Hurwitz.
(2) The measurement noise process $e_t$ is i.i.d zero-mean Gaussian processes. Furthermore, $\mathbb{E}\left[e_t e_q^\top\right] = O\delta_{tq}$ for all $t, q$, where $\delta_{tq}$ is the Kronecker delta, which is 1 for $t = q$ and 0 otherwise.
(3) The disturbance signals $d_t^i$, $i = 1, \ldots, M$, are i.i.d zero-mean jointly Gaussian processes and $\mathbb{E}\left[d_t d_q^\top\right] = Q\delta_{tq}$.

[2] Despite the usual claim, we believe that the generalization of the existing results into a case with higher dimensions becomes a nontrivial task.

[3] This is imposed due to the interconnection existing among subsystems

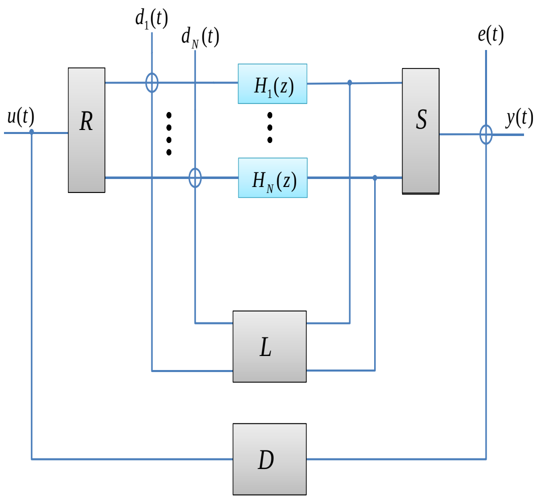


Fig. 1. *The whole networked system.*

We further assume that unknown entries in the system matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ and the disturbance and noise covariance matrices i.e. $O$ and $Q$ are appropriately parameterized by the elements of a vector $\theta \in \mathbb{R}^{n_\theta}$. Details on how this may be achieved are provided in what follows.

Then given the model structure (3), and data records

$$\begin{aligned} U_N &:= [u_1 \ \dots \ u_N], \\ Y_N &:= [y_1 \ \dots \ y_N], \end{aligned} \tag{5}$$

our objective is to find an estimate of the vector $\theta$ such that the input-output series produced by the associated model best match $Y_N$ (in sense that will be explained later). To this end, we use the well-known ML technique Ljung (1987) which is shortly explained in the next section.

## 3. MAXIMUM-LIKELIHOOD ESTIMATION

The ML technique is very well-developed see e.g. Ljung (1987) and Söderström and Stoica (1988). We now in short review this approach for the purpose of the current paper. The ML technique requires the formulation of the joint density $p_\theta(Y_N)$ of the observations, which may be expressed using Bayes' rule as

$$p_\theta(Y_N) = p_\theta(y_0) \prod_{t=1}^{N} p_\theta(y_t \mid y_{t-1}, \cdots, y_1). \tag{6}$$

A maximum likelihood estimate $\widehat{\theta}_{\mathrm{ML}}$ of the parameter vector $\theta$ is then defined as any value $\widehat{\theta}_{\mathrm{ML}}$ satisfying

$$\widehat{\theta}_{\mathrm{ML}} \in \{\theta \in \Theta : \mathcal{L}(\theta) \le \mathcal{L}(\theta'), \forall \theta' \in \Theta\}, \tag{7}$$

where

$$\mathcal{L}(\theta) \triangleq -\log p_\theta(Y_N) \tag{8}$$

and $\Theta \subseteq \mathbb{R}^{n_\theta}$ is a user chosen compact subset. In the above, the common notation $p_\theta(Y_N)$ is the likelihood of the observed data $Y_N$ given a choice $\theta$ of the parameterization of the system (3). Furthermore, note that while the form of $p_\theta(Y_N)$ will depend on $U_N$, it is not formally an argument since it is either a deterministic process or its data generating mechanism is independent of the choice of $\theta$ (see Goodwin and Payne (1977) page 115). Since the by assumption $e_t$ and $d_t$ are Gaussian processes, the same holds for $y_t$ and $x_t$ (Anderson and Moore (1979) page 18). Hence, the density $p_\theta(y_t \mid Y_{t-1})$ will also be a Gaussian density and lead to, save for some constants that do not affect the minimization (7) Åström (1980),

$$\mathcal{L}(\theta) = \sum_{t=1}^{N} \log\det\{\Lambda_t(\theta)\} + \epsilon_t^\top(\theta)\Lambda_t(\theta)^{-1}\epsilon_t(\theta) \tag{9}$$

where

$$\begin{aligned} \epsilon_t(\theta) &= y_t - \widehat{y}_{t|t-1}(\theta), \\ \widehat{y}_{t|t-1}(\theta) &:= \mathbb{E}\left[y_t|Y_{t-1},\theta\right] \\ \Lambda_t(\theta) &:= \mathbb{E}\left[\epsilon_t(\theta)\epsilon_t(\theta)^\top|Y_{t-1},\theta\right]. \end{aligned} \tag{10}$$

To compute the above quantities, one can use the well-known Kalman filter Anderson and Moore (1979)

$$\begin{aligned} \widehat{y}_{t|t-1}(\theta) &= \mathbf{C}_\theta \widehat{x}_{t|t-1}(\theta) + Du_t, \\ \widehat{x}_{t+1|t}(\theta) &= \mathbf{A}_\theta \widehat{x}_{t|t-1}(\theta) + \mathbf{B}_\theta u_t + K_\theta \epsilon_t(\theta), \\ K_\theta &= \left(\mathbf{A}_\theta P_{t|t-1}(\theta)\mathbf{C}_\theta^\top\right)\Lambda_t^{-1}(\theta), \\ P_{t+1|t} &= \mathbf{A}_\theta P_{t|t-1}(\theta)\mathbf{A}_\theta^\top + B_\theta Q_\theta B_\theta^\top - K_\theta \Lambda_t(\theta) K_\theta^\top, \\ \Lambda_t(\theta) &= \mathbf{C}_\theta P_{t|t-1}(\theta)\mathbf{C}_\theta^\top + O_\theta, \end{aligned} \tag{11}$$

where

$$\begin{aligned} \widehat{x}_{t|t-1} &:= \mathbb{E}\left[x_t|Y_{t-1},\theta\right] \\ P_{t|t-1}(\theta) &:= \left[\left(x_t - \widehat{x}_{t|t-1}\right)\left(x_t - \widehat{x}_{t|t-1}\right)^\top|Y_{t-1},\theta\right]. \end{aligned} \tag{12}$$

## 4. PARAMETER ESTIMATION

In this section, we estimate the unknown parameters of the system (3) via the ML method (6)-(12). Note that the system matrices of (3) are very well-structured. We now exploit the Gauss-Newton approach to solve the minimization (7).

The Gauss-Newton approach requires that we run a time-varying Kalman filter in order to evaluate the cost in (9). It also requires that we run a time-varying Kalman filter for each gradient computation.

Consider the closed-loop system (3). We parameterize the unkonw entries of as

$$\begin{aligned} \theta &:= \left[\alpha^\top \ \beta^\top \ \psi^\top \ \omega^\top\right]^\top \\ \alpha &:= \mathrm{vec}\left[A_1, \dots, A_M\right] \\ \beta &:= \mathrm{vec}\left[B_1, \dots, B_M\right] \\ \psi &:= \mathrm{vec}\left[C_1, \dots, C_M\right] \\ \omega &:= \mathrm{vec}\left[Q, O\right] \end{aligned} \tag{13}$$

Now given $N$ observations of the input $u_t$ and output $y_t$, we would like to form an expression for the likelihood of seeing these output observations as a function of the parameters $\theta$. As explained earlier, such an expression can then be maximized over the unknown parameters to find the ML estimate. To this end, one can equivalently minimize the negative logarithm [4] of that likelihood, as expressed in (9).

[4] Note that the logarithm is a monotonic function and is numerically more preferable.

In general, $\mathcal{L}(\theta)$ cannot be minimized in closed-form and we will require an iterative procedure to find the minimum. Possibly the most popular way to achieve this is to use gradient-based search techniques. An essential ingredient of these methods involves the computation of the gradient (and quite possibly higher order terms, which will be discussed later). In terms of the gradient, it is straightforward to derive that [5]

$$
\begin{aligned}
\frac{\partial \mathcal{L}(\theta)}{\partial \theta} &= \sum_{t=1}^{N} \operatorname{tr}\left(\Lambda_t^{-1}\frac{\partial \Lambda_t}{\partial \theta}\right) + 2\frac{\partial \epsilon_t^\top}{\partial \theta}\Lambda_t^{-1}\epsilon_t \\
&\quad -\epsilon_t^\top \Lambda_t^{-1}\frac{\partial \Lambda_t}{\partial \theta}\Lambda_t^{-1}\epsilon_t. \qquad (14)
\end{aligned}
$$

These derivative terms can be found recursively via

$$
\begin{aligned}
\frac{\partial \mathcal{L}(\theta)}{\partial \theta} &:= \sum_{t=1}^{N} \operatorname{tr}\left(\Lambda_t^{-1}\frac{\partial \Lambda_t}{\partial \theta}\right) + 2\frac{\partial \epsilon_t^\top}{\partial \theta}\Lambda_t^{-1}\epsilon_t \\
&\quad -\epsilon_t^\top \Lambda_t^{-1}\frac{\partial \Lambda_t}{\partial \theta}\Lambda_t^{-1}\epsilon_t, \\
\frac{\partial \epsilon_t}{\partial \theta} &= -\frac{\partial \mathbf{C}}{\partial \theta}\widehat{x}_{t|t-1} - \mathbf{C}\frac{\partial \widehat{x}_{t|t-1}}{\partial \theta} \\
\frac{\partial \widehat{x}_{t+1|t}}{\partial \theta} &= \frac{\partial \mathbf{A}}{\partial \theta}\widehat{x}_{t|t-1} + \mathbf{A}\frac{\partial \widehat{x}_{t|t-1}}{\partial \theta} + \frac{\partial \mathbf{B}}{\partial \theta}u_t + \frac{\partial K_t}{\partial \theta}\epsilon_t \\
&\quad + K_t\frac{\partial \epsilon_t}{\partial \theta}, \\
\frac{\partial K_t}{\partial \theta} &= \Bigg(\frac{\partial \mathbf{A}}{\partial \theta}P_{t|t-1}\mathbf{C}^\top + \mathbf{A}\frac{\partial P_{t|t-1}}{\partial \theta}\mathbf{C}^\top \\
&\quad + \mathbf{A}P_{t|t-1}\frac{\partial \mathbf{C}^\top}{\partial \theta}\Lambda_t^{-1} - K_t\frac{\partial \Lambda_t}{\partial \theta}\Bigg)\Lambda_t^{-1}, \\
\frac{\partial P_{t+1|t}}{\partial \theta} &= \frac{\partial \mathbf{A}}{\partial \theta}P_{t|t-1}\mathbf{A}^\top + \mathbf{A}\frac{\partial P_{t|t-1}}{\partial \theta}\mathbf{A}^\top \\
&\quad + \mathbf{A}P_{t|t-1}\left(\frac{\partial \mathbf{A}}{\partial \theta}\right)^\top + \frac{\partial B}{\partial \theta}QB^\top + B\frac{\partial Q}{\partial \theta}B^\top \\
&\quad + BQ\left(\frac{\partial B}{\partial \theta}\right)^\top - \frac{\partial K_t}{\partial \theta}\Lambda_t K_t^T - K_t\frac{\partial \Lambda_t}{\partial \theta}K_t^\top \\
&\quad - K_t\Lambda_t\left(\frac{\partial K_t}{\partial \theta}\right)^\top, \\
\frac{\partial \Lambda_t}{\partial \theta} &= \frac{\partial \mathbf{C}}{\partial \theta}P_{t|t-1}\mathbf{C}^T + \mathbf{C}\frac{\partial P_{t|t-1}}{\partial \theta}\mathbf{C}^\top + \mathbf{C}P_{t|t-1}\frac{\partial \mathbf{C}^\top}{\partial \theta} \\
&\quad + \frac{\partial O}{\partial \theta}.
\end{aligned}
$$

We now provide an algorithm that demonstrates required steps for estimation of $\theta$.

**Algorithm 1** Algorithm for estimation of $\theta$

initialize $\theta$
**repeat**
compute Kalman filter (11)
compute (14)
update $\theta$
**until** convergence of $\theta$

[5] As it is clear that all the expressions depend on the choice of $\theta$, in the upcoming parts we omit $\theta$.

There are several approaches that can be used for implementing the update step in Algorithm 1 see e.g. Chi et al. (2006). These approaches are divided into two main groups of first order methods and second order ones. It is not recommended to use first order approaches as they only use the gradient by itself that lead to a poor performance of the search algorithm. In the second order approaches, the gradient is generally scaled by a positive-definite matrix $H_i$, which sometimes turns out to be the inverse of Hessian matrix, in order to improve the efficiency of the search procedure. Thus, in this paper we use the following update role

$$
\theta_{i+1} = \theta_i - H_i\frac{\partial \mathcal{L}(\theta)}{\partial \theta_i} \qquad (15)
$$

where $\theta_i$ represents $\theta$ at the $i-$th search step.

The matrix $H_i$ can be assigned in several ways. We now discuss one possibility for this assignment; viz. Quasi-Newton approach. This approach iteratively delivers the matrix $H_i$ as an approximation for the inverse of Hessian matrix.

### *4.1 Quasi-Newton Method*

There exist variations of Quasi-Newton method. Here, we briefly discuss one of the most effective ones called Broyden-Fletcher-Goldfarb-Shanno (BFGS) techinque. In this approach matrix $H_i$ turns about to be an approximation for the Hessian matrix and is computed in an iterative way.

The BFGS method starts with a positive definite proposal for initialization of $H_i$. Then this matrix is updated at each step based on the following equations Chi et al. (2006).

$$
\begin{aligned}
\delta_i &= \theta_i - \theta_{i-1}, && (16)\\
\gamma_i &= \frac{\partial \mathcal{L}(\theta)}{\partial \theta_i} - \frac{\partial \mathcal{L}(\theta)}{\partial \theta_{i-1}}, && (17)\\
\xi_i &= \frac{\gamma_i^\top H_{i-1}\gamma_i}{\gamma_i^\top \delta_i}, && (18)\\
H_i &= H_{i-1} + \frac{1}{\delta_i^\top \gamma_i}\left[(1+\xi_i)\gamma_i\gamma_i^\top - \delta_i\gamma_i^\top H_{i-1}\right. && (19)\\
&\quad \left. -H_{i-1}\gamma_i\delta_i^\top\right].
\end{aligned}
$$

The matrix $H_i$ is required to be a positive definite matrix. The following theorem proves that under the update law (19), it remains to be positive definite after being initialized by a positive definite candidate.

*Lemma 1.* Suppose that the matrix $H_{i-1}$ is positive definite, then the matrix $H_i$ generated under the update role (19) is also positive definite given that $\delta_i$ and $\gamma_i$ are nonzero vectors.

**Proof.** Refer to Appendix2F in Chi et al. (2006).

## 5. NUMERICAL SIMULATION

In this section, we consider a network comprising of three subsystems and the topology as in Fig. 2.

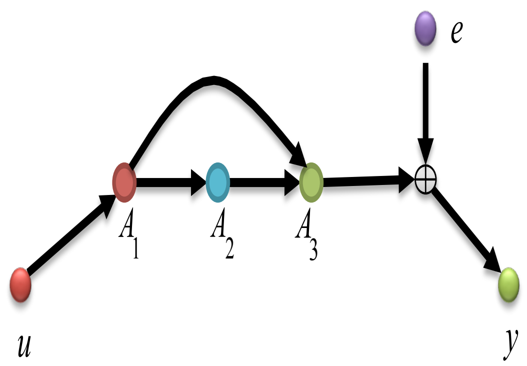


Fig. 2. *A network consisting of three SISO subsystems. The subsystems, the external input, measurement and its associated noise are depicted by symbols $A_i$, $i = 1, 2, 3$, $u$ , $y$ and $e$, accordingly. The whole network is stable and relative weights are set to be unity.*

We suppose that each subsystem has the following simple dynamics.

$$\begin{aligned} x_{t+1}^i &= A_i x_t^i + B_i v_t^i \\ w_t^i &= x_t^i,\ i = 1, \ldots, 3. \end{aligned} \tag{20}$$

where $A_1 = -0.2$, $A_2 = -0.6$, $A_3 = -0.8$, $B_1 = 1.5$, $B_2 = 3.4$ and $B_3 = 2.6$. The linking paramaters are set to be unity i.e. $L = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}$, $R = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $S = [0\ 0\ 1]$. Then it is easy to verify that the whole networked system depectid in Fig. 2 can be written as

$$\begin{aligned} \begin{bmatrix} x_{t+1}^1 \\ x_{t+1}^2 \\ x_{t+1}^3 \end{bmatrix} &= \begin{bmatrix} -0.2 & 0 & 0 \\ 3.4 & -0.6 & 0 \\ 1.5 & 2.6 & -0.8 \end{bmatrix} \begin{bmatrix} x_t^1 \\ x_t^2 \\ x_t^3 \end{bmatrix} + \begin{bmatrix} 1.5 \\ 0 \\ 0 \end{bmatrix} u_t \\ y_t &= [0\ 0\ 1]\, x_t + e_t \end{aligned} \tag{21}$$

where $e_t$ is a zero-mean Guassian noise and $\mathbb{E}\left[e_t e_t^\top\right] = 0.1$. The parameter estimation is done with the help of the package introduced in Ninness et al. (2013). The frequency domain and the time domain responses for both estimated system and the actual system are depicted in Fig. 3, Fig. 5 and Fig. 4, respectively. It is very clear from the simulation results that the response of the estimated model is a close match to that of the actual system.

## 6. CONCLUSION

In this paper, the problem of identification for networked systems with time-invariant interconnection topology was studied. The topology was assumed to be available and chosen such that the whole network becomes stable. The ML technique was employed for obtaining estimations of parameter matrices. Finally, solutions of the ML problem were attained through a gradient-based search. There are several important questions that can be considered as a part of future research study. For instance, as seen in Section 5, the interconnection topology in networked systems has a sparse structure and this constraint needs to be contemplated in the identification of the whole system. Furthermore, this paper only examined the problem using

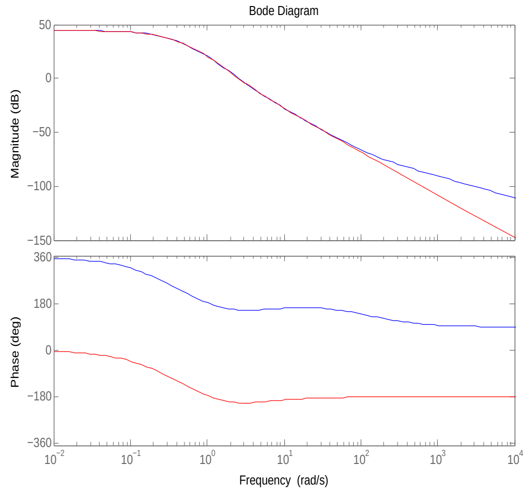


Fig. 3. *The bode diagram of the actual system* (21) *and that of the associated estimated system are sketched in blue and red colors, accordingly.*

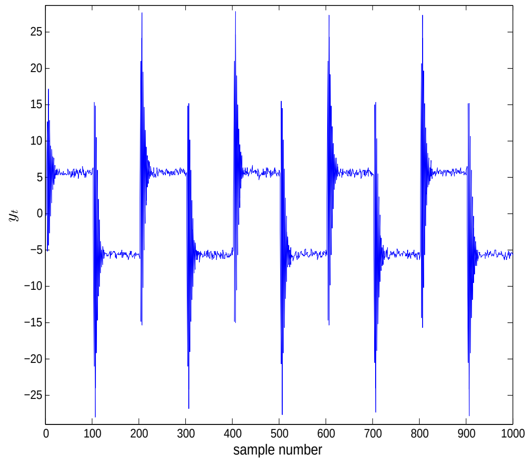


Fig. 4. *The output response of the system* (21).

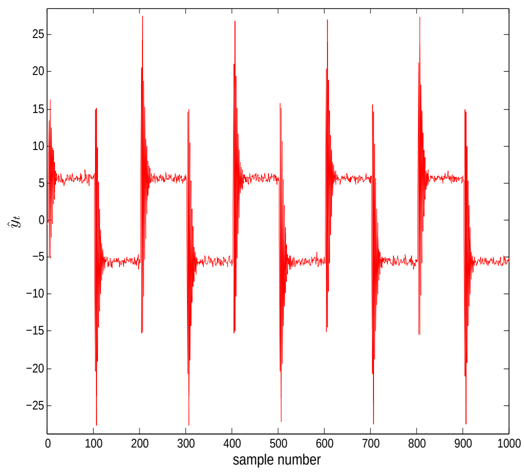


Fig. 5. *The output response of the estimated model.*

one of the possible available techniques that is not scalable. It would be very interesting to study this problem using other available approaches that are able to handle this issue.

## ACKNOWLEDGEMENTS

Supports by the Australian Research Council, the National Natural Science Foundation of China under grant number 61203119 and the Chilean Research Council (CONICYT) through Basal Project FB0008, and FONDECYT (grant no. 1150954) are gratefully acknowledged.

## REFERENCES

Agüero, J., Tang, W., Yuz, J.I., Delgado, R., and Goodwin, G.C. (2012). Dual timefrequency domain system identification. *Automatica*, 48(12), 3031–3041.

Amin, S., Cárdenas, A.A., and Sastry, S. (2009). Safe and secure networked control systems under denial-of-service attacks. In *Proc Hybrid Systems: Computation and Control*, 31–45. Springer.

Anderson, B.D.O. and Moore, J.B. (1979). *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall.

Anderson, B.D.O., Yu, C., Fidan, B., and Hendrickx, J.M. (2008). Rigid graph control architectures for autonomous formations. *IEEE Control Systems*, 28(6), 48–63.

Åström, K. (1980). Maximum likelihood and prediction error methods. *Automatica*, 16(5), 551–574.

Cárdenas, A.A., Amin, S., Lin, Z., Huang, Y., Huang, C., and Sastry, S. (2011). Attacks against process control systems: risk assessment, detection, and response. In *Proc ACM Symposium on Information, Computer and Communications Security*, 355–366.

Chi, C.Y., , Feng, C.C., Chen, C.H., and Chen, C.Y. (2006). Blind equalization and system identification. *London: SpringerVerlag*.

Dankers, A.G. (2014). *System Identification in Dynamic Networks*. Ph.D. thesis, TU Delft, Delft University of Technology.

Dankers, A.G., Van den Hof, P.M.J., and Heuberger, P.S.C. (2013). Predictor input selection for direct identification in dynamic networks. In *Proc the IEEE Conference on Decision and Control*, 4541–4546.

Fax, J.A. and Murray, R.M. (2004). Information flow and cooperative control of vehicle formations. *IEEE Transactions on Automatic Control*, 49(9), 1465–1476.

Fazlyaband, M. and Preciado, V.M. (2014). Robust topology identification and control of lti networks. *arXiv preprint arXiv:1406.4201*.

Fuhrmann, P.A. and Helmke, U.R. (2013). Observability and strict equivalence of networked linear systems. *Mathematics of Control, Signals, and Systems*, 25(4), 437–471.

Gibson, S. and Ninness, B. (2005). Robust maximum-likelihood estimation of multivariable dynamic systems. *Automatica*, 41(10), 1667–1682.

Goodwin, G. and Payne, L. (1977). *Dynamic system identification: experiment design and data analysis*. Academic press.

Haber, A. and Verhaegen, M. (2013). Moving horizon estimation for large-scale interconnected systems. *IEEE Transactions on Automatic Control*, 58(11), 2834–2847.

Haber, A. and Verhaegen, M. (2014). Subspace identification of large-scale interconnected systems. *IEEE Transactions on Automatic Control*, 59(10), 2754–2759.

Hayden, D., Yuan, Y., and Goncalves, J. (2014). Network reconstruction from intrinsic noise: Non-minimum-phase systems. In *Proc of the 19th IFAC World Congress*.

Hespanha, P.J., Naghshtabrizi, P., and Xu, Y. (2007). A survey of recent results in networked control systems. *IEEE Proceedings*, 95(1), 138–162.

Ji, M. and Egerstedt, M. (2007). Distributed coordination control of multi-agent systems while preserving connectedness. *IEEE Transactions on Robotics*, (4).

Ljung, L. (1987). *System Identification Theory for the User*. Prentice Hall, New Jersey.

Ninness, B., Wills, A., and Mills, A. (2013). Unit: A freely available system identification toolbox. *Control Engineering Practice*, 21(5), 631–644.

Olfati-Saber, R. and Murray, R.M. (2002). Distributed cooperative control of multiple vehicle formations using structural potential functions. In *IFAC World Congress*, 346–352.

Ren, W., Beard, R.W., and Atkins, E.M. (2007). Information consensus in multivehicle cooperative control. *IEEE Control Systems Magazine*, 27(2), 71–82.

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

Sandberg, H., Högg, P., and Wahlberg, B. (2014). Approximative model reconstruction of cascade systems. *Systems & Control Letters*, 69, 90–97.

Shahrampour, S. and Preciado, V.M. (2014). Topology identification of directed dynamical networks via power spectral analysis. *IEEE Transactions on Automatic Control*, (99).

Shumway, R. and Stoffer, D. (1982). An approach to time series smoothing and forecasting using the em algorithm. *Journal of Time Series Analysis*, 3(4), 253–264.

Sinopoli, B., Sharp, C., Schenato, L., Schaffert, S., and Sastry, S. (2003). Distributed control applications within sensor networks. *Proceedings of the IEEE*, 91(8), 1235–1246.

Söderström, T. and Stoica, P. (eds.) (1988). *System Identification*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Van den Hof, P.M.J., Dankers, A.G., Heuberger, P., and Bombois, X. (2012). Identification in dynamic networks with known interconnection topology. In *Proc the IEEE Conference on Decision and Control*, 895–900.

Zamani, M., Helmke, U., and Anderson, B.D.O. (2014). Zeros of networked systems with time-invariant interconnections. To appear in Automatica (http://arxiv.org/abs/1408.6889).

Zamani, M., Helmke, U., and Anderson, B. (2013). Zeros of networks of linear multi-agent systems: Time-invariant interconnections. In *Proc the IEEE 52nd Annual Conference on Decision and Control*, 5939–5944.

Zamani, M. and Lin, H. (2009). Structural controllability of multi-agent systems. In *Proc the American Control Conference*, 5743–5748.

Zhang, W., Branicky, M.S., and Phillips, S.M. (2001). Stability of networked control systems. *IEEE Control Systems*, 21(1), 84–99.