Subspace identification of local systems in 1D homogeneous networks

Yu, Chengpu; Verhaegen, Michel; Hansson, A.

DOI
10.1109/TAC.2017.2738919

Publication date
2017

Document Version
Final published version

Published in
IEEE Transactions on Automatic Control

Subspace Identification of Local Systems in One-Dimensional Homogeneous Networks

Chengpu Yu, Michel Verhaegen, and Anders Hansson

***Abstract*—This note considers the identification of large-scale one-dimensional networks consisting of identical LTI dynamical systems. A subspace identification method is developed that only uses local input-output information and does not rely on knowledge about the local state interaction. The proposed identification method estimates the Markov parameters of a locally lifted system, following the state-space realization of a single subsystem. The Markov-parameter estimation is formulated as a rank minimization problem by exploiting the low-rank property and the two-layer Toeplitz structural property in the data equation, whereas the state-space realization of a single subsystem is formulated as a structured low-rank matrix-factorization problem. The effectiveness of the proposed identification method is demonstrated by simulation examples.**

***Index Terms*—Large-scale 1-D distributed systems, rank minimization problem, two-layer Toeplitz structure.**

I. Introduction

We consider the identification of large-scale one-dimensional (1-D) homogeneous networked systems, which often result by discretizing dynamical systems described via partial differential equations [1]–[3]. Instead of identifying the global networked system in a centralized manner, a distributed identification method is developed, i.e., the local system dynamics is identified by using local data only. The difficulty of this identification problem is that the local system states and the interconnection signals between neighboring subsystems are unmeasurable. In other words, some input sequences are missing for the concerned local network identification problem.

To cope with the challenging local network identification problem, the property that the inverse of the observability Grammian is off-diagonally decaying was used in [2] to approximate the unmeasurable interconnections, influencing the local system dynamics to be identified, via an (unknown) linear combination of locally neighboring inputs and outputs. The selection of these neighboring input and output quantities requires, however an exhaustive search, which is computationally demanding. As a complementary work to [2], a nuclear norm identification solution was provided in [4] to separate the local dynamics and global dynamics by exploiting their distinct rank and order properties; however, it did not consider the identification of the interconnections between the subsystems.

Manuscript received November 20, 2014; revised July 1, 2015, May 23, 2016, February 14, 2017, and May 23, 2017; accepted August 3, 2017. Date of publication August 11, 2017; date of current version March 27, 2018. This work was supported by the European Research Council under European Union's Seventh Framework Programme (FP7/2007-2013)/ERC under Grant 339681. Recommended by Associate Editor C. M. Lagoa. *(Corresponding author: Chengpu Yu.)*

C. Yu is with the School of Automation, Beijing Institue of Technology, Beijing 100081, China, and was with the Delft Center for Systems and Control, Delft University, Delft 2628CD, The Netherlands (e-mail: c.yu-4@tudelft.nl).

M. Verhaegen is with the Delft Center for Systems and Control, Delft University, Delft 2628CD, The Netherlands (e-mail: m.verhaegen@tudelft.nl).

A. Hansson is with the Division of Automatic Control, Department of Electrical Engineering, Linkoping University, Linkoping SE-581 83, Sweden (e-mail: anders.g.hansson@liu.se).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2017.2738919

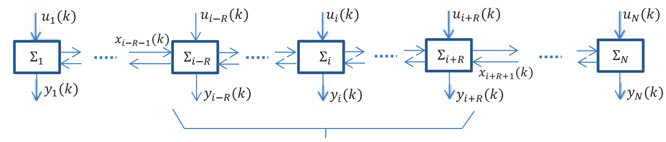

Fig. 1. Illustration of a cluster of subsystems in a neighborhood of the subsystem Σ_i with radius R. The states $x_{i-R-1}(k)$ and $x_{i+R+1}(k)$ are explicitly indicated. They are like all other states unmeasurable.

A new subspace identification method is presented, which is to identify the Markov parameters of a locally lifted system first, following the state-space realization of the single subsystems. To identify the Markov parameters, we fully exploit the *two-layer* structure of the block Toeplitz matrices in the data equation of a lifted state-space system and formulate a rank minimization problem for which the optimal solution can yield (parts of) the true Markov parameters of the locally lifted system. The state-space realization of an individual subsystem from the estimated Markov parameters is inherently a structured state-space realization problem for which the optimal solution can yield the estimates of system matrices up to a similarity transformation.

The rest of this note is organized as follows. Section II describes the concerned identification problem and shows the challenge of dealing with the identification of a small cluster of subsystems in a large-scale network. Section III presents a method for identifying the Markov parameters of locally lifted state-space models. Section IV provides a solution to the state-space realization of a single subsystem. Section V provides numerical simulation results. The conclusions are provided in Section VI.

II. Preliminaries and Problem Definition

We consider the linear time-invariant (LTI) systems $\{\Sigma_i\}_{i=1}^N$ connected in a homogeneous 1-D network as shown in Fig. 1. Denote by $x_i(k) \in \mathbb{R}^n$, $u_i(k) \in \mathbb{R}^m$, $y_i(k) \in \mathbb{R}^p$, and $e_i(k) \in \mathbb{R}^p$ the state, input, output, and measurement noise of the ith subsystem, respectively. By lifting all states $x_i(k)$ into the vector $x(k)$ as $x(k) = [x_1^T(k) \quad \cdots \quad x_N^T(k)]^T$ and doing the same for the inputs, outputs, and noises defining, respectively the vectors $u(k)$, $y(k)$, and $e(k)$, the global networked system has the following state-space model:

$$\begin{aligned} x(k+1) &= \mathcal{A}x(k) + \mathcal{B}u(k) \\ y(k) &= \mathcal{C}x(k) + e(k) \end{aligned} \tag{1}$$

where $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are $N \times N$ block matrices which have the following forms:

$$\mathcal{A} = \begin{bmatrix} A & A_r & & \\ A_l & A & \ddots & \\ & \ddots & \ddots & A_r \\ & & A_l & A \end{bmatrix},$$

$$\mathcal{B} = \begin{bmatrix} B & & & \\ & B & & \\ & & \ddots & \\ & & & B \end{bmatrix}, \mathcal{C} = \begin{bmatrix} C & & & \\ & C & & \\ & & \ddots & \\ & & & C \end{bmatrix}.$$

For a large-scale distributed system, we always assume that $N \gg n$ and $n > \max\{p, m\}$.

To identify the system model in (1), the existing system identification (SID) methods only estimate the triplet $(\mathcal{A}, \mathcal{B}, \mathcal{C})$ *up to a similarity transformation*, thereby not preserving the special block-diagonal and block-tridiagonal structures of these system matrices. In addition, the computational complexity of SID methods, which is at least $O(N^3)$ with N being the number of subsystems in the network, may easily disqualify their use for the identification of large-scale networks.

To deal with the high computational complexity for identifying the global system model in (1), we consider the identification of a cluster of subsystems $\{\Sigma_j\}_{j=i-R}^{i+R}$ in a neighborhood of Σ_i with radius R satisfying $R < i < N - R$ and $R \ll N$, as shown in Fig. 1. The lifted state-space model of this cluster is represented as

$$\begin{aligned} \underline{x}_i(k+1) &= \underline{A}_R \underline{x}_i(k) + \underline{B}_R \underline{u}_i(k) + \underline{D}_R \underline{v}_i(k) \\ \underline{y}_i(k) &= \underline{C}_R \underline{x}_i(k) + \underline{e}_i(k) \end{aligned} \tag{2}$$

with $\underline{x}_i(k), \underline{u}_i(k), \underline{y}_i(k)$ the subparts of $x(k), u(k), y(k)$, respectively, from the block rows $(i - R)$ to $(i + R)$; $\underline{A}_R, \underline{B}_R, \underline{C}_R$ are $(2R + 1) \times (2R + 1)$ block matrices which have forms similar to $\mathcal{A}, \mathcal{B}, \mathcal{C}$ in (1), respectively; the $(2R + 1) \times 2$ block matrix $\underline{D}_R$ and the vector $\underline{v}_i(k)$ are defined as

$$\underline{D}_R = \begin{bmatrix} A_l & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & A_r \end{bmatrix}, \quad \underline{v}_i(k) = \begin{bmatrix} x_{i-R-1}(k) \\ x_{i+R+1}(k) \end{bmatrix}.$$

The *problem of interest* is to identify the system matrices A, A_r, A_l, B, C of a single subsystem using only the local input sequence $\underline{u}_i(k)$ and the local output sequence $\underline{y}_i(k)$. More specifically, the system matrices are to be identified up to a similarity transformation, i.e., the estimates $\hat{A}, \hat{A}_r, \hat{A}_l, \hat{B}, \hat{C}$ of A, A_r, A_l, B, C satisfy $\hat{A} = Q^{-1}AQ,\ \hat{A}_l = Q^{-1}A_lQ,\ \hat{A}_r = Q^{-1}A_rQ,\ \hat{B} = Q^{-1}B,\ \hat{C} = CQ$ with $Q \in \mathbb{R}^{n \times n}$ being a nonsingular transformation matrix.

To address the concerned identification problem, we stipulate the following assumption.

Assumption A.1: The global system $(\mathcal{A}, \mathcal{B}, \mathcal{C})$ and the locally lifted system $(\underline{A}_R, \underline{B}_R, \underline{C}_R)$ are assumed to be minimal.

The persistent excitation of the input signal $u(k)$, which will be used for the identifiability analysis in the sequel, is defined below.

Definition 1: A time sequence $u(k) \in \mathbb{R}^{Nm}$ is persistently exciting of order s if there exists an integer h such that the (block-) Hankel matrix

$$\begin{bmatrix} u(k) & u(k+1) & \cdots & u(k+h-1) \\ u(k+1) & u(k+2) & \cdots & u(k+h) \\ \vdots & \vdots & \iddots & \vdots \\ u(k+s-1) & u(k+s) & \cdots & u(k+s+h-2) \end{bmatrix}$$

has full row rank for any positive integer k.

III. Identifying the Markov Parameters of the Locally Lifted State-Space Model (2)

In this section, the Markov parameters of the state-space model in (2) are to be estimated for which a rank minimization problem will be formulated by exploiting the low-rank property of the unmeasurable-state related terms and the specific block Toeplitz structure of the convolution matrix in the data equation. It will be shown in Theorem 1 that, under some mild conditions, the optimal solution to the proposed rank minimization problem can yield (parts of) the true Markov parameters.

A. Formulation of a Rank Minimization Problem

The data equation of the local state-space model (2) is given as follows:

$$Y_{s,h}^i = \mathcal{O}_s x_h^i + \mathcal{T}_s^{B_R} U_{s,h}^i + \mathcal{T}_s^{D_R} V_{s,h}^i + E_{s,h}^i. \tag{3}$$

In this equation, the (block-) Hankel matrix $Y_{s,h}^i$ is defined as

$$Y_{s,h}^i = \begin{bmatrix} \underline{y}_i(1) & \cdots & \underline{y}_i(h) \\ \vdots & \iddots & \vdots \\ \underline{y}_i(s) & \cdots & \underline{y}_i(h+s-1) \end{bmatrix}$$

with the superscript i being the spatial index of the subsystem Σ_i, the subscripts s, h, respectively, being the number of block rows and the number of block columns. Analogous to $Y_{s,h}^i$, we define the block-Hankel matrices $U_{s,h}^i$, $V_{s,h}^i$, $E_{s,h}^i$ from the sequences $\underline{u}_i(k), \underline{v}_i(k), \underline{e}_i(k)$, respectively. The matrix $\mathcal{T}_s^{B_R}$ is a block Toeplitz matrix defined from the triplet $(\underline{A}_R, \underline{B}_R, \underline{C}_R)$ as

$$\mathcal{T}_s^{B_R} = \begin{bmatrix} 0 & & & \\ \underline{C}_R \underline{B}_R & 0 & & \\ \vdots & \ddots & \ddots & \\ \underline{C}_R \underline{A}_R^{s-2} \underline{B}_R & \cdots & \underline{C}_R \underline{B}_R & 0 \end{bmatrix}$$

and $\mathcal{T}_s^{D_R}$ is defined in a similar way from the triplet $(\underline{A}_R, \underline{D}_R, \underline{C}_R)$. The final matrix definitions in (3) are

$$\mathcal{O}_s = \begin{bmatrix} \underline{C}_R \\ \underline{C}_R \underline{A}_R \\ \vdots \\ \underline{C}_R \underline{A}_R^{s-1} \end{bmatrix}$$

and

$$x_h^i = [\underline{x}_i(k) \cdots \underline{x}_i(k+h-1)].$$

Due to the unmeasurable state sequences in a network, the matrix sum $\mathcal{O}_s x_h^i + \mathcal{T}_s^{D_R} V_{s,h}^i$ in (3) is unknown. However, this matrix sum has a low-rank property that will be exploited as a solution in the new subspace identification method.

Lemma 1: For the data equation (3), when $h > ps$ or $Y^i_{s,h}$ is a fat matrix, the sum $\mathcal{O}_s x^i_h + \mathcal{T}^{D_R}_s V^i_{s,h}$ satisfies the following rank property

$$\text{rank}\left(\mathcal{O}_s x^i_h + \mathcal{T}^{D_R}_s V^i_{s,h}\right) \leq (2R+1)n + \min\{(s-1)sp, 2(s-1)n\} \quad (4)$$

where $(s-1)sp$ and $2(s-1)n$ denote the number of non-zero rows and columns of $\mathcal{T}^{D_R}_s$, respectively.

Proof: From the structures of $\mathcal{O}_s$ and $\mathcal{T}^{D_R}_s$, we can get that $\text{rank}\,(\mathcal{O}_s x^i_h) \leq \text{rank}\,(\mathcal{O}_s) \leq (2R+1)n$ and $\text{rank}\left(\mathcal{T}^{D_R}_s V^i_{s,h}\right) \leq \text{rank}\left(\mathcal{T}^{D_R}_s\right) \leq \min\{(s-1)sp, 2(s-1)n\}$. Thus, the result in the lemma is straightforward. ■

From Lemma 1, we can derive a condition to select the parameter s in the data (3) and the cluster radius R, defined above (2), such that the sum $\mathcal{O}_s x^i_h + \mathcal{T}^{D_R}_s V^i_{s,h}$ is of low rank (or rank deficient). This condition reads

$$(2R+1)sp > (2R+1)n + \min\{(s-1)sp, 2(s-1)n\}. \quad (5)$$

The above condition means that the number of the rows of the matrix $\mathcal{O}_s x^i_h + \mathcal{T}^{D_R}_s V^i_{s,h}$ is larger than an upper bound of its rank. In practice, by fixing a value of s satisfying that $s > \frac{n}{p}$, we can always find a value of R such that the above inequality holds. Therefore, in the sequel, we assume that the matrix sum $\mathcal{O}_s x^i_h + \mathcal{T}^{D_R}_s V^i_{s,h}$ is of low rank (or rank deficient).

Denote the noise-free output as $\hat{y}_i(k) = \underline{y}_i(k) - \underline{e}_i(k)$ and its related block Hankel matrix $\hat{Y}^i_{s,h}$. Based on the rank property discussed above, a low-rank regularized optimization problem is then proposed as follows:

$$\min_{\Theta^{B_R}_s \in \mathcal{T}, \hat{Y}^i_{s,h} \in \mathcal{H}} \sum_{t=1}^{h+s-1} \|\hat{y}_i(t) - \underline{y}_i(t)\|^2 + \lambda \cdot \text{rank}\left[\hat{Y}^i_{s,h} - \Theta^{B_R}_s U^i_{s,h}\right] \quad (6)$$

where $\mathcal{T}$ and $\mathcal{H}$ denote, respectively, the set of block Toeplitz and block Hankel matrices with appropriate block sizes, and the regularization parameter λ allows to make a tradeoff between the two terms in the cost function. It is noted that the matrix set $\mathcal{T}$ has a two-layer block Toeplitz structure: the first layer is the block Toeplitz structure of $\mathcal{T}^{B_R}_s$ with respect to its block entries $\underline{C}_R \underline{A}^j_R \underline{B}_R$; the second layer is the partial block Toeplitz structure inside the block entries $\underline{C}_R \underline{A}^j_R \underline{B}_R$, as highlighted in the following example and lemma.

Example 1: If we take $R = 3$ and assume that each block in $\underline{A}_R, \underline{B}_R, \underline{C}_R$ has size 2×2, then a visual illustration of the structures of the matrices $\{\underline{M}_j = \underline{C}_R \underline{A}^j_R \underline{B}_R\}^3_{j=1}$ is given in Fig. 2.

Based on the matrices $\underline{A}_R, \underline{B}_R, \underline{C}_R$ defined in (2), it can be verified that the matrix product $\underline{M}_j = \underline{C}_R \underline{A}^j_R \underline{B}_R$ has the following properties:
1) $\underline{M}_j$ is a banded block matrix, with block bandwidth j;
2) the submatrices of $\underline{M}_j$, for $j < 2R+1$, at the lth block row and qth block column with $l, q \in \{1, \cdots, 2R+1\}$, are inside the partial block-Toeplitz region for the index-pair (l, q) satisfying

$$i + 1 \leq l + q \leq 4R + 3 - i.$$

For the uniqueness property of the rank minimization problem, use will be made of the following time-varying observability matrix $\mathbf{O}_{j,k}$, which is a submatrix of the extended observability matrix $\mathcal{O}_j$ defined in (3), consisting of the block rows corresponding to the second-layer block Toeplitz part of $\mathcal{T}^{B_R}_s$.

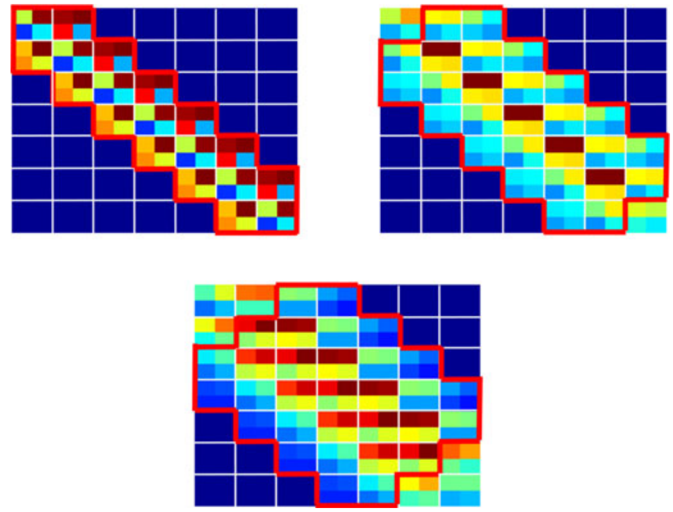

Fig. 2. Illustration of the partial Toeplitz structure of the matrices $\{\underline{M}_j = \underline{C}_R \underline{A}^j_R \underline{B}_R\}^3_{j=1}$: Top left for $j = 1$ with block bandwidth 1; Top right for $j = 2$ with block bandwidth 2; Bottom for $j = 3$ with block bandwidth 3. The deep blue color represents zero entries. The parts surrounded by red curves have block Toeplitz structures.

Definition 2: Let G_j be a $j \times (j+2)$ block Toeplitz matrix of the form

$$G_j = \begin{bmatrix} A_l & A & A_r & & & \\ & A_l & A & A_r & & \\ & & \ddots & \ddots & \ddots & \\ & & & A_l & A & A_r \end{bmatrix}. \quad (7)$$

Denotes $\underline{C}_j = I_j \otimes C$. A time-varying observability matrix $\mathbf{O}_{j,k}$, for $j > 2(k-1)$, is defined in terms of the matrix pair $(\underline{C}_j, G_j)$ as [5, Ch. 3]

$$\mathbf{O}_{j,k} = \begin{bmatrix} \underline{C}_j \\ \underline{C}_{j-2} G_{j-2} \\ \underline{C}_{j-4} G_{j-4} G_{j-2} \\ \vdots \\ \underline{C}_{j-2(k-1)} G_{j-2(k-1)} \cdots G_{j-2} \end{bmatrix}.$$

Theorem 1: Suppose that the following assumptions are satisfied.
1) The global system input $u(k)$ to (1) is persistently exciting of order $Nn + s$.
2) The measurement noise is absent, i.e., $\hat{y}_i(k) = \underline{y}_i(k)$.
3) The matrix pairs (A_l, B) and (A_r, B) have the full row rank.
4) Assumption A.1 holds and ν_o is the observability index of the pair $(\underline{A}_R, \underline{C}_R)$.
5) The cluster *radius* R of the lifted model in (2) and the dimension parameter s in (3) satisfy

$$s > \nu_o, \quad R \geq s - 1.$$

6) The time-varying observability matrix $\mathbf{O}_{2R+1,s-1}$, defined in Definition 2, has the full column rank.

Then, the submatrices of the Markov parameters $\underline{M}_j = \underline{C}_R \underline{A}^j_R \underline{B}_R$, for $j = 0, 1, \ldots, s-2$, contained in block-Toeplitz regions of the matrix $\mathcal{T}^{B_R}_s$ in (3) can be computed in a unique manner from the following low-rank optimization problem:

$$\min_{\Theta^{B_R}_s \in \mathcal{T}_f} \text{rank}\left[Y^i_{s,h} - \Theta^{B_R}_s U^i_{s,h}\right], \quad \text{for } R + s < i < N - R - s. \quad (8)$$

The proof of the above theorem can be found in [6]. The correct recovery of certain parts of the Markov parameters as shown in Theorem 1 implies that the rank minimization problem in (6) is a reasonable formulation.

Since the optimization problem in (6) is nonconvex, it is difficult to obtain an optimal solution under a mild computational burden. In this note, the reweighted nuclear norm optimization method [7] is adopted, which is an iterative heuristic for the rank minimization problem (6).

IV. State-Space Realization of a Single Subsystem

In this section, we shall study the final realization of the system matrices $\{C, A, A_l, A_r, B\}$ from the estimated submatrices of the Markov parameters $\underline{M}_j = \underline{C}_R \underline{A}_R^j \underline{B}_R$, for $j = 0, 1, \ldots, s-2$, contained in second-layer partial Toeplitz regions of the matrix $\mathcal{T}_s^{B_R}$ in (3).

We start the solution by developing expressions of the submatrices in the second-layer Toeplitz regions in terms of the system matrices $\{C, A, A_l, A_r, B\}$. This is done in the following Lemma.

Lemma 2: Consider the block matrices $\underline{A}_R, \underline{B}_R$, and $\underline{C}_R$ defined in (2). Let the sequence of non-zero block entries from left to right of the $(j+1)$th block row of the matrix $\underline{C}_R \underline{A}_R^j \underline{B}_R$ be denoted as $\{F_{j,-j}, F_{j,1-j}, \ldots, F_{j,j-1}, F_{j,j}\}$, then these matrix entries satisfy the following relationship:

$$\sum_{k=-j}^{j} F_{j,k} z^{-k} = C(A_l z^{-1} + A + A_r z)^j B \tag{9}$$

where $z \in \mathbb{C}$.

Proof: The above result can be derived using the filter bank theory in [8].

As $F_{j,k}$ are the Markov parameters inside the second-layer block Toeplitz part of $\mathcal{T}_s^{B_R}$, the values of $F_{j,k}$ for $j \in \{0, 1, \ldots, s-2\}, k \in \{-j, \ldots, j\}$ are assumed to be available in this section. Based on these values $F_{j,k}$, we will address the problem of estimating the matrices $\{C, A, A_l, A_r, B\}$ up to a similarity transformation.

Dual to Definition 2, we shall define a time-varying controllability matrix $\mathbf{C}_{j,k}$, which is a submatrix of the extended controllability matrix determined by $(\underline{A}_R, \underline{B}_R)$.

Definition 3: Let Γ_j be a $(j+2) \times j$ block Toeplitz matrix of the form

$$\Gamma_j = \begin{bmatrix} A_r & & & \\ A & A_r & & \\ A_l & A & \ddots & \\ & A_l & \ddots & A_r \\ & & \ddots & A \\ & & & A_l \end{bmatrix}. \tag{10}$$

Denote $\underline{B}_j = I_j \otimes B$. A time-varying controllability matrix $\mathbf{C}_{j,k}$, for $j > 2(k-1)$, is defined in terms of the matrix pair $(\Gamma_j, \underline{B}_j)$

$$\mathbf{C}_{j,k} = \left[\underline{B}_j \middle| \Gamma_{j-2} \underline{B}_{j-2} \middle| \cdots \middle| \Gamma_{j-2} \cdots \Gamma_{j-2(k-1)} \underline{B}_{j-2(k-1)} \right].$$

In the sequel, let s be an even integer such that $s/2$ is an integer as well. The solution to the realization of the system matrices $\{C, A_l, A, A_r, B\}$ is done in two phases. In the first phase, the structured time-varying observability matrix $\mathbf{O}_{2R+1,s/2}$ and the structured time-varying controllability matrix $\mathbf{C}_{2R+1,s/2}$ are to be estimated from the available matrix values $F_{i,j}$. In the second phase, the system matrices $\{C, A, A_l, A_r, B\}$ are derived from these time-varying observability and controllability matrices. For the notational simplicity, the subscripts are denoted by $M = 2R+1$ and $L = s/2$.

A. Determining the Time-Varying Observability and Controllability Matrices

In this section, the determination of $\mathbf{O}_{M,L}$ and $\mathbf{C}_{M,L}$ will be formulated as a structured low-rank matrix-factorization problem. More importantly, we show that the optimal solution to this matrix-factorization problem can yield the estimates of $\mathbf{O}_{M,L}$ and $\mathbf{C}_{M,L}$ up to a block-diagonal ambiguity matrix with identical block-diagonal entries. This is crucial to identify the system matrices $\{C, A, A_l, A_r, B\}$ up to a similarity transformation.

First, by the definitions of $\mathbf{O}_{M,L}$ and $\mathbf{C}_{M,L}$, we can find that the product of $\mathbf{O}_{M,L}$ and $\mathbf{C}_{M,L}$ is equal to a matrix constructed from $\{F_{j,k}\}_{k=-j}^{j}$ for $j = 0, 1, \ldots, s-2$. Here, the product of $\mathbf{O}_{M,L}$ and $\mathbf{C}_{M,L}$ is represented as

$$\mathbf{O}_{M,L} \mathbf{C}_{M,L} = \mathbf{H} \tag{11}$$

where $\mathbf{H}$ is a $(2R+1) \times (2R+1)$ block matrix constructed by the block entries $F_{j,k}$ that are assumed to be known.

Given the matrix $\mathbf{H}$, the problem of interest is to determine $\mathbf{O}_{M,L}$ and $\mathbf{C}_{M,L}$ from (11). Denote by $\mathcal{O}_{M,L}$, the set of block matrices having the same structure as $\mathbf{O}_{M,L}$, and $\mathcal{C}_{M,L}$ the set of block matrices having the same structure as $\mathbf{C}_{M,L}$. We then propose the following structured low-rank matrix-factorization problem:

$$\begin{aligned} \min_{\mathbf{O},\mathbf{C}} \quad & \|\mathbf{H} - \mathbf{O}\mathbf{C}\|_F^2 \\ \text{s.t.} \quad & \mathbf{O} \in \mathcal{O}_{M,L}, \mathbf{C} \in \mathcal{C}_{M,L}. \end{aligned} \tag{12}$$

According to the structures of $\mathbf{O}_{M,L}$ and $\mathbf{C}_{M,L}$ described in Definitions 2 and 3, we will show in the following theorem that the optimal solution to (12) can yield the estimates of $\mathbf{O}_{M,L}$ and $\mathbf{C}_{M,L}$ up to a block-diagonal ambiguity matrix with identical block-diagonal entries.

Theorem 2: Consider the optimization problem in (12). Suppose that the following assumptions are satisfied.

1) The values of R and s satisfy $R \geq s-2$, and s is a positive even integer.
2) The matrices $\mathbf{O}_{j,s/2}$ and $\mathbf{C}_{j,s/2}$, for any $j \geq \min\{2s-3, 2R\}$, have full column and row rank, respectively.
3) The matrix $\mathbf{H}$ satisfying (11) is known exactly.

Then, any optimal solution pair $\{\hat{\mathbf{O}}, \hat{\mathbf{C}}\}$ to the optimization problem (12) satisfies

$$\begin{aligned} \hat{\mathbf{O}} &= \mathbf{O}_{M,L} \mathbf{Q} \\ \hat{\mathbf{C}} &= \mathbf{Q}^{-1} \mathbf{C}_{M,L} \end{aligned} \tag{13}$$

where $\mathbf{Q} = I_{2R+1} \otimes Q$ with $Q \in \mathbb{R}^{n \times n}$ being a nonsingular ambiguity matrix.

The proof of this theorem can also be found in [6].

The concerned optimization problem in (12) is bilinear and nonconvex. In this note, this problem is solved using the method developed in our previous work [9]: the bilinear optimization problem (12) is recast into a rank-constrained optimization problem, and further a difference-of-convex optimization problem, which is then solved using the sequential convex programming method.

B. Determining the System Matrices $\{A, A_l, A_r, B, C\}$

We assume that the obtained estimates $\hat{\mathbf{O}}_{M,L}$ and $\hat{\mathbf{C}}_{M,L}$ satisfy

$$\begin{aligned} \hat{\mathbf{O}}_{M,L} &= \mathbf{O}_{M,L} \mathbf{Q} \\ \hat{\mathbf{C}}_{M,L} &= \mathbf{Q}^{-1} \mathbf{C}_{M,L} \end{aligned} \tag{14}$$

Algorithm 1: Local identification for 1D distributed systems.

Step 1 Construct a spatially stacked state-space model (2) and its temperally stacked equation (3) based on local observations;
Step 2 Estimate $\mathcal{T}_s^{B_R}$ from the optimization problem (6);
Step 4 Estimate $\mathbf{O}_{M,L}$ and $\mathbf{C}_{M,L}$ from the optimization problem (12);
Step 5 Extract the estimates of C and B from the estimates of $\mathbf{O}_{M,L}$ and $\mathbf{C}_{M,L}$, respectively;
Step 6 Estimate A_l, A, A_r by solving the optimization problem in (16).

where $\mathbf{Q} = I_{2R+1} \otimes Q$ with $Q \in \mathbb{R}^{n\times n}$ being nonsingular. Based on these estimates, we will address the identification of the system matrices $\{A, A_l, A_r, B, C\}$ up to a similarity transformation.

First, the shifting structure of the time-varying observability matrix $\mathbf{O}_{M,L}$ will be explored. Denote

$$\mathbf{O}_{j,k_1:k_2} = \begin{bmatrix} \underline{C}_{j-2k_1} G_{j-2k_1} \cdots G_{j-2} \\ \underline{C}_{j-2(k_1+1)} G_{j-2(k_1+1)} \cdots G_{j-2} \\ \vdots \\ \underline{C}_{j-2k_2} G_{j-2k_2} \cdots G_{j-2} \end{bmatrix}$$

where $0 \le k_1 < k_2 \le s/2 - 2$ and $2k_2 \le j \le 2R+1$. The matrix $\mathbf{O}_{j,k_1:k_2}$ above is constructed by the block rows of $\mathbf{O}_{j,k}$ with block-row indices from k_1 to k_2. Then, the *structure-shifting property* of $\mathbf{O}_{M,L}$ can be represented as

$$\mathbf{O}_{2R-1,0:s/2-2} G_{2R-1} = \mathbf{O}_{2R+1,1:s/2-1} \tag{15}$$

where $\mathbf{O}_{2R-1,0:s/2-2}$ and $\mathbf{O}_{2R+1,1:s/2-1}$ are submatrices of $\mathbf{O}_{2R+1,s/2}$, and G_{2R-1} is a block Toeplitz matrix defined in Definition 2.

Based on (15), we formulate the following structured least-squares optimization problem to identify the matrices A_l, A, A_r based on the estimate $\hat{\mathbf{O}}_{2R+1,s/2}$

$$\begin{aligned} \min_{G} \quad & \|\hat{\mathbf{O}}_{2R-1,0:s/2-2} G - \hat{\mathbf{O}}_{2R+1,1:s/2-1}\|_F^2 \\ \text{s.t.} \quad & G \in \mathcal{G}_{2R-1} \end{aligned} \tag{16}$$

where $\mathcal{G}_{2R-1}$ denotes a set of matrices having the same structure as G_{2R-1}, as shown in Definition 2.

The optimal solution to (16) has properties shown in the following lemma.

Lemma 3: Let $\hat{\mathbf{O}}_{2R+1,s/2}$ satisfy (14). Assume that $\hat{\mathbf{O}}_{2R-1,0:s/2-2}$ has full column rank. Then, the optimal solution $\hat{G}$ to the optimization problem in (16) satisfies

$$\hat{G} = \left(I_{2R-1} \otimes Q^{-1}\right) G_{2R-1} \left(I_{2R+1} \otimes Q\right) \tag{17}$$

where $Q \in \mathbb{R}^{n\times n}$ is a nonsingular transformation matrix.

The above lemma can be derived straightforwardly based on (14) and the optimization formulation (16). Lemma 3 implies that the matrices A_l, A, A_r can be determined up to a similarity transformation, i.e.,

$$\hat{A}_l = Q^{-1} A_l Q, \quad \hat{A} = Q^{-1} A Q, \quad \hat{A}_r = Q^{-1} A_r Q.$$

In addition, according to (14), the estimates $\hat{C}$ and $\hat{B}$ can be extracted, respectively, from $\hat{\mathbf{O}}_{M,L}$ and $\hat{\mathbf{C}}_{M,L}$, satisfying that

$$\hat{C} = CQ, \quad \hat{B} = Q^{-1} B.$$

To ease the reference, the proposed local-network identification algorithm is summarized in Algorithm 1.

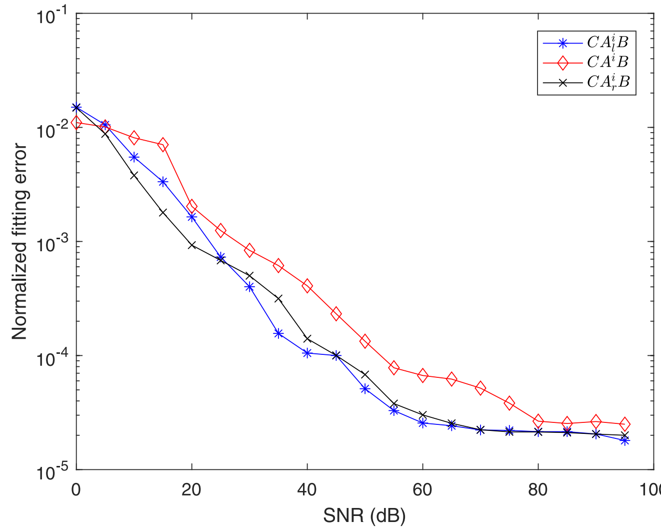

Fig. 3. Normalized fitting errors under different noise levels.

V. Numerical Simulation

In this section, numerical simulations are provided to demonstrate the effectiveness of the proposed identification method—Algorithm 1. In the simulation, the distributed system is constructed by connecting 40 identical subsystems in a line, and the identification for the 20th subsystem is performed. The system matrices (A, A_l, A_r, B, C) with $A, A_l, A_r \in \mathbb{R}^{3\times 3}$, $B \in \mathbb{R}^{3\times 2}$, and $C \in \mathbb{R}^{2\times 3}$ are randomly generated such that Assumption A.1 is satisfied and the 1-D networked system is stable.

The system input and the measurement noise are generated as white noise sequences, and the data length is set to 1200. The regularization parameter λ in (6) is chosen using the cross-validation method [10, Ch. 5], i.e., λ is determined by computing solutions using half of the available data for 20 logarithmically-spaced values of λ in the interval 10^{-4} to 10^2, and selecting the value of λ yielding the best fit on the other half of the available data.

To evaluate the identification performance against the noise effect, the criterion signal-to-noise ratio (SNR) is adopted, which is defined as

$$\text{SNR} = 10 \log \left(\frac{\text{var}(y_i(k) - e_i(k))}{\text{var}(e_i(k))} \right).$$

In the sequel, we shall carry out numerical simulations with the SNR ranging from 0 to 95 dB.

We use the criterion *impulse-response fitting* to evaluate the performance of the proposed identification method. The normalized fitting error of the impulse-response sequence CA^iB is defined by

$$\frac{1}{T} \sum_{j=1}^{T} \frac{\sum_{i=0}^{10} \|CA^iB - \hat{C}_j \hat{A}_j^i \hat{B}_j\|_F}{\sum_{i=0}^{10} \|CA^iB\|_F}$$

where T is the number of randomly generated networked systems and $\{\hat{C}_j, \hat{A}_j, \hat{B}_j\}$ are the estimates of $\{C, A, B\}$ of the jth generated network model. Similarly, we define the normalized fitting errors for the impulse-response sequences CA_l^iB and CA_r^iB.

Fig. 3 shows the impulse-response fittings of individual subsystems against the SNR, where the dimension parameters involved in the proposed identification method are set to $R = 5$ and $s = 8$. The normalized fitting errors are calculated by averaging the results of 50 randomly generated networked systems. It can be observed from Fig. 3

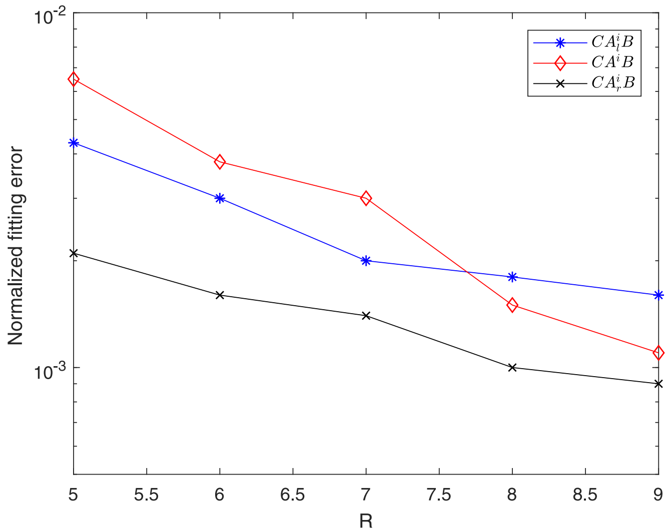

Fig. 4. Normalized fitting errors against the dimension parameter R.

that, using the proposed identification method, the normalized fitting error decreases along with the increase of the SNR. When the SNR is larger than 50 dB, the normalized fitting error can be smaller than 10^{-4}, indicating that the state-space model of a single subsystem can be accurately identified at a low noise level.

Fig. 4 shows the impulse-response fidelity of individual subsystems against the dimension parameter R, where $s = 8$ and SNR$=15$ dB. It can be seen that the better identification performance can be achieved by choosing a larger value of R. This can be explained by that, as the radius of the local cluster R becomes larger, more input and output data can be used for the system identification.

VI. Conclusion

The local identification of 1-D large-scale distributed systems has been studied. Compared with the classical system identification problems, the challenging point of the local system identification is that there are two unknown system inputs, which are the states of its neighboring subsystems. By exploiting both the spatial and temporal structures of the distributed system, especially the two-layer Toeplitz structure of the Markov-parameter matrix, a rank minimization problem has been provided for identifying the Markov parameters of a local cluster of identical subsystems, where the associated optimal solution can yield (parts of) the true Markov parameters. Moreover, the structured state-space realization of the local cluster is formulated as a structured low-rank matrix-factorization problem, showing that the system matrices can be determined up to a similarity transformation by enforcing the structure of the generalized observability/controllability matrix.

Although we only consider the identification of 1-D homogeneous networked systems, it can be extended to 2-D homogeneous networks by using the same identification framework, namely exploiting both the spatial and temporal structures of the concerned distributed system. In our future work, the identification of large heterogeneous networks will be investigated.

References

[1] J. Rice and M. Verhaegen, "Distributed control in multiple dimensions: A structure preserving computational technique," *IEEE Trans. Autom. Control*, vol. 56, no. 3, pp. 516–530, Mar. 2011.

[2] A. Haber and M. Verhaegen, "Subspace identification of large-scale interconnected systems," *IEEE Trans. Autom. Control*, vol. 59, no. 10, pp. 2754–2759, Oct. 2014.

[3] J. Kulkarni, R. DAndrea, and B. Brandl, "Application of distributed control techniques to the adaptive secondary mirror of cornells large atacama telescope," in *Proc. SPIE Astronomical Telescopes Instrum. Conf.*, 2002, pp. 750–756.

[4] N. Matni and A. Rantzer, "Low-rank and low-order decompositions for local system identification," 2014, arXiv: 1403.7175.

[5] P. Dewilde and A.-J. Van der Veen, *Time-Varying Systems and Computations*. New York, NY, USA: Springer Science & Business Media, 2013.

[6] C. Yu and M. Verhaegen, "Subspace identification of large-scale 1d homogeneous networks," 2017, arXiv: 1702.03539.

[7] K. Mohan and M. Fazel, "Reweighted nuclear norm minimization with application to system identification," in *Proc. IEEE Amer. Control Conf.*, 2010, pp. 2953–2959.

[8] G. Strang, "Fast transforms: Banded matrices with banded inverses," *Proc. Nat. Academy Sci. United States Amer.*, vol. 107, no. 28, pp. 12413–12416, 2009.

[9] C. Yu, M. Verhaegen, S. Kovalsky, and R. Basri, "Identification of structured lti mimo state-space models," in *Proc. 2015 IEEE 54th IEEE Conf. Decision Control*, 2015, pp. 2737–2742.

[10] P. Hansen, *Discrete Inverse Problems: Insight and Algorithms*. Philadelphia, PA, USA: SIAM, 2010.