2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

Distributed Proper Orthogonal Decomposition for Large-Scale Networked Dynamical Systems

Chiaki Kojima[1], Issei Kawasaki[1], Satoshi Moriyama[1] and Jun Wada[1]

Abstract— **Recently, dynamical systems in engineering and science problems become drastically larger and too complex. One of the ways to solve the difficulty is to model systems with hierarchical network structures. Proper orthogonal decomposition (POD) is a model reduction method using available data and its singular value decomposition. In this paper, we apply the POD to a networked linear dynamical system and propose the distributed POD which can specify the degree of approximate of each subsystem and preserve the network structure. We also characterize an upper bound of the approximation error and give a criterion to determine the optimal degrees of reduced-order subsystems. As an application to the distributed POD, the ℓ_1-norm minimizing POD is also proposed for a construction of a simple network structure which approximates the behavior of the entire system appropriately. A numerical example is provided to demonstrate the distributed and ℓ_1-norm minimizing PODs and to show their efficiency.**

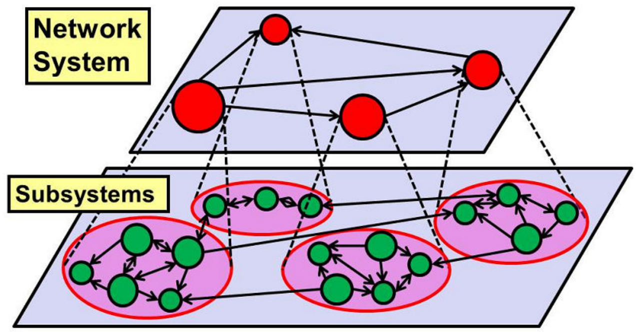

Fig. 1. A dynamical system with a hierarchical network structure

I. Introduction

Recently, due to a performance improvement of computers and the variety of social requirements, dynamical systems in engineering, e.g. smart grids, sensor network, formation control, Internet and economics, and science, e.g. meteorological system, biological system, fluid dynamics and chemical reaction, become increasingly larger and too complex [11]. It is not necessarily desirable to deal with each component of a system in a uniformed way, since there exists a limitation in computational amount and available information about the systems. Hence, if the system of our interest is highly large-scale and complex, we should create an innovation which reduces the amount within the range which does not lose an accuracy.

One of the ways to solve these difficulties pointed out in the last paragraph is to divide a large-scale system into some hierarchical layers by corresponding to its physical scale, and to give a comprehensive notion of a dynamical system modeling [5]. For this purpose, hierarchical and network structures can be an efficient solution to the difficulties [4][6][12][13][14]. The structures may enable us to transform a large-scale system into a networked interconnection of subsystems and to deal with them easier. The figure 1 illustrates the notion of a dynamical system with a hierarchical network structure, where circles and arrows represent the states and their interactions, respectively.

Proper orthogonal decomposition (POD) can be an efficient model reduction method for large-scale systems using available data [10][9][1]. The method has been widely developed in analyzes of fluid dynamics so far. This method can be considered as an application of singular value decomposition (SVD) to a model reduction combined with a mathematical model of a dynamical system. The method has two strong points explained in the following. The first point is that the method is applicable to dynamical systems with nonlinearities and parameters contrary to the ordinary model reduction methods. As the second point, the method can adapt to parameter variations which ordinary reduction methods may be difficult to catch up by using real-time snapshots of the state variables. Although we recognize such merits, we could not find an expansion of the method to networked dynamical system to the best of authors' knowledge so far.

Based on the above observations, we focus on network structures and the POD for an efficient modeling of large-scale systems and propose network POD as a new theoretical framework of the model reduction. In addition, a difficulty of large-scale networked dynamical system can be viewed as a combination of high-dimensionality of subsystems and complexity of a network structure. Hence, a model reduction for such systems should be completed by simplifying the network structure with a reduction of subsystems. We should remark that similar directions has been also proposed by Sandberg and Murray [12], Ishizaki et.al. [7] and etc so far. However, a data based reduction method for such systems has not been proposed to the best of the authors' knowledge. In addition, a network simplification of a dynamical networked system has not also been considered so far.

In this paper, we apply the POD to a networked dynamical system and propose the distributed POD which can specify the degree of approximation of each subsystem and preserve the network structure. The ℓ_1-norm minimizing POD is also

[1]C. Kojima, I. Kawasaki, S. Moriyama and J. Wada are with Department of Information Physics and Computing, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan chiaki_kojima@ipc.i.u-tokyo.ac.jp

978-3-952-41734-8/©2013 EUCA

proposed for a construction of a simple network structure, which approximates the behavior of the entire system in an appropriate sense.

We use the following notations throughout this paper. The set of $m \times n$ real and $n \times n$ symmetric matrices are denoted by $\mathbb{R}^{m\times n}$ and $\mathbb{R}_{\mathrm{s}}^{n\times n}$, respectively. We denote the matrix $\begin{bmatrix} A_1^\top & A_2^\top & \cdots & A_n^\top \end{bmatrix}^\top$ by $\mathrm{col}\,(A_1, A_2, \cdots, A_n)$. We define $\mathrm{diag}(A_1, A_2, \cdots, A_n)$ as the $q \times q$ (block) diagonal matrix with (block) diagonal elements $\{A_1, A_2, \cdots, A_n\}$.

II. Proper Orthogonal Decomposition

In this section, we introduce preliminaries for the model reduction based on the proper orthogonal decomposition of a continuous-time linear dynamical system [10][9][1].

Consider a continuous-time linear dynamical system Σ described by the state-space equation

$$\frac{d}{dt}x(t) = Ax(t) + Bu(t), \tag{1}$$

$$y(t) = Cx(t) + Du(t), \tag{2}$$

where $x \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^n)$, $u \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^m)$, $y \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^p)$ are the state, input and output variables and $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$ are constant matrices. Using the projection $\Pi = WW^\top$, $W^\top W = I_k$, $W \in \mathbb{R}^{n\times k}$, $k \ll n$, we would like to construct the reduced-order system $\hat{\Sigma}$ by

$$\frac{d}{dt}\hat{x}(t) = W^\top AW\hat{x}(t) + W^\top Bu(t), \tag{3}$$

$$y(t) = CW\hat{x}(t) + Du(t), \tag{4}$$

which minimizes the approximation error

$$\int_0^T \left\| x(t) - WW^\top x(t) \right\|_2^2 dt \text{ s.t. } W^\top W = I_k. \tag{5}$$

where $\hat{y} \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^p)$ and $\hat{x}(t) = W^\top x(t) \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^k)$ evolves in k-dimensional space. The columns of W form as orthonormal set, Π is orthogonal and is called a *Galerkin projection*. Then, $x(t)$ is approximated as $x(t) \approx WW^\top x(t)$.

A Galerkin projection is constructed as follows [9] if the continuous snapshots $\{x(t); t \in [0, T]\}$. The Galerkin projection is constructed as the first k dominant eigenvectors of the symmetric matrix $R \in \mathbb{R}_{\mathrm{s}}^{n\times n}$ defined by

$$R := \int_0^T x(t)x(t)^\top dt.$$

However, it contains an essential difficulty that the snapshots in continuous-time are hard to obtain.

Due to the above difficulty, we would like to minimize the following error in discrete-time defined by

$$\sum_{k=0}^{N} \left\| x(t_k) - WW^\top x(t_k) \right\|_2^2, \; t_k := kh, h := \frac{1}{N+1}T$$
$$\text{s.t. } W^\top W = I_k \tag{6}$$

instead of the error in (5). Then the Galerkin projection W can be constructed as follows. Define the data matrix $X \in \mathbb{R}^{n\times(N+1)}$ of the state variable by

$$X := \begin{bmatrix} x(t_0) & x(t_1) & \cdots & x(t_N) \end{bmatrix},$$

which is constructed by the discrete snapshots $x(t_k)$ $(k = 0, 1, \cdots, N)$ of the state variable. Consider the SVD of X given by

$$X = U_1\Sigma_1 V_1^\top + U_2\Sigma_2 V_2^\top, \tag{7}$$

where $U_1 \in \mathbb{R}^{n\times k}$, $U_2 \in \mathbb{R}^{n\times(N-k)}$, $V_1^\top \in \mathbb{R}^{k\times L}$, $V_2^\top \in \mathbb{R}^{(N-k)\times N}$ are orthogonal matrices and $\Sigma_1 \in \mathbb{R}^{k\times k}$, $\Sigma_2 \in \mathbb{R}^{(N-k)\times(N-k)}$ are diagonal matrices defined by

$$\Sigma_1 := \mathrm{diag}\,(\sigma_1, \sigma_2, \cdots, \sigma_k)\,,$$
$$\Sigma_2 := \mathrm{diag}\,(\sigma_{k+1}, \sigma_{k+2}, \cdots, \sigma_N)\,,$$
$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k \gg \sigma_{k+1} \geq \cdots \geq \sigma_N,$$

respectively. Suppose that the singular values of X decays rapidly and only k of them are significant [1]. By truncating the series of the values at k, X is approximated by

$$X \simeq U_1\Sigma_1 V_1^\top. \tag{8}$$

The Galerkin projection W consists of the leading k left singular vectors of the data matrix X, that is $W = U_1 \in \mathbb{R}^{n\times k}$. Thus, we obtain the state-space equation of the reduced order system $\hat{\Sigma}$ by (3) and (4).

The connection between the POD based on continuous and discrete snapshots can be explained as follows [9]. When the sampled snapshots are sufficiently dense in $[0, T]$, i.e. the sampling interval h as sufficiently small. The projection W computed from the SVD (7) asymptotically minimizes the error (5) in continuous-time.

III. Problem Formulation

In this section, we formulate a networked dynamical system which is dealt in this paper. Next, we give the problem formulation of the distributed POD problem.

A. Networked Dynamical System

Networked dynamical system consists of L-tuple of subsystems in lower layer and network system in higher layer. Each subsystem represents local dynamics of the hierarchical network system. On the other hand, the network system plays a role which communicates the internal information of each subsystem. We give mathematical formulation of these system in this subsection.

We give indices $i = 1, \cdots, L$ for each subsystem and define ith subsystem Σ_i. Suppose that Σ_i is represented by the state-space equation

$$\frac{d}{dt}x_i(t) = A_i x_i(t) + B_i r_i(t) + u_i(t), \tag{9}$$

$$y_i(t) = C_i x_i(t) + D_i r_i(t) \tag{10}$$

$A_i \in \mathbb{R}^{n_i\times n_i}$, $B_i \in \mathbb{R}^{n_i\times m_i}$, $C_i \in \mathbb{R}^{p_i\times n_i}$, $D_i \in \mathbb{R}^{p_i\times m_i}$ are constant matrices, and $r_i \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^{r_i})$, $u_i \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^{n_i})$, $x_i \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^{n_i})$, $y_i \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^{p_i})$ are the input variable, the input variable from the network, the state variable, and the output variable of Σ_i, respectively.

[1]We can also determine the value of k to what extent we would like to approximate Σ.

Subsystems are interconnected with a network structure in the upper layer via a contraction of state information. We call the structure as *network system*. The network system Σ_0 is represented by a static equation

$$\begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_L(t) \end{bmatrix} = \begin{bmatrix} M_{1,1} & M_{1,2} & \cdots & M_{1,L} \\ M_{2,1} & M_{2,2} & \cdots & M_{2,L} \\ \vdots & \vdots & \ddots & \vdots \\ M_{L,1} & M_{L,2} & \cdots & M_{L,L} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_L(t) \end{bmatrix}, \quad (11)$$

where $M_{i,j} \in \mathbb{R}^{n_i \times n_j}$ $(i, j = 1, 2, \cdots, L)$. In (11), the matrix $M \in \mathbb{R}^{\sum_{i=1}^{L} n_i \times \sum_{j=1}^{L} n_j}$ is called the *network matrix*.

The entire system Σ can be regarded as a feedback interconnection of subsystems and network system. Then, the ith subsystem is represented by

$$\frac{d}{dt}x_i(t) = A_i x_i(t) + \sum_{i=1}^{L} M_{i,j} x_j(t) + B_i r_i(t), \quad (12)$$

$$y_i(t) = C_i x_i(t) + D_i r_i(t), \quad (13)$$

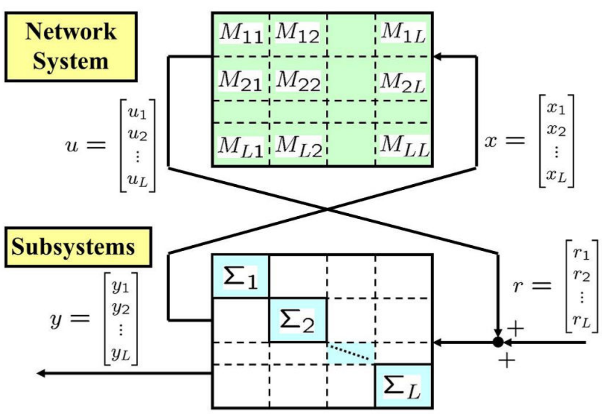

Fig. 2. entire system as a feedback interconnection of subsystems and network system

B. Problem Formulation

We first assume a ideal situation that the snapshots

$$\{x_i(t); t \in [0, T]\} \quad (14)$$

are available in continuous-time.

We formulate the distributed POD problem in the following.

Problem 1: Consider the subsystems Σ_i defined by state-space equation (12) and (13) for $i = 1, \cdots, L$. Assume that continuous-time snapshots in (15) are available. Let the dimension $k \in \mathbb{Z}$, $k > 0$ of the entire system Σ be given. Then, we find a reduced-order system $\hat{\Sigma}$ satisfying the following requirements (i) and (ii).

(i) The Galerkin projection $W_i \in \mathbb{R}^{n_i \times k_i}$ to $\hat{\Sigma}_i$ minimizes the approximation error

$$\sum_{i=1}^{L} \int_0^T \left\| x_i(t) - W_i W_i^\top x_i(t) \right\|_2^2 dt \quad (15)$$

such that

$$W_i W_i^\top = I_{k_i} \text{ and } \sum_{i=1}^{L} k_i = k. \quad (16)$$

(ii) The reduced network system $\hat{\Sigma}_0$ preserves the original structure of Σ_0, i.e.

$$M_{i,j} = 0_{n_i \times n_j} \implies \hat{M}_{i,j} = 0_{k_i \times k_j}$$

for all $i, j = 1, \cdots, L$, where $\hat{M}_{i,j} \in \mathbb{R}^{k_i \times k_j}$ is the (i, j)th block of the reduced-order network matrix corresponding to $M_{i,j}$.

Remark 1: In actual situation, the continuous snapshots are unavailable as we have pointed out in Section II. Hence, the approximation error in (15) is going to be replaced by the following error on discrete snapshots in Section IV-B.

$$\sum_{i=1}^{L} \sum_{k=0}^{N} \left\| x_i(t_k) - W_i W_i^\top x_i(t_k) \right\|_2^2,$$

$$t_k := kh, \ h := \frac{1}{N+1}T \quad (17)$$

Remark 2: The proposed reduction method is still applicable to the cases where subsystems $\Sigma_1, \cdots, \Sigma_L$ are nonlinear and the network system Σ_0 is dynamical and/or nonlinear. However, the main purpose of this paper is to clarify the approximation errors of the distributed and ℓ_1-norm minimizing PODs, and to give a policy for the dimension selection of reduced-order subsystems and network simplification. Hence, we restrict our attention to linear subsystems and static network system. It remains a future work to generalize the characterization of approximation error and dimension selection to nonlinear subsystems and dynamic network system.

IV. Distributed Proper Orthogonal Decomposition

In this section, we propose the distributed POD as a solution to the problem shown in the last section. We first consider an ideal situation that continuous snapshots are available for the POD and give a theoretical bound for the approximation error. This leads a criterion for the selection of optimal dimension of reduced-order subsystems. Based on the error bound, we derive an algorithm for the distributed POD using the discrete snapshots.

A. Distributed POD

Using the snapshots in (14), we define the symmetric matrix $R_i \in \mathbb{R}_s^{n_i \times n_i}$ by

$$R_i = \int_0^T x_i(t) x_i(t)^\top dt. \quad (18)$$

The SVD of R_i in (18) is given by

$$R_i = U_{i,1} \Sigma_{i,1} U_{i,1}^\top + U_{i,2} \Sigma_{i,2} U_{i,2}^\top, \quad (19)$$

where $U_{i,1} \in \mathbb{R}^{n_i \times k_i}$, $U_{i,2} \in \mathbb{R}^{n_i \times (n_i - k_i)}$ are orthogonal matrices and $\Sigma_{i,1} \in \mathbb{R}^{k_i \times k_i}$ is the diagonal matrix whose diagonal elements are the large singular values of

R_i, $\Sigma_{i,2} \in \mathbb{R}^{(n_i-k_i)\times(n_i-k_i)}$ has sufficiently small singular values. Thus, the Galerkin projection W_i is computed as

$$W_i = U_{i,1} \tag{20}$$

based on the above truncation in an appropriate dimension.

By using W_i in (20), the state variable $\hat{x}_i \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^{k_i})$ of the reduced-order subsystem $\hat{\Sigma}_i$ is given by $\hat{x}_i(t) = W_i^\top x_i(t)$. Then, $\hat{\Sigma}_i$ is described by the state-space equation

$$\begin{aligned}\dot{\hat{x}}_i(t) =& W_i^\top A_i W_i \hat{x}_i(t) + W_i \sum_{j\neq i} M_{i,j} W_j \hat{x}_j(t) \\ &+ W_i^\top B_i r_i(t) \quad (21)\\ y_i(t) =& W_i \hat{x}_i(t) + D_i r_i(t) \quad (22)\end{aligned}$$

Then, we have the network matrix of reduced-order network system $\hat{\Sigma_0}$

$$\hat{M} := \begin{bmatrix} \hat{M}_{1,1} & \hat{M}_{1,2} & \cdots & \hat{M}_{1,L} \\ \hat{M}_{2,1} & \hat{M}_{2,2} & \cdots & \hat{M}_{2,L} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{M}_{L,1} & \hat{M}_{L,2} & \cdots & \hat{M}_{L,L} \end{bmatrix}, \tag{23}$$

where $\hat{M}_{i,j} \in \mathbb{R}^{k_i \times k_j}$ $(i, j = 1, 2, \cdots, L)$ is defined by

$$\hat{M}_{ij} = W_i^\top M_{i,j} W_j.$$

We have the following error bounds for the distributed POD by an analogous discussion to the derivation of Theorem 3.1 in Chaturantabut and Sorensen [2]. The proof is omitted due to a space limitation.

Theorem 1: Let $\lambda_{i,j} \in \mathbb{R}$ $(i = 1, \cdots, L;\ j = 1, \cdots, n_i)$ be singular values of R_i in (18) such that

$$\lambda_{i,1} \geq \cdots \geq \lambda_{i,k_i} \geq \lambda_{i,k_i+1} \geq \cdots \geq \lambda_{i,n_i}. \tag{24}$$

Then, the error bound in (15) for the distributed POD is characterized by the inequality

$$\begin{aligned}&\sum_{i=1}^{L} \int_0^T \|x_i(t) - W_i \hat{x}_i(t)\|_2^2 \, dt \\ &\quad \leq \left(1 + T c_\mu(T) \alpha^2\right) \sum_{i=1}^{L} \sum_{j=k_i+1}^{n_i} \lambda_{i,j}, \quad (25)\end{aligned}$$

where $c_\mu(T)$ is defined by

$$c_\mu(T) := \begin{cases} \frac{1}{a}\left(e^{2at} - 1\right) & (a \neq 0) \\ 2t & (a = 0) \end{cases},$$

$$a := \mu(M) + \sum_{i=1}^{L} \left\|W_i^\top\right\|_2 \lambda_{\max}(A_i),$$

$$\mu(M) := \max\left\{\mu \in \mathbb{R} \;\middle|\; \mu\text{: eigenvalue of } \tfrac{1}{2}\left(M + M^\top\right)\right\}.$$

We give a criterion for choosing a dimension k_i of reduced-order subsystem $\hat{\Sigma}_i$ in the remainder of this subsection.

B. Algorithm for Distributed POD

As we have pointed out in Section III-B, the continuous snapshots are unavailable in the actual setting. For this reason, we replace the approximation error in (15) with the error in (17).

Let the discrete snapshots

$$\{x_i(t_0), x_i(t_1), \cdots, x_i(t_{N-1})\}$$

be given. Construct the data matrix $X_i \in \mathbb{R}^{n_i \times (N+1)}$, $i = 1, 2, \cdots, L$ of discrete snapshots of Σ_i by

$$X_i := \begin{bmatrix} x_i(t_0) & x_i(t_1) & \cdots & x_i(t_N) \end{bmatrix}, \tag{26}$$

where $t_k := kh$ and $h := \frac{1}{N+1}T$. The SVD of X_i in (26) is given by

$$X_i = U_{i,1}\Sigma_{i,1}V_{i,1}^\top + U_{i,2}\Sigma_{i,2}V_{i,2}^\top, \tag{27}$$

where $U_{i,1} \in \mathbb{R}^{n_i \times k_i}$, $U_{i,2} \in \mathbb{R}^{n_i \times (N-k_i)}$, $V_{i,1}^\top \in \mathbb{R}^{k_i \times n_i}$, $V_{i2}^\top \in \mathbb{R}^{(N-k_i)\times N}$ are orthogonal matrices and $\Sigma_{i,1} \in \mathbb{R}^{k_i \times k_i}$ is the diagonal matrix whose diagonal elements are the large singular values of R_i, $\Sigma_{i,2} \in \mathbb{R}^{(n_i-k_i)\times(n_i-k_i)}$ has sufficiently small singular values. Then, the singular value $\lambda_{i,j}$ and the Galerkin projection W_i are approximated as $\lambda_{i,j} \approx \sigma_{i,j}^2$ and $W_i \approx U_{i,1}$, where $\sigma_{i,j}$ is the singular value of X_i with $\sigma_{i,1} \geq \cdots \geq \sigma_{i,k_i} \geq \sigma_{i,k_i+1} \geq \cdots \geq \sigma_{i,n_i}$.

We summarize the algorithm for the distributed POD based on the SVD on discrete snapshots and Theorem 1.

Algorithm 1:

- Step 1: Compute the SVD (27) for the data matrix X_i $(i = 1, 2, \cdots, L)$. Compare the singular values of among the X_i's. Then, specify the dimension k_i satisfying (16) based on the error bound (25).
- Step 2: Compute the Galerkin projection $W_i \in \mathbb{R}^{k_i \times N}$ by (27) for $i = 1, 2, \cdots, L$.
- Step 3: Construct the reduced-order subsystem $\hat{\Sigma}_i$ by (10) and (11) for $i = 1, 2, \cdots, L$.
- Step 4: Compute the network matrix $\hat{M} \in \mathbb{R}^{k \times k}$ of the reduced-order network system Σ_0 by (23).

Remark 3: As we have explained Section II. if the sampling interval h is sufficiently small and the sampled snapshots are sufficiently dense. the projection W_i minimizes the continuous-time approximation error in (15).

V. ℓ_1-Norm Minimizing POD

In Section IV, we have considered a model reduction of each subsystem with a reduction of the network system simultaneously. However, it can be considered that, corresponding to our interests and requirements, the reduced network system should be minimal which can explain the entire behavior of the original system as simple as possible from the viewpoint of both tractability and computational amount. Such view point has also been proposed in the identification of a genetic network [8]. We propose the ℓ_1-norm minimizing POD in this section for this purpose as a possible evolution of the distributed POD. This is formulated as a slight modification of Problem 1 in the following.

Problem 2: Consider the subsystems Σ_i defined by state-space equation (12) and (13) for $i = 1, \cdots, L$. Assume that continuous-time snapshots in (15) are available. Let the dimension $k \in \mathbb{Z}$, $k > 0$ of the entire system Σ and $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$ for the truncation be given. Then, we find a reduced-order system $\hat{\Sigma}$ satisfying the requirements (i) and (ii) in Problem 1 and the following requirements (iii).

(iii) The network matrix $\hat{M} \in \mathbb{R}^{k\times k}$ of the reduced-order network system $\hat{\Sigma}_0$ has the elements satisfying

$$\hat{m}_{i,j} < \varepsilon, \ \forall\, i, j = 1, \cdots, k,$$

where $\hat{m}_{i,j}$ denotes the (i, j)th element of $\hat{M}$.

A. ℓ_1-norm Relaxation

In this paper, a simplification of a network structure should be defined as getting more sparse network matrix as possible. Such a sparsity can be expressed as minimizing the ℓ_0-norm [3] of the network matrix [2] However, ℓ_0-norm optimization is known as NP-hard problem. Hence, we relax the problem to the optimization of ℓ_1-norm. For this purpose, we introduce the notion of ℓ_1-norm as a preliminary notion.

Consider the constant matrix $M \in \mathbb{R}^{L\times L}$ given by

$$M = \begin{bmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,L} \\ m_{2,1} & m_{2,2} & \cdots & m_{2,L} \\ \vdots & \vdots & \ddots & \vdots \\ m_{L,1} & m_{L,2} & \cdots & m_{L,L} \end{bmatrix} \in \mathbb{C}^{L\times L},$$
$$m_{i,j} \in \mathbb{R}\ (i, j = 1, 2, \cdots, L).$$

Then, the $\ell_1$norm of M is defined by the summation of the magnitude of all elements as

$$\|M\|_1 := \sum_{i=1}^{L}\sum_{j=1}^{L} |m_{i,j}|\,.$$

If we can find the network matrix which minimizes the ℓ_1-norm via the POD, we can get accordingly sparse network matrix which is not possibly the most sparse. As we can easily see, there exists a freedom of orthogonal matrices in the state variable, equivalently the network matrix, in the reduced order system. Hence, thanks to this freedom, we can reconstruct the reduced network matrix which minimizes the ℓ_1-norm. We show the procedure for the construction.

B. ℓ_1-Norm Minimizing POD

Let the network matrix $\hat{M} \in \mathbb{R}^{k\times k}$ of the reduced-order network system $\hat{\Sigma}_0$ in (23) be given. Then, we can construct the simplified network matrix $\hat{M}_{\rm simple} \in \mathbb{R}^{k\times k}$ by the following strategies.

We first consider the ℓ_1-norm optimization

$$T_{\rm opt} = \operatorname{argmin}\|T^\top \hat{M} T\|_1 \quad (28)$$

such that

$$T := \operatorname{diag}(T_{1,1}, T_{2,1}, \cdots, T_{L,1}),$$
$$T^\top T = I,\ T_{i,1} \in \mathbb{R}^{k_i\times k_i}\ (i = 1, 2, \cdots, L),$$

where $T \in \mathbb{R}^{k\times k}$ is a block diagonal similarity transformation matrix. This problem can be solved efficiently as a convex optimization problem. The optimal network matrix $\hat{M}_{\rm opt}$ is defined as follows by using the solution to the above optimization.

$$\hat{M}_{\rm opt} = T_{\rm opt}^\top \hat{M} T_{\rm opt}. \quad (29)$$

We have shown a construction of the network matrix minimizing the ℓ_1-norm and corresponding subsystems. Then, it may not affect the entire behavior so much if we truncate the sufficient small element for a given criteria. We explain this idea mathematically.

We set $\varepsilon > 0$ as a criterion for the truncation. If the element is smaller than or equal to ε after a normalization, we can think of that the influence between subsystems is small. Let $\hat{m}_{s,i,j} \in \mathbb{R}$ $(i, j = 1, 2, \cdots, k)$ be the (i, j)th element of $\hat{M}_{\rm simple}$. This procedure is done by

$$\frac{|\hat{m}_{i,j}|}{\|\hat{M}_{\rm opt}\|_1} \le \varepsilon \implies \hat{m}_{s,i,j} = 0. \quad (30)$$

On the other hand, if the inequality

$$\frac{|\hat{m}_{i,j}|}{\|\hat{M}_{\rm opt}\|_1} > \varepsilon \implies \hat{m}_{s,i,j} = \hat{m}_{i,j}. \quad (31)$$

holds, we remain the element. We define $\hat{M}_{\rm truncated} \in \mathbb{R}^{k\times k}$ as the constant matrix which consists of the truncated elements of $\hat{M}_{\rm opt}$ due to (30). Thus, we can construct the simplified network matrix $\hat{M}_{\rm simple}$ of the reduced-order network system $\hat{\Sigma}_0$. Then, the reduced-order subsystem $\hat{\Sigma}_i$ is expressed as

$$\begin{aligned}\dot{\check{x}}_i(t) =& T_i^\top W_i^\top A_i W_i T_i \check{x}_i(t) + T_i^\top W_i^\top \sum_{j\neq i} M_{i,j} W_j T_j \check{x}_j(t) \\ & + T_i^\top W_i^\top B_i r_i(t) \quad (32)\\ y_i(t) =& C_i W_i T_i \check{x}_i(t) + D_i r_i(t), \quad (33)\end{aligned}$$

where $\check{x}_i \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^{k_i})$ is defined by $\check{x}_i = T_i^\top W_i^\top x_i(t)$.

It remains a future work to derive an error bound for ℓ_1-norm minimizing POD, which corresponds to Theorem 1 for the distributed POD.

Using Theorem 1, (29) and $\hat{M}_{\rm opt} = \hat{M}_{\rm simple} + \hat{M}_{\rm truncated}$, we have the approximation error bound for the ℓ_1-norm minimizing POD.

Theorem 2: Let $\lambda_{i,j} \in \mathbb{R}$ $(i = 1, \cdots, L;\ j = 1, \cdots, n_i)$ be singular values of R_i in (18) such that (24). Then, the error bound in (15) for the ℓ_1-norm minimizing POD is characterized by the inequality

$$\sum_{i=1}^{L}\int_0^T \left\| x_i(t) - T_i W_i \left(T_i W_i\right)^\top \check{x}_i(t)\right\|_2^2 dt \le \left(1 + T c'_\mu(T)\alpha^2\right)\sum_{i=1}^{L}\sum_{j=k_i+1}^{n_i} \lambda_{i,j}, \quad (34)$$

[2]In fact, ℓ_0-norm is not defined as a norm, since it does not satisfy the principle of norm.

where $c'_\mu(T)$ is defined by

$$c'_\mu(T) := \begin{cases} \frac{1}{b}\left(e^{2bt}-1\right) & (b \neq 0) \\ 2t & (b = 0) \end{cases},$$

$$b := \mu(M) + k\varepsilon + \sum_{i=1}^{L} \left\|T_i W_i^\top\right\|_2 \lambda_{\max}(A_i).$$

C. Algorithm for ℓ_1-Norm Minimizing POD

We summarize the ℓ_1-norm minimizing POD based on the discussions in Section V-B.

Algorithm 2:

Step 0: Set the criterion $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$ for the truncation (30).

Step 1: We apply the distributed POD to the original system Σ and compute the reduced-order system $\hat{\Sigma}$ by (10) and (11) and the network matrix $\hat{M}$ in (23) by Algorithm 1.

Step 2: For $\hat{M}$ in Step 1, we solve the ℓ_1 norm minimization problem and compute the optimal network matrix $\hat{M}_{\rm opt}$ by (28) and (29).

Step 3: For $\hat{M}_{\rm opt}$ in Step 2, we consider the truncation (30) and (31). By using this truncation, construct the simplified network matrix $\hat{M}_{\rm simple}$.

VI. Numerical Example

In this section, we show an efficiency of the distributed and network simplifying PODs by applying them to a linear networked dynamical system.

We consider a 200th-order linear networked dynamical system defined by subsystems Σ_i $(i = 1, \cdots, L)$ with

$$g(x_i, r_i) = \mathrm{diag}\left(h(x_{20(i-1)+1}, h(x_{20(i-1)+2}, \cdots, h(x_{20i}))\right),$$
$$h(x_j) = -0.25x_j,$$
$$n = 200,\ n_i = 40,\ L = 5$$

and network system Σ_0 with

$$M := \sigma\bar{M} \text{ and } \sigma = 0.7,$$

where $\bar{M}$ is a constant matrix generated by substituting $m_{i,j} = 1$ and $m_{j,i} = -1$ randomly to the 2.47% of the elements. This formulation implies that the subsystem Σ_i consists of the state variables $(x_{20(i-1)+1}, x_{20(i-1)+2}, \cdots, x_{20i})$. In addition, we set the initial condition $x_i(0)$ $(i = 1, \cdots, L)$ as uniformly distributed pseudo random values in $[0, 1.25]$ and no input is add, i.e. $r_i(t) = 0$ for all $i = 1, \cdots, L$. Based on the criterion (25), the dimensions of subsystems after the reduction are selected as $k_2 = 6$ and $k_i = 5$ $(i \neq 2)$ and the criteria for the truncation is given by $\varepsilon = 0.0116$.

The simulation result of original system, distributed POD, ℓ_1-norm optimization, and ℓ_1-norm minimizing POD after simplification are illustrated by the following 3-dimensional (3-D) plots. We can see that the results by proposed methods approximates the original system with an appropriate accuracy.

More details can be observed by showing the simulations of the magnitude of errors. Then, we see that both distributed POD and ℓ_1-norm optimization approximates Σ_4 which is the subsystem of our interest. However, the network simplifying POD does not necessarily approximates Σ_4 so good, but approximates the behavior of the entire system with a reasonable accuracy.

We also observe the preservation and simplification of the network structure. The results are illustrated as follow. The nonzero (i, j)th elements of the network matrices of original system and each method are plotted by blue circles, where the vertical and horizontal axes correspond to rows and columns of each matrix. Then, distributed POD preserves a network structure appropriately. Moreover, ℓ_1-norm minimizing POD simplifies the network structure with preserving the behavior of the entire system appropriately. The number of nonzero elements is 988 in the original system, however this was reduced to 96 by the network simplifying POD.

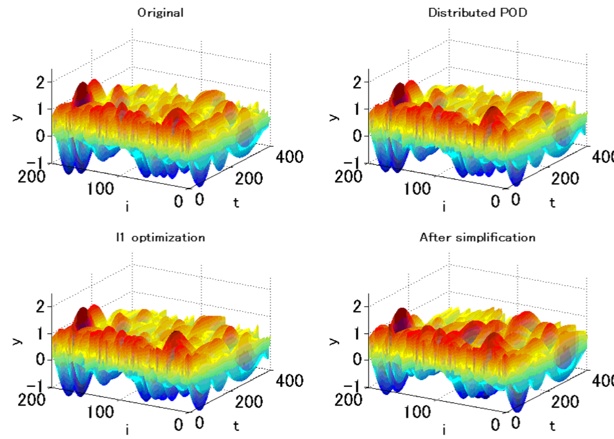

Figure	System
Upper left	Original system
Upper right	Distributed POD
Lower left	Distributed POD with ℓ_1-norm optimization
Lower right	ℓ_1-norm minimizing POD after simplification

Fig. 3. 3-D plots of original system and reduced systems by each reduction methods.

VII. Conclusions

In this paper, we have proposed the distributed POD which can specify the degree of approximate of each subsystem and preserve the network structure. In addition, ℓ_1-norm minimizing POD is proposed for constructing a simple network structure. We have seen that proposed reduction methods works well in a linear network dynamical system by a numerical example, however we could not have obtained a good approximation in the nonlinear case. This difficulty should be resolved in our future work.

The authors greatly appreciate Professor Shinji Hara of the University of Tokyo for his valuable comments and suggestions. This work was supported by Grant-in-Aids for Young Scientists (B) 22760313 and 25820177 of Japan Society for the Promotion of Science.

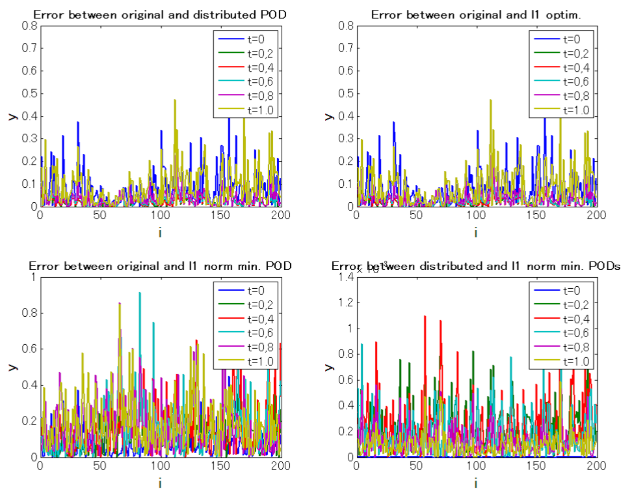

Figure	System
Upper left	Original system and distributed POD
Upper right	Original system and distributed POD with ℓ_1-norm optimization
Lower left	Original system and ℓ_1-norm minimizing POD
Lower right	Distributed POD and ℓ_1-norm minimizing POD

Fig. 4. Snapshots of errors at $t = 0, 0.2, 0.4, 0.6, 0.8, 1.0$

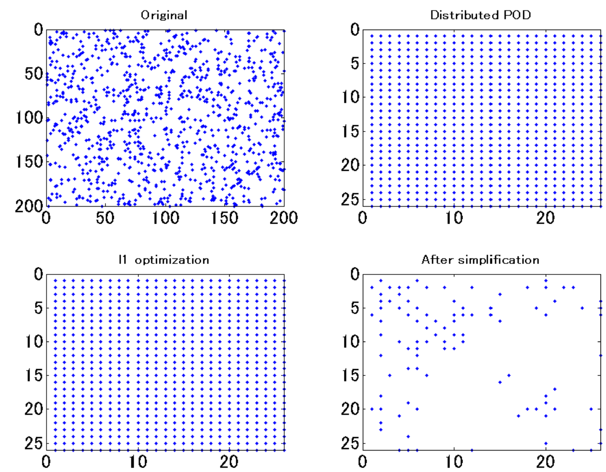

Figure	System	Nonzero elements
Upper left	Original system	988
Upper right	Distributed POD	676
Lower left	Distributed POD with ℓ_1-norm optimization	676
Lower right	ℓ_1-norm minimizing POD	96

Fig. 5. Nonzero elements of the network matrices of original system and proposed reduction methods

REFERENCES

[1] A.C. Antoulas. *Approximation of Large-Scale Dynamical Systems*. SIAM, 2005.

[2] S. Chaturantabut and D.C. Sorensen. A State Space Error Estimate for POD-DEIM Nonlinear Model Reduction. *SIAM Journal on Numerical Analysis*, Vol. 50, No. 1, pp. 46–63, 2012.

[3] R. Gonin and A.H. Money. *Nonlinear L_p-norm Estimation*. Marcel Dekker, 1989.

[4] S. Hara, T. Hayakawa, and H. Sugata. LTI Systems with Generalized Frequency Variables: A Unified Framework for Homogeneous Multi-agent Dynamical Systems. *SICE Journal of Control, Measurement, and System Integration*, Vol. 2, pp. 299–306, September 2009.

[5] S. Hara, J. Imura, and K. Tsumura. Multi-Resolved Dynamical System Theory for Large Scale Complex Systems. In *Proceedings of the SICE Annual Conference 2008*, Chofu, Japan, August 2008.

[6] S. Hara, H. Shimizu, and T-H. Kim. Consensus in Hierarchical Multi-agent Dynamical Systems with Low-rank Interconnections: Analysis of stability and convergence rates. In *Proceedings of 2009 American Control Conference*, St. Louis, USA, June 2009.

[7] T. Ishizaki, K. Kashima, J. Imura, and K. Aihara. Model Reduction of Multi-Input Dynamical Networks Based on Clusterwise Controllability. In *Proceedings of 2012 American Control Conference*, pp. 2301–2306.

[8] A.A. Julius, M. Zavlanos, S. Boyd, and G.J. Pappas. Genetic Network Identification Using Convex Programming. *IET Systems Biology*, Vol. 3, No. 3, p. 155.166, 2009.

[9] K. Kunisch and S. Volkwein. Galerkin Proper Orthogonal Decomposition Methods for a General Equation in Fluid Dynamics. *SIAM Journal on Numerical Analysis*, Vol. 40, No. 2, pp. 492–515, 2002.

[10] E.N. Lorenz. *Empirical Orthogonal Functions and Statistical Weather Prediction*, Vol. 1. 1956.

[11] T. Samad and A. Annaswamy. In *The Impact of Control Technology-Overview, Success Stories, and Research Challenges*. IEEE Control Systems Society, February 2011.

[12] H. Sandberg and R.M. Murray. Model Reduction of Interconnected Linear Systems. *Optimal Control Applications and Methods*, Vol. 30, pp. 225–245, 2009.

[13] H. Shimizu and S. Hara. Rank Properties of Inter-layer Incidence Matrix and Convergence Performance in Hierarchical Consensus. *Transactions of the Society of Instrument and Control Engineers*, Vol. 45, No. 9, pp. 476–483, September 2009. (in Japanese).

[14] R. Smith. A Hierarchical Cyclic Pursuit Scheme for Vehicle Networks. *Automatica*, Vol. 41, pp. 1045–1053, 2005.