Reaction-Diffusion Clustering of Single-Input Dynamical Networks

Takayuki Ishizaki, Kenji Kashima, Jun-ichi Imura* and Kazuyuki Aihara**

***Abstract*— A novel clustering method for single-input dynamical networks is proposed, where we aggregate state variables that behave similarly for any input signals. This clustering method is based on the Reaction-Diffusion transformation, which can be applied to large-scale networks, and preserves the stability as well as a kind of network structure of the original system. In addition, the upper bound of the state discrepancy caused by the clustering is evaluated in terms of H_∞-norm.**

I. Introduction

Dynamical systems on large-scale complex networks, whose behaviors are determined by the interaction of a large number of subsystems, have been widely studied over the past decades. Examples of such dynamical networks include World-Wide-Web, gene regulatory networks, spread of infection; see [1], [2], [3] for an overview. For such dynamical networks, it is crucial to address a clustering-based model reduction problem, i.e., a model reduction method in which each set of state variables clustered in a certain way is aggregated. This enables us to efficiently analyze the coarse properties of the original large-scale system such as the mean behaviors. This is because the new state variables of the reduced model express the system behaviors given by aggregating a certain set of the original state variables.

As one of such possible approaches, the state aggregation based on singular perturbation of dynamical networks have been intensively developed in [4], [5], [6]. However, this kind of approach cannot explicitly take account of the effect of the external input. Furthermore, a kind of structure-preserving model reduction methods have been developed. The papers [7], [8] address this kind of problem, more specifically, the problem of the order reduction of a dynamical network as well as preserving some underlying structure of systems such as the Lagrangian structure and the second-order structure. However, these methods only deal with the preservation of certain formula of differential equations. In addition, even though [9] discuss the reduction problem of each subsystem interconnected by a network, it requires a priori knowledge on clustering of the subsystems ([10] has somewhat relaxed the assumption) and it does not give a theoretical evaluation of the approximation accuracy. Egerestdt in [11] has also solved a similar problem from the controllability and graph theory points of view for a limited class of linear dynamical networks.

On the other hand, we propose a new type of network clustering method for reducing the dimension of a linear dynamical network and approximating its input-to-state mapping within a specified approximation error precision. In the proposed approach, the network structure transformation called a Reaction-Diffusion transformation, which has been proposed in [12] by the authors, is fully exploited to find a set of state variables that behaves similarly for any input signals, called a cluster set of nodes. Thus, the proposed approach does not need a priori knowledge on cluster sets. Furthermore, thanks to the numerical efficiency of the Reaction-Diffusion transformation, the proposed method can be applied to large-scale dynamical networks.

As the first step of the above approach, we have preliminarily discussed in [13] a network clustering problem. However in that paper, no theoretical error evaluation has been provided. In this paper, we formulate a network clustering problem in more general setting by introducing the notion of weak reducibility, and provides a solution to this problem.

This paper is organized as follows: In section II, we describe a systems, or the system to be investigated and recap fundamental results on the Reaction-Diffusion transformation. Section III poses and solves a network clustering problem, where the properties of the Reaction-Diffusion realization are utilized to solve the problem. In the last of the section, a numerical example demonstrates and validates the proposed method. Finally Section IV concludes this paper.

NOTATION: Let v be a vector, and $M_1, \ldots, M_n$ matrices. The following notation is used in this paper:

$\mathbb{R}$	the set of real numbers
I_n	the unit matrix of the size $n \times n$
e_k^n	the k-th column vector of I_n
$e_{k_1:k_2}^n$	the k_1-th to k_2-th columns of I_n
$\|M_1\|$	the maximum singular value of M_1
$\mathrm{diag}\,(v)$	the diagonal matrix whose diagonal entries are the entries of v
$\mathrm{Diag}\,(M_1, \ldots, M_n)$	the block diagonal matrix composed of $M_1, \ldots, M_n$

The H_∞-norm of a stable proper transfer function matrix $G(s)$ is defined by $\|G(s)\|_\infty := \sup_{\omega\in\mathbb{R}} \sigma_{\max}(G(j\omega))$. Let $\mathcal{I}$ be the set of integers, for which $|\mathcal{I}|$ denotes the cardinality of $\mathcal{I}$ and $e_{\mathcal{I}}^n \in \mathbb{R}^{n\times|\mathcal{I}|}$ denotes the matrix whose column vectors are composed of e_k^n for $k \in \mathcal{I}$ (in some order of k), i.e., $e_{\mathcal{I}}^n = \left[e_{k_1}^n, \ldots, e_{k_m}^n\right] \in \mathbb{R}^{n\times m}$ for $\mathcal{I} = \{k_1, \ldots, k_m\}$.

II. Introduction of Reaction-Diffusion Transformation

In this paper, we deal with linear systems on large-scale complex networks whose general form is given as follows:

*Graduate School of Information Science and Engineering, Tokyo Institute of Technology; 2-12-1, Meguro ward, Tokyo {ishizaki,kashima,imura}@cyb.mei.titech.ac.jp

**Institute of Industrial Science, University of Tokyo; 4-6-1 Komaba, Meguro ward, Tokyo aihara@sat.t.u-tokyo.ac.jp

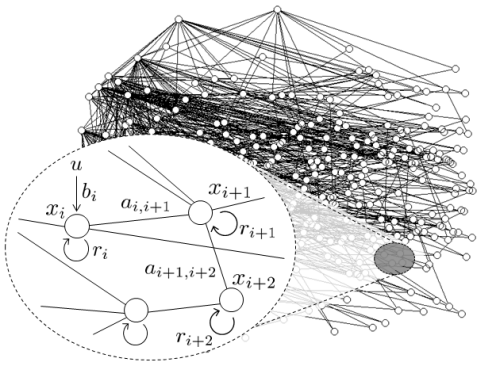

Fig. 1. Depiction of dynamical networks.

Definition 1: The linear system

$$\dot{x} = Ax + bu \tag{1}$$

with $A = \{a_{i,j}\} \in \mathbb{R}^{n\times n}$ and $b = \{b_i\} \in \mathbb{R}^n$ is said to be a *dynamical network* (A, b) if A is stable and symmetric. Moreover, if $a_{i,j}$ for $i \neq j$ and b_i are all non-negative, we call a *positive* dynamical network.

This is a generalization of *undirected reaction-diffusion* systems depicted in Fig. 1:

$$\dot{x}_i = -r_i x_i + \sum_{j=1, j\neq i}^{n} a_{i,j}\left(x_j - x_i\right) + b_i u \tag{2}$$

where $r_i(\geq 0)$ denotes the intensity of the reaction (chemical dissolution) of x_i, and $a_{i,j}(\geq 0)$ for $i \neq j$ denotes the intensity of the diffusion between x_i and x_j. This coupled dynamics is stable if at least one r_i is strictly positive and the graph is connected. See, e.g., [3] for a survey on networked systems and multi-agent systems. This reaction-diffusion structure over the network can be represented in the following *spatially one-dimensional* manner:

Definition 2: Let (A, b) be a dynamical network in (1). Then, unitary H is said to be *Reaction-Diffusion transformation matrix* if $\mathcal{A} := HAH^{\mathsf{T}}$ and $\mathcal{B} := Hb$ are in the form of

$$\begin{aligned}
\mathcal{A} &= \begin{bmatrix} \alpha_1 & \beta_1 & & & & \\ \beta_1 & \alpha_2 & \beta_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \beta_{n-1} & \\ & & & \beta_{n-1} & \alpha_n & \end{bmatrix} \in \mathbb{R}^{n\times n} \\
\mathcal{B} &= \begin{bmatrix} \beta_0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^n
\end{aligned} \tag{3}$$

with some negative constant α_i for $i \in \{1, \ldots, n\}$ and some non-negative constant β_i for $i \in \{0, \ldots, n-1\}$. Moreover, the realization $(\mathcal{A}, \mathcal{B})$ is called *Reaction-Diffusion realization*.

Hereafter, the term "Reaction-Diffusion" is denoted as "**RD**-". As shown in Theorem 1 in [12], we can effectively construct a **RD**-transformation matrix H. Moreover

$$\bar{i} := \begin{cases} \min_i\{i : \beta_i = 0\} \text{ if } \prod_{i=1}^{n-1}\beta_i = 0, \\ n, \text{ otherwise} \end{cases} \tag{4}$$

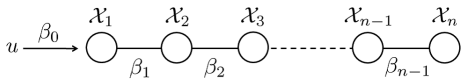

Fig. 2. Depiction of **RD**-realization.

does not depend on the choice of H. Actually, the column vectors of $H^{\mathsf{T}}e^n_{1:\bar{i}}$ span the controllable sub-space. That is, $\bar{i} = n$ if and only if (A, b) is controllable.

In what follows, we denote

$$\begin{cases} g(s) := (sI_n - A)^{-1} b \\ g_i(s) := (e^n_i)^{\mathsf{T}} g(s), \end{cases} \quad \begin{cases} \mathcal{G}(s) := (sI_n - \mathcal{A})^{-1} \mathcal{B} \\ \mathcal{G}_i(s) := (e^n_i)^{\mathsf{T}} \mathcal{G}(s). \end{cases} \tag{5}$$

Then, we see that the following *low-pass* property in the **RD**-realization:

Proposition 1 (Theorem 2 in [12]): Let $(\mathcal{A}, \mathcal{B})$ be **RD**-realization in (3). Then, $\mathcal{G}_i$ in (5) satisfies

$$\left\|\mathcal{G}_i(s)\right\|_\infty = \mathcal{G}_i(0), \quad \forall i \in \{1, \ldots, n\}. \tag{6}$$

Corollary 1: Consider the **RD**-transformation of the dynamical network (1). For $k \in \{1, \ldots, n\}$ and the **RD**-transformation matrix $H = \{h_{i,j}\}$

$$h_{1,i} = \cdots = h_{k-1,i} = 0, \quad h_{k,i} \neq 0 \tag{7}$$

holds if and only if the relative degree of g_i is k.

Proof: For the transfer function $\mathcal{G}_i$ in (5), we have $g_i(s) = \sum_{k=1}^n h_{k,i}\mathcal{G}_k(s)$. The result follows from the fact that the relative degree of $\mathcal{G}_k$ is k. ■

Corollary 1 indicates that we can identify the relative degree of $g_i(s)$, which is the transfer function from u to x_i in the original dynamical networks, by examining the column vectors of the **RD**-transformation matrix H. This further implies that if (7) holds, the distance (the smallest number of the edges) between the node having the input and the i-th node is k.

III. Application to Network Clustering

A. Network Clustering based on State Aggregation

In this subsection, we outline a network clustering method based on the aggregation of states. First, we define the following notion of network clustering:

Definition 3: Consider the dynamical network (A, b) in (1). A family of index sets $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ for $\mathbb{L} := \{1, \ldots, L\}$ is called a *cluster set* (its element is referred to as a cluster) if each element is a disjoint subset of $\{1, \ldots, n\}$ and $\bigcup_{l\in\mathbb{L}} \mathcal{I}_{[l]} = \{1, \ldots, n\}$. An *aggregation matrix* (*compatible with* $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$) is defined by

$$\mathsf{P} := \mathrm{Diag}\left(\mathsf{p}_{[1]}, \ldots, \mathsf{p}_{[L]}\right) Q \in \mathbb{R}^{\Delta\times n}, \quad \Delta := \sum_{l=1}^{L} \delta_l \tag{8}$$

with $\mathsf{p}_{[l]} \in \mathbb{R}^{\delta_l \times \left|\mathcal{I}_{[l]}\right|}$ such that $\delta_l \leq \left|\mathcal{I}_{[l]}\right|$ and $\mathsf{p}_{[l]}\mathsf{p}_{[l]}^{\mathsf{T}} = I_{\delta_l}$, and the permutation matrix

$$Q = \left[e^n_{\mathcal{I}_{[1]}}, \ldots, e^n_{\mathcal{I}_{[L]}}\right]^{\mathsf{T}} \in \mathbb{R}^{n\times n}, \quad e^n_{\mathcal{I}_{[l]}} \in \mathbb{R}^{n\times\left|\mathcal{I}_{[l]}\right|}. \tag{9}$$

Then, *the aggregated model* (*associated with* P) of the dynamical network (A, b) in (1) is given by

$$\left(\mathsf{P}A\mathsf{P}^{\mathsf{T}}, \mathsf{P}b\right). \tag{10}$$

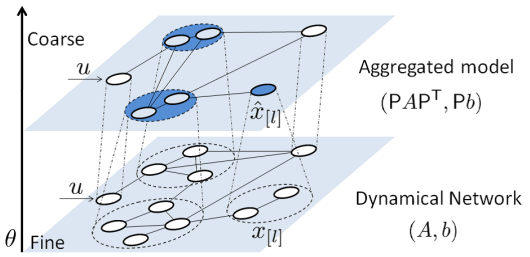

Fig. 3. Illustration of network clustering based on state aggregation.

In this definition, there are L clusters labeled by $\mathbb{L}$. Then, each node (state variable) belongs to exactly one of them, or equivalently, the behavior of l-th cluster in the original system is represented by $x_{[l]} := (e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} x$. On the other hand, the aggregated model has the same number of clusters as that of the original system with the state variable $\hat{x}_{[l]} = \mathsf{p}_{[l]} x_{[l]}$. Note that $\mathsf{P}A\mathsf{P}^{\mathsf{T}}$ is symmetric and the aggregation matrix P clearly satisfies $\mathsf{P}\mathsf{P}^{\mathsf{T}} = I_\Delta$. In what follows, we derive a condition under which $x_{[l]}$ can be recovered from $\hat{x}_{[l]}$ in a suitable sense. From the model reduction points of views, a small δ_l is desirable for reducing the order of dynamical networks.

Remark 1: Most of traditional model reduction methods, such as the balanced truncation, Hankel-norm approximation and Krylov projection [14], which provide a reduced model appropriately approximating the input-to-output mapping of a given system, does not preserve the *spatial information* of the original system. In other words, the network structure of the system is destroyed through the reduction. On the contrary, we propose a model reduction based on the state aggregation. As shown in Fig. 3, the network structure (spatial distribution) of internal states is retained through the reduction. We refer to such state aggregation of dynamical networks as *network clustering*.

Hereafter, we denote the transfer function from the input to the state of the aggregated model by

$$\begin{cases} \mathsf{g}(s) := \mathsf{P}^{\mathsf{T}} \left(sI_\Delta - \mathsf{P}A\mathsf{P}^{\mathsf{T}}\right)^{-1} \mathsf{P}b \\ \mathsf{g}_i(s) := (e^n_i)^{\mathsf{T}} \mathsf{g}(s). \end{cases} \tag{11}$$

Furthermore, the DC gain vector of the states of the **RD**-realization, whose elements represent the maximum gain of $\mathcal{G}_i(s)$ as shown in Proposition 1, is denoted by

$$\mathfrak{g} := -\mathcal{A}^{-1}\mathcal{B} \in \mathbb{R}^n, \tag{12}$$

which can be efficiently obtained by solving $\mathcal{A}\mathfrak{g} + \mathcal{B} = 0$ with exploiting the structure of $\mathcal{A}$ and $\mathcal{B}$ [14], [15]. Then, let us begin with the simple situation where some of the clusters in the original dynamical network have redundancy as in the following sense:

Definition 4: Under Definition 3, if there exists a row-fullrank matrix $\mathsf{q}_{[l]} \in \mathbb{R}^{\left(|\mathcal{I}_{[l]}|-\delta_l\right)\times|\mathcal{I}_{[l]}|}$ such that

$$\mathsf{q}_{[l]} \left(e^n_{\mathcal{I}_{[l]}}\right)^{\mathsf{T}} g(s) = 0, \tag{13}$$

then the cluster $\mathcal{I}_{[l]}$ is said to be *reducible*.

The following theorem characterizes the reducibility of $\mathcal{I}_{[l]}$ via the **RD**-transformation:

Theorem 1: Consider the **RD**-transformation of the dynamical network (1). Define

$$\mathsf{H}^{\mathfrak{g}}_{[l]} := \mathrm{diag}(\mathfrak{g})\, He^n_{\mathcal{I}_{[l]}} \in \mathbb{R}^{n\times|\mathcal{I}_{[l]}|}. \tag{14}$$

Then, (13) is equivalent to

$$\mathsf{q}_{[l]} \left(\mathsf{H}^{\mathfrak{g}}_{[l]}\right)^{\mathsf{T}} = 0. \tag{15}$$

Furthermore, for each $l \in \mathbb{L}$, take a unitary matrix $[\mathsf{p}^{\mathsf{T}}_{[l]}, \mathsf{q}^{\mathsf{T}}_{[l]}]^{\mathsf{T}} \in \mathbb{R}^{|\mathcal{I}_{[l]}|\times|\mathcal{I}_{[l]}|}$ satisfying (15). Then, the aggregated model associated with P in (8) is stable and satisfies

$$g(s) = \mathsf{g}(s). \tag{16}$$

Proof: [Necessity of (15)] We have

$$\mathsf{q}_{[l]} \left(e^n_{\mathcal{I}_{[l]}}\right)^{\mathsf{T}} g(s) = \mathsf{q}_{[l]} \left(e^n_{\mathcal{I}_{[l]}}\right)^{\mathsf{T}} H^{\mathsf{T}} \mathcal{G}(s).$$

The necessity follows from the following facts:

- $\mathcal{G}_i = 0$ for $i > \bar{i}$, and $\{\mathcal{G}_i\}_{i=1}^{\bar{i}}$, where $\bar{i}$ is defined in (4), are linearly independent,
- the i-th entry $\mathfrak{g}_i$ of $\mathfrak{g}$ is 0 if and only if $i > \bar{i}$.

[Sufficiency of (15)] Denoting

$$\left(e^{|\mathcal{I}_{[l]}|-\delta_l}_i\right)^{\mathsf{T}} \mathsf{q}_{[l]} = \left[\mathsf{q}^i_{1[l]}, \ldots, \mathsf{q}^i_{|\mathcal{I}_{[l]}|[l]}\right] \in \mathbb{R}^{1\times|\mathcal{I}_{[l]}|}$$

$$He^n_{\mathcal{I}_{[l]}} = \begin{bmatrix} \mathsf{h}^1_{1[l]} & \cdots & \mathsf{h}^1_{|\mathcal{I}_{[l]}|[l]} \\ \vdots & \ddots & \vdots \\ \mathsf{h}^n_{1[l]} & \cdots & \mathsf{h}^n_{|\mathcal{I}_{[l]}|[l]} \end{bmatrix} \in \mathbb{R}^{n\times|\mathcal{I}_{[l]}|},$$

condition (15) implies

$$\sum_{j=1}^{|\mathcal{I}_{[l]}|} \mathsf{q}^i_{j[l]} \mathsf{h}^k_{j[l]} \mathfrak{g}_k = 0, \quad \begin{cases} \forall i \in \{1, \ldots, |\mathcal{I}_{[l]}| - \delta_l\} \\ \forall k \in \{1, \ldots, n\}. \end{cases}$$

By Proposition 1, we have

$$\left\| \left(e^{|\mathcal{I}_{[l]}|-\delta_l}_i\right)^{\mathsf{T}} \mathsf{q}_{[l]} \left(e^n_{\mathcal{I}_{[l]}}\right)^{\mathsf{T}} g(s) \right\|_\infty$$

$$= \left\| \sum_{k=1}^{n} \sum_{j=1}^{|\mathcal{I}_{[l]}|} \mathsf{q}^i_{j[l]} \mathsf{h}^k_{j[l]} \mathcal{G}_k(s) \right\|_\infty \le \sum_{k=1}^{n} \left| \sum_{j=1}^{|\mathcal{I}_{[l]}|} \mathsf{q}^i_{j[l]} \mathsf{h}^k_{j[l]} \mathfrak{g}_k \right|$$

where the right-hand side is 0.

[Proof of (16)] The stability of g is trivial from the negative definiteness of A. Considering the coordinate transformation by unitary $[\mathsf{P}^{\mathsf{T}}, \overline{\mathsf{P}}^{\mathsf{T}}]^{\mathsf{T}}$, we have

$$\begin{aligned} g(s) &= \mathsf{g}(s) + \Xi(s)\, \overline{\mathsf{P}} \left(sI_n - A\right)^{-1} b \\ \Xi(s) &= \mathsf{P}^{\mathsf{T}} \left(sI_\Delta - \mathsf{P}A\mathsf{P}^{\mathsf{T}}\right)^{-1} \mathsf{P}A\overline{\mathsf{P}}^{\mathsf{T}} + \overline{\mathsf{P}}^{\mathsf{T}}. \end{aligned} \tag{17}$$

Note that $\Xi(s)$ is stable. Define $\overline{\mathsf{P}}$ by replacing $\mathsf{p}_{[l]}$ by $\mathsf{q}_{[l]}$ for each $l \in \mathbb{L}$ in (8). Then, $[\mathsf{P}^{\mathsf{T}}, \overline{\mathsf{P}}^{\mathsf{T}}]^{\mathsf{T}}$ is unitary, and $\overline{\mathsf{P}}\left(sI_n - A\right)^{-1} b = 0$ by the reducibility of (13). ■

Theorem 1 implies that the reducibility of the cluster $\mathcal{I}_{[l]}$ is characterized by the (column) rank deficiency of the matrix $\mathsf{H}^{\mathfrak{g}}_{[l]}$, which is composed of $|\mathcal{I}_{[l]}|$ column vectors of $\mathrm{diag}(\mathfrak{g})H$ obtained through the **RD**-transformation. This further implies that a lower order aggregated model is obtained if $\mathsf{H}^{\mathfrak{g}}_{[l]}$ for each $l \in \mathbb{L}$ has lower rank.

B. Cluster Determination

Hereafter, we suppose $\delta_l = 1$ for all $l \in \mathbb{L}$. This means that every cluster is aggregated into only one variable. In this section, aiming at significant order reduction, we relax the reducibility of $\mathcal{I}_{[l]}$ in (13) through its equivalent characterization in (15). Let $\mathsf{h}_{j[l]}$ and $\mathsf{h}^{\mathfrak{g}}_{j[l]}$ for $j \in \{1, \ldots, |\mathcal{I}_{[l]}|\}$ denote a column vector of $He^n_{\mathcal{I}_{[l]}}$ and $\mathsf{H}^{\mathfrak{g}}_{[l]}$, namely

$$\begin{aligned} He^n_{\mathcal{I}_{[l]}} &= \left[\mathsf{h}_{1[l]}, \ldots, \mathsf{h}_{|\mathcal{I}_{[l]}|[l]}\right], \\ \mathsf{H}^{\mathfrak{g}}_{[l]} &= \left[\mathsf{h}^{\mathfrak{g}}_{1[l]}, \ldots, \mathsf{h}^{\mathfrak{g}}_{|\mathcal{I}_{[l]}|[l]}\right], \quad \mathsf{h}_{j[l]}, \mathsf{h}^{\mathfrak{g}}_{j[l]} \in \mathbb{R}^n. \end{aligned} \tag{18}$$

We impose the following less-restrictive assumption: for at least one $j \in \{1, \ldots, |\mathcal{I}_{[l]}|\}$, $\mathfrak{g}^{\mathsf{T}}\mathsf{h}_{j[l]} \neq 0$ for all $l \in \mathbb{L}$; thus without loss of generality $\mathfrak{g}^{\mathsf{T}}\mathsf{h}_{1[l]} \neq 0$ for all $l \in \mathbb{L}$. Actually, $\mathfrak{g}^{\mathsf{T}}\mathsf{h}_{j[l]}$ is always positive for all j in the case of *positive* dynamical networks; see Corollary 2 below.

Definition 5: Consider the dynamical network (1). The cluster $\mathcal{I}_{[l]}$ is *θ-weakly reducible* if $\mathfrak{g}^{\mathsf{T}}\mathsf{h}_{1[l]} \neq 0$ and

$$\left\| \mathsf{h}^{\mathfrak{g}}_{j[l]} - \frac{\mathfrak{g}^{\mathsf{T}}\mathsf{h}_{j[l]}}{\mathfrak{g}^{\mathsf{T}}\mathsf{h}_{1[l]}} \mathsf{h}^{\mathfrak{g}}_{1[l]} \right\| \leq \theta, \ \forall j \in \{1, 2, \ldots, |\mathcal{I}_{[l]}|\}. \tag{19}$$

In this definition, the constant θ represents the distance from the reducibility. The following lemma shows that (19) is equivalent to (15) if $\theta = 0$:

Lemma 1: Let $\delta_l = 1$ and $\mathfrak{g}^{\mathsf{T}}\mathsf{h}_{1[l]} \neq 0$. Then, $\mathcal{I}_{[l]}$ is reducible if and only if it is 0-weakly reducible.

Proof: When $\delta_l = 1$, the $\mathcal{I}_{[l]}$ is reducible if and only if the (column) rank of $\mathsf{H}^{\mathfrak{g}}_{[l]}$ is 1. Namely, for all j, there exist $c_j \in \mathbb{R}$ such that $\mathsf{h}^{\mathfrak{g}}_{j[l]} = c_j \mathsf{h}^{\mathfrak{g}}_{1[l]}$. Note $\mathfrak{g}^{\mathsf{T}}\mathsf{h}_{j[l]}$ is equal to the sum of all entries of $\mathsf{h}^{\mathfrak{g}}_{j[l]}$. Hence, c_j should be given by $c_j = \mathfrak{g}^{\mathsf{T}}\mathsf{h}_{j[l]} / \mathfrak{g}^{\mathsf{T}}\mathsf{h}_{1[l]}$. ∎

Here, we propose to construct a cluster set such that all clusters are θ-weakly reducible. This notion yields an aggregated model having the following properties:

Theorem 2: Consider the **RD**-transformation of the dynamical network (1). Suppose the cluster $\mathcal{I}_{[l]}$ for all $l \in \mathbb{L}$ is θ-weakly reducible, and define

$$\mathsf{p}_{[l]} = \frac{\hat{p}_{[l]}}{\left\|\hat{p}_{[l]}\right\|} \in \mathbb{R}^{1 \times |\mathcal{I}_{[l]}|}, \quad \hat{p}_{[l]} := \mathfrak{g}^{\mathsf{T}} He^n_{\mathcal{I}_{[l]}}. \tag{20}$$

Then, the aggregated model associated with P in (8) satisfies

$$g(0) = \mathsf{g}(0), \quad \|g(s) - \mathsf{g}(s)\|_\infty \leq \alpha\theta \tag{21}$$

for a positive constant α.

Proof: [Proof of the preservation of the DC gain] The desired result is $A^{-1}b = \mathsf{P}^{\mathsf{T}} \left(\mathsf{P}A\mathsf{P}^{\mathsf{T}}\right)^{-1} \mathsf{P}b$. Clearly, it suffices to show

$$b = A\mathsf{P}^{\mathsf{T}} \left(\mathsf{P}A\mathsf{P}^{\mathsf{T}}\right)^{-1}. \tag{22}$$

By direct calculation, we have $QA^{-1}b = \left[\hat{p}_{[1]}, \ldots, \hat{p}_{[L]}\right]^{\mathsf{T}}$. Therefore, (22) can be rewritten as

$$\left[\hat{p}_{[1]}, \ldots, \hat{p}_{[L]}\right]^{\mathsf{T}} = \left[\mathsf{p}_{[1]}\mathsf{p}^{\mathsf{T}}_{[1]}\hat{p}_{[1]}, \ldots, \mathsf{p}_{[L]}\mathsf{p}^{\mathsf{T}}_{[L]}\hat{p}_{[L]}\right]^{\mathsf{T}}.$$

This equality can be easily verified.

[Proof of the error evaluation] We prove based on (17). Note that $\|\mathsf{P}\| = \left\|\overline{\mathsf{P}}\right\| = 1$ and

$$\left\| \left(sI_\Delta - \mathsf{P}A\mathsf{P}^{\mathsf{T}}\right)^{-1} \right\|_\infty = \left\| \left(\mathsf{P}A\mathsf{P}^{\mathsf{T}}\right)^{-1} \right\| \leq \left\|A^{-1}\right\|,$$

which follows from Lemma 1 in [12] and Cauchy interlacing theorem (see Proposition 3.26 in [14]). This means that $\|\Xi(s)\|_\infty$ in (17) is bounded by a positive constant that does not depend on P. Thus, it suffices to show

$$\left\| \overline{\mathsf{P}} (sI_n - A)^{-1} b \right\|_\infty \leq c\theta \tag{23}$$

for a positive constant c. The matrix $\mathsf{H}^{\mathfrak{g}}_{[l]}$ in (18) can be rewritten as

$$\mathsf{H}^{\mathfrak{g}}_{[l]} = \frac{\mathsf{h}^{\mathfrak{g}}_{1[l]}}{\mathfrak{g}^{\mathsf{T}}\mathsf{h}_{1[l]}} \mathfrak{g}^{\mathsf{T}} He^n_{\mathcal{I}_{[l]}} + \left[0, v_2, \ldots, v_{|\mathcal{I}_{[l]}|}\right]$$

where

$$v_j := \mathsf{h}^{\mathfrak{g}}_{j[l]} - \frac{\mathfrak{g}^{\mathsf{T}}\mathsf{h}_{j[l]}}{\mathfrak{g}^{\mathsf{T}}\mathsf{h}_{1[l]}} \mathsf{h}^{\mathfrak{g}}_{1[l]}.$$

Noting $\mathfrak{g}^{\mathsf{T}} He^n_{\mathcal{I}_{[l]}} = \hat{p}_{[l]}$, we take a unitary matrix $[\mathsf{p}^{\mathsf{T}}_{[l]}, \mathsf{q}^{\mathsf{T}}_{[l]}]^{\mathsf{T}}$, for which

$$\mathsf{q}_{[l]} \left(\mathsf{H}^{\mathfrak{g}}_{[l]}\right)^{\mathsf{T}} = \mathsf{q}_{[l]} \left[0, v_2, \ldots, v_{|\mathcal{I}_{[l]}|}\right]^{\mathsf{T}}$$

holds. The definition of θ-weakly reducible clusters implies $\|v_j\| \leq \theta$ for all $j \in \{2, \ldots, |\mathcal{I}_{[l]}|\}$. Hence, (23) follows from the same argument as that in the proof of Theorem 1. ∎

Theorem 2 indicates that by taking the aggregation matrix as in (20), we can construct the aggregated model such that the difference between g and g is linearly bounded by θ and their DC gains are identical.

Remark 2: In this network clustering method, the value of the transfer function at 0-frequency is matched. This property is similar to that of moment matching methods [16], including the Krylov projection methods [14]. It should be emphasized that most of the moment matching methods do not provide the *global* error bound like as (21) instead of matching the transfer function at some points in the complex plane. On the contrary, this method enables to provide the error bound by exploiting the particular properties of the **RD**-transformation. In addition, since we evaluated the state discrepancy, $\hat{y} = C\mathsf{P}^{\mathsf{T}}\hat{x}$ is apparently close to $y = Cx$ for any C and input.

Furthermore, for positive dynamical networks (see Definition 1), the proposed clustering method can retain the relative degree of the original system as follows:

Corollary 2: Consider the **RD**-transformation of the *positive* dynamical network (1). Under the same notation in Theorem 2, assume that g_i for all $i \in \mathcal{I}_{[l]}$ have the same relative degree. Then, the aggregated model is again a positive dynamical network for which g_i and g_i have the same relative degree for every $i \in \{1, \ldots, n\}$.

Proof: For the relative degree, it suffices to show that the edges connecting the clusters $\mathcal{I}_{[i]}$ and $\mathcal{I}_{[j]}$ do not vanish

through the aggregation. The condition for preserving edges between $\mathcal{I}_{[i]}$ and $\mathcal{I}_{[j]}$ is represented as

$$\mathsf{p}_{[i]} \left(e^n_{\mathcal{I}_{[i]}}\right)^{\mathsf{T}} A e^n_{\mathcal{I}_{[j]}} \mathsf{p}^{\mathsf{T}}_{[j]} \neq 0 \tag{24}$$

for all $i, j \in \{1, \ldots, n\}$ such that $(e^n_{\mathcal{I}_{[i]}})^{\mathsf{T}} A e^n_{\mathcal{I}_{[j]}} \neq 0, i \neq j$. Here, the negative definiteness, the irreducibility and the non-negativity of the off-diagonal entries of A imply the positivity of the entries of $-A^{-1}$ [18]. Note that $\hat{p}_{[i]}$ in (20) can be rewritten as $\hat{p}_{[i]} = -b^{\mathsf{T}} A^{-1} e^n_{\mathcal{I}_{[i]}}$, which implies that $\mathsf{p}_{[i]}$ has positive entries. Hence, (24) follows from the non-negative property of A. The non-negativity of $\mathsf{P}A\mathsf{P}^{\mathsf{T}}$ and $\mathsf{P}b$ follows from the positivity of $\mathsf{p}_{[i]}$. ■

From the viewpoint of the approximation of the input-output properties, Corollary 2 guarantees the preservation of the high-frequency properties. In particular for dynamical networks, it further implies the preservation of the distance (the smallest number of the edges) between the node having an input and all the other nodes. Therefore, by combining Theorem 2 and Corollary 2, for a positive dynamical network, we can derive a clustered positive dynamical network not only approximating the overall input-to-output mapping of the original system with a specified error bound, but also having the same DC gain, and the same relative degree as the original ones.

The algorithm of the proposed network clustering method is as follows:

(a) Calculate **RD**-transformation by applying Theorem 1 in [12], and then find $\mathrm{diag}\,(\mathfrak{g})\, H$ in (14).
(b) Fix the value of a positive constant θ in Theorem 2 as the coarseness of the aggregated model.
(c) Find a θ-weakly reducible cluster set $\{\mathcal{I}_{[l]}\}_{l \in \mathbb{L}}$ (such that g_i for all $i \in \mathcal{I}_{[l]}$ have the same relative degree, for positive dynamical network case.)
(d) Derive the aggregation matrix P in (8) with (20).
(e) Construct the aggregated model $(\mathsf{P}A\mathsf{P}^{\mathsf{T}}, \mathsf{P}b)$ with $C\mathsf{P}^{\mathsf{T}}$.

C. Numerical Example; Network Clustering for Dynamical System on Complex Network

We consider a dynamical system on the complex network of a Holme-Kim model in Fig. 4, which is well-known as an extension of the Barabasi-Albert model, and has the high cluster coefficient as well as the scale-free and small-world properties [1]. The model in Fig. 4 has 3000 nodes and 6000 edges in which some hubs are included and the first node is connected to every other node within 6 edges. In the figure, each node is ordered accordingly to the distance from the first node. Construct the positive dynamical network (A, b) in (2) by taking

$$a_{i,j} = \begin{cases} 1, \text{ if nodes } i \text{ and } j \text{ are connected} \\ 0, \text{ otherwise,} \end{cases} \quad \text{for } i \neq j$$
$$r_1 = 1, \quad r_i = 0, \quad i \neq 1$$
$$b = [1, 0, \ldots, 0]^{\mathsf{T}} \in \mathbb{R}^{3000}.$$

By implementing a *greedy* algorithm, we determine θ-weakly reducible clusters inside each layer in Fig. 4. Fig. 5 shows the number of clusters versus the coarseness index θ. Fig. 6 and Fig. 7 show the clustered networks for $\theta = 0.3$ and $\theta = 3$, respectively. The order of the dynamical network is reduced to $|\mathbb{L}| = 344$ and $|\mathbb{L}| = 49$ from $n = 3000$. Fig. 8 shows the distribution of cluster size $\left|\mathcal{I}_{[l]}\right|$. In both cases, the maximal clusters are in the 5-th layer.

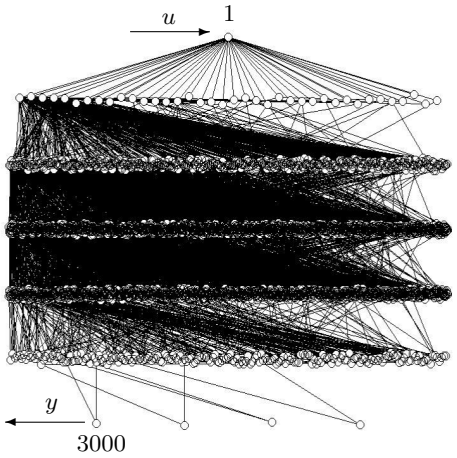

Fig. 4. Dynamical network on Holme-Kim model (3000 nodes).

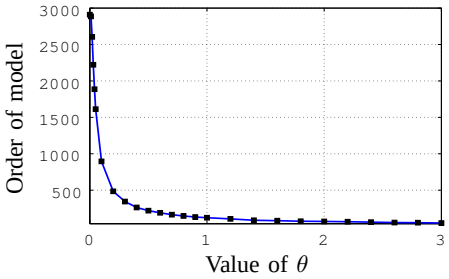

Fig. 5. Order of resultant models versus values of θ.

Fig. 9 shows the Bode diagrams of the original dynamical network (3000-th order, solid line) and the aggregated models (344-th and 49-th order, the line of $*$ and $\circ$) at the 3000-th node. From this figure, we can see that the low- and high-frequency properties are exactly retained, and also the overall properties are almost identical in both cases.

IV. Conclusion

In this paper, a network clustering method for linear dynamical networks has been proposed by using Reaction-Diffusion transformation. In this method, from the control theory points of view, the sets of states that behave similarly for any input signals are interpreted as sets of uncontrollable (or weakly controllable) states. Moreover, it has been shown that such states are efficiently found via the Reaction-Diffusion transformation. The method aggregates these states without loss of the network structure, in which a cluster-wise system description is obtained. In addition, the stability and an H_∞-norm approximation error bound is guaranteed.

Acknowledgment

This research is partially supported by the Aihara Innovative Mathematical Modelling Project, the Japan Society for the Promotion of Science (JSPS) through the "Funding

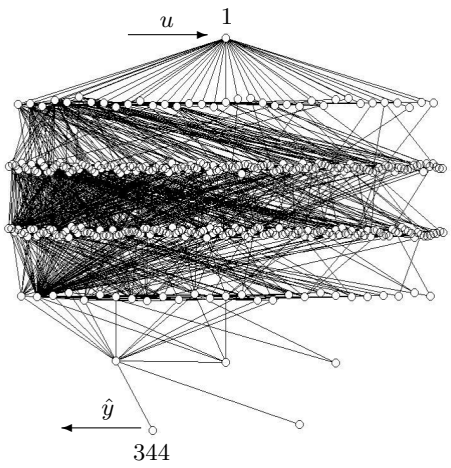

Fig. 6. Aggregated model of Holme-Kim model ($\theta = 0.3$), (344 clusters).

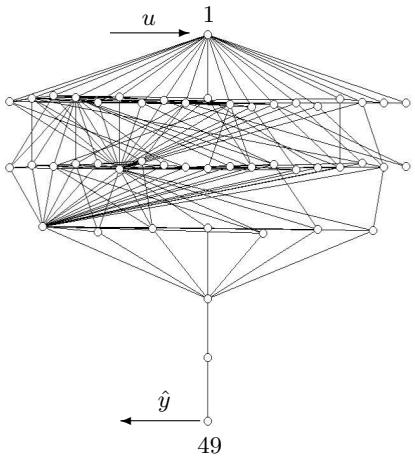

Fig. 7. Aggregated model of Holme-Kim model ($\theta = 3$) (49 clusters).

Program for World-Leading Innovative R&D on Science and Technology (FIRST Program)," initiated by the Council for Science and Technology Policy (CSTP). The authors also would like to thank Prof. M. Egerstedt for fruitful discussions.

REFERENCES

[1] S. Boccalettia, V. Latorab, Y. Morenod, M. Chavezf, and D. U. Hwang, "Complex networks: Structure and dynamics," *Physics Reports*, vol. 424-4-5, pp. 175–308, 2006.

[2] N. Masuda and N. Konno, *Complex Network*. Kindai Kagaku sha Co., Ltd, 2010.

[3] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*. Princeton Univerwity Press, 2010.

[4] N. Aoki, "Control of large-scale dynamic systems by aggregation," *IEEE Transactions on Automatic Control*, vol. 13-3, pp. 246–253, 1968.

[5] J. Chow and P. Kokotovic, "Time scale modeling of sparse dynamic networks," *IEEE Transactions on Automatic Control*, vol. 30-8, pp. 714–722, 1985.

[6] E. Biyik and M. Arcak, "Area aggregation and time scale modeling for sparse nonlinear networks," in *45th IEEE Conference on Decision and Control*, 2006, pp. 4046–4051.

[7] S.Lall, P. Krysl, and J. E. Marsden, "Structure-preserving model reduction for mechanical systems," *Physica D: Nonlinear Phenomena*, vol. 184-1-4, pp. 304–318, 2003.

[8] R. C. Li and Z. Bai, "Structure-preserving model reduction," *Applied Parallel Computing*, vol. 3732, pp. 323–332, 2006.

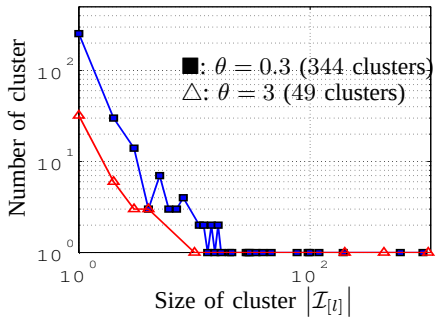

Fig. 8. Distribution of number of clusters versus size of clusters.

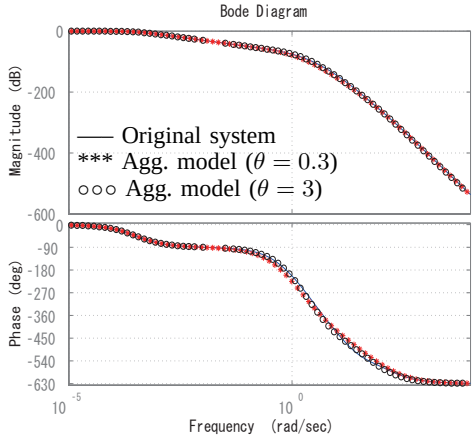

Fig. 9. Bode diagrams of dynamical network on Holme-Kim model and aggregated models.

[9] H. Sandberg and M. Murray, "Model reduction of interconnected linear systems," *Optimal Control Applications and Methods*, vol. 30-3, pp. 225–245, 2009.

[10] E. Yeung, J. Gonalves, H. Sandberg, and S. Warnick, "Network structure preserving model reduction with weak a priori structural information," in *Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, 2009, pp. 3256–3263.

[11] M. Egerstedt, "Controllability of networked systems," in *Mathematical Theory of Networks and Systems*, 2010.

[12] T. Ishizaki, K. Kashima, J. Imura, and K. Aihara, "Model order reduction for MIMO linear dynamical networks via reaction-diffusion transformation," in *Proceedings of American Control Conference*, 2011 (to appear).

[13] ——, "Network clustering for SISO linear dynamical networks via reaction-diffusion transformation," in *18th IFAC World Congress*, 2011 (to appear).

[14] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Society for Industrial Mathematics, 2005.

[15] H. S. Stone, "An efficient parallel algorithm for the solution of a tridiagonal linear system of equations," *Journal of the ACM*, vol. 20-1, pp. 27–38, 1973.

[16] A. Astolfi, "Model reduction by moment matching for linear and nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 55-10, pp. 2321–2336, 2010.

[17] S. Gugercin, "An iterative SVD-Krylov based method for model reduction of large-scale dynamical systems," in *44th IEEE Conference on Decision and Control*, 2005, pp. 5905–5910.

[18] L. Farina and S. Rinaldi, *Positive Linear Systems: Theory and Applications*, P. H. abd H. Hochstadt, Ed. Wiley-Interscience Publication, 2000.