Clustering-based $\mathcal{H}_2$-State Aggregation of Positive Networks and Its Application to Reduction of Chemical Master Equations

Takayuki Ishizaki[†1], Kenji Kashima[2], Antoine Girard[3], Jun-ichi Imura[1], Luonan Chen[4] and Kazuyuki Aihara[5]

***Abstract*— In this paper, based on a notion of network clustering, we propose a state aggregation method for positive systems evolving over directed networks, which we call positive networks. In the proposed method, we construct a set of clusters (i.e., disjoint sets of state variables) according to a kind of local uncontrollability of systems. This method preserves interconnection topology among clusters as well as stability and some particular properties, such as system positivity and steady-state characteristic (steady-state distribution). In addition, we derive an $\mathcal{H}_2$-error bound of the state discrepancy caused by the aggregation. The efficiency of the proposed method is shown through the reduction of a chemical master equation representing the time evolution of the Michaelis-Menten chemical reaction system.**

I. Introduction

Dynamical systems evolving on large-scale networks, whose behaviors are determined by the interaction of a number of subsystems, have been widely studied over the past decades. Examples of the network systems include social networks, biology networks, spread of infection; see [1] for an overview. For such systems, it is crucial to address a network structure preserving model reduction problem. This structured reduction enables us to analyze coarse properties of the given large-scale networks such as averaged behavior.

As one of possible approaches, singular perturbation-based state aggregation techniques have been studied in, e.g., [2], [3]. However, this kind of approach does not explicitly take into account the effect of external inputs. Furthermore, many kinds of structure-preserving model reduction methods have been considered in literature. For example, the papers [4], [5] address a model reduction problem while preserving some underlying structure of systems such as the Lagrangian structure and the second-order structure. However, these problems are not formulated based on the premise of the network structure. In addition, even though the paper [6] has proposed a network structure preserving reduction for network systems, a priori knowledge on partitioning (clustering) of the subsystems is required. Moreover, the relation between partitioning and the reduction error is not discussed.

In contrast with these existing approaches, this paper proposes a state aggregation method for positive systems evolving on directed networks, which we call positive networks, based on a notion of network clustering. In this method, we construct a set of clusters (i.e., disjoint subsets of state variables) according to local uncontrollability of systems. The aggregation of the constructed clusters under suitable weights provides a reduced model that preserves interconnection topology among clusters as well as stability and some particular properties, such as system positivity and steady-state characteristic. In addition, we derive an $\mathcal{H}_2$-error bound of the state discrepancy caused by the aggregation. The method proposed in this paper coincides with a generalization of our state aggregation approach proposed in [8], where systems evolving on undirected networks are considered and the resultant aggregation error is evaluated in terms of the $\mathcal{H}_\infty$-norm.

Furthermore, this paper also provides an application to the reduction of chemical master equations (CMEs), which describe the time evolution of chemical reactions by a set of linear ordinary differential equations. The CME expression of chemical reactions belongs to a class of continuous-time Markovian processes, where each state variable represents the realization probability of a molecule distribution. The state transition diagram can be regarded as a positive network on a multidimensional lattice (see Section IV for the details). The proposed method provides a reduced model that preserves specific properties of the CMEs such as system positivity and steady-state distribution.

This paper is organized as follows: In Section II, we describe a positive system evolving on directed networks, called positive networks, and formulate a clustering-based state aggregation problem for this class of systems. In Section III, using a projected controllability gramian of positive networks, we devise a state aggregation method where we construct clusters redundant in a suitable sense. In addition, we provide an algorithm to construct the redundant clusters. Section IV shows the efficiency of the proposed method by applying it to the reduction of chemical master equations, which belong to a class of the positive networks. Finally, Section V concludes this paper.

NOTATION The following notation is to be used: $\mathbb{R}$: the set of real numbers, $\mathbb{R}_+$ ($\mathbb{R}_{0+}$): the set of positive (non-negative) real numbers, $\mathbb{N}$: the set of natural numbers, I_n:

[†]Research Fellow of the Japan Society for the Promotion of Science

[1]Department of Mechanical and Environmental Informatics, Graduate School of Information Science and Engineering, Tokyo Institute of Technology; 2-12-1, Meguro, Tokyo: {ishizaki@cyb., imura@}mei.titech.ac.jp

[2]Department of Systems Innovation, Graduate School of Engineering Science, Osaka University; 1-3, Machikaneyama, Toyonaka, Osaka: kashima@sys.es.osaka-u.ac.jp

[3]Laboratory Jean Kuntzmann, University of Grenoble; B.P. 53, 38041 Grenoble, France: Antoine.Girard@imag.fr

[4]Affiliatioin with position: Chinese Academy of Sciences, The Universityof Tokyo: lnchen@sibs.ac.cn

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

the unit matrix of the size $n \times n$, e_i^n: the ith column vector of I_n, $\mathrm{span}(M)$: the range space spanned by the column vectors of a matrix M, $\|M\|_F$: the Frobenius norm of M, i.e., $\sqrt{\mathrm{tr}[MM^*]}$, $\|M\|$: the induced 2-norm of a matrix M, i.e., the maximum singular value of M, $\mathrm{diag}(v)$: the diagonal matrix whose diagonal entries are the entries of a vector v, $\mathrm{Diag}(M_1, \ldots, M_n)$, the block diagonal matrix whose diagonal blocks are composed of matrices $M_1, \ldots, M_n$.

For a set of natural numbers $\mathcal{I}$, the cardinality of $\mathcal{I}$ is denoted by $|\mathcal{I}|$. Furthermore, $e_{\mathcal{I}}^n \in \mathbb{R}^{n \times |\mathcal{I}|}$ denotes the matrix whose column vectors are composed of e_i^n for each $i \in \mathcal{I}$, i.e., $e_{\mathcal{I}}^n = [e_{i_1}^n, \ldots, e_{i_m}^n] \in \mathbb{R}^{n \times m}$ for $\mathcal{I} = \{i_1, \ldots, i_m\}$.

A matrix M (respectively, a transfer matrix G) is said to be *marginally stable* if its eigenvalues (poles) are all in the closed left-half plane, and all eigenvalues (poles) with zero real value are simple roots. In particular, M (respectively, G) is said to be *stable* if they are in the open left-half plane. Furthermore, M is said to be *reducible* if it can be placed into block upper-triangular form by simultaneous row and column permutations. Conversely, M is said to be *irreducible* if it is not reducible. Finally, the $\mathcal{H}_2$ and $\mathcal{H}_\infty$-norm of a stable transfer matrix G are denoted by $\|G(s)\|_{\mathcal{H}_2} := (\int_{-\infty}^{\infty} \|G(j\omega)\|_F^2 \frac{d\omega}{2\pi})^{1/2}$ and $\|G(s)\|_{\mathcal{H}_\infty} := \sup_{\omega \in \mathbb{R}} \|G(j\omega)\|$, respectively.

II. Problem Formulation

In this paper, we deal with the following positive linear systems evolving on directed networks:

Definition 1: Define

$$\mathcal{M}_n := \{M = \{m_{i,j}\} \in \mathbb{R}^{n \times n} : m_{i,j} \geq 0,\ \forall i \neq j,\ \text{irreducible, marginally stable}\}. \quad (1)$$

A linear system

$$\dot{x} = Ax + bu \quad (2)$$

is said to be *positive network* (A, b) if $A \in \mathcal{M}_n$ and $b \in \mathbb{R}_{0+}^n$.

In this paper, for simplicity of discussion, we only deal with single-input systems, even though the generalization to multi-input systems is available. Furthermore, the assumption of the irreducibility in (1) can be relaxed under a suitable situation (see Section III-D below for the details).

This class of systems includes, e.g., spatially-discrete reaction-diffusion systems, electric circuit systems and continuous time Markovian processes. The state trajectory of these systems does not escape from the non-negative orthant, i.e., $\mathbb{R}_{0+}^n$, under any non-negative input signals and initial conditions. Systems having such a non-negative property, called positive systems, often appear in ecology, industrial engineering and socio-economics [9], [10]. For $A = \{a_{i,j}\}$ and $b = \{b_i\}$, Fig. 1 depicts the interconnection topology of states (network structure) of positive networks.

In matrix theory, the largest real part eigenvalue of the Metzler matrices is called the Frobenius eigenvalue, and the associated eigenvector is called the Frobenius eigenvector, whose entries are all non-negative [9], [10]. In what follows, $\lambda_F(M)$ denotes the Frobenius eigenvalue of $M \in \mathcal{M}_n$. Furthermore, $v_l(M) \in \mathbb{R}_{0+}^{1 \times n}$ and $v_r(M) \in \mathbb{R}_{0+}^n$ denote the

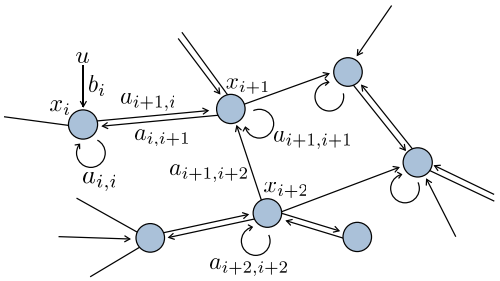

Fig. 1. Depiction of Positive Networks ($a_{i,j} \geq 0,\ i \neq j$).

left and right Frobenius eigenvectors such that $\|v_l(M)\| = \|v_r(M)\| = 1$. First, we state the following lemma that sums up the properties of $A \in \mathcal{M}_n$:

Lemma 1: Any $A \in \mathcal{M}_n$ is equipped with the following properties:

- the Frobenius eigenvalue $\lambda_F(A)$ is simple and real
- the entries of both left and right Frobenius eigenvectors are all positive, i.e., $v_l^{\mathsf{T}}(A), v_r(A) \in \mathbb{R}_+^n$
- A has no eigenvalue on the imaginary axis except the origin.

This lemma ensures that if $A \in \mathcal{M}_n$ has the zero-eigenvalue $\lambda_F(A) = 0$, its left and right eigenspaces are necessarily one-dimensional. Furthermore, define

$$\mathcal{M}_n^\dagger := \{M \in \mathcal{M}_n : v_r(M) = v_l^{\mathsf{T}}(M)\}.$$

The following *diagonal* transformation is to be useful for system analyses:

Lemma 2: Given $A \in \mathcal{M}_n$, define

$$D := \left[\mathrm{diag}^{-1}\left(v_r(A)\right) \mathrm{diag}(v_l^{\mathsf{T}}(A))\right]^{\frac{1}{2}}. \quad (3)$$

Then, $DAD^{-1} \in \mathcal{M}_n^\dagger$ holds.

This lemma guarantees that any $A \in \mathcal{M}_n$ is diagonally similar to $DAD^{-1} \in \mathcal{M}_n^\dagger$. Note that DAD^{-1} has a same network topology (Boolean structure) as that of A. The match of the left and right Frobenius eigenvectors of $A \in \mathcal{M}_n^\dagger$ works important roles in the following arguments. Based on this lemma, we, as necessary, assume $A \in \mathcal{M}_n^\dagger$ without loss of generality. Next, in order to formulate a clustering-based state aggregation problem, we give the following definition:

Definition 2: Let (A, b) be a positive network. The family of an index set $\{\mathcal{I}_{[l]}\}_{l \in \mathbb{L}}$ for $\mathbb{L} := \{1, \ldots, L\}$ is called a *cluster set* (its element is referred to as a cluster) if each element $\mathcal{I}_{[l]}$ is a disjoint subset of $\{1, \ldots, n\}$ and it satisfies $\bigcup_{l \in \mathbb{L}} \mathcal{I}_{[l]} = \{1, \ldots, n\}$. Furthermore, an *aggregation matrix* compatible with $\{\mathcal{I}_{[l]}\}_{l \in \mathbb{L}}$ is defined by

$$\mathsf{P} := \mathrm{Diag}(\mathsf{p}_{[1]}, \ldots, \mathsf{p}_{[L]})\Pi \in \mathbb{R}^{L \times n} \quad (4)$$

with $\mathsf{p}_{[l]} \in \mathbb{R}^{1 \times |\mathcal{I}_{[l]}|}$ such that $\|\mathsf{p}_{[l]}\| = 1$, and the permutation matrix $\Pi := [e_{\mathcal{I}_{[1]}}^n, \ldots, e_{\mathcal{I}_{[L]}}^n]^{\mathsf{T}} \in \mathbb{R}^{n \times n}$ for $e_{\mathcal{I}_{[l]}}^n \in \mathbb{R}^{n \times |\mathcal{I}_{[l]}|}$. Then, the *aggregated model* of (A, b) associated with P is given by $(\mathsf{P}A\mathsf{P}^{\mathsf{T}}, \mathsf{P}b)$.

The aggregation matrix P clearly satisfies $\mathsf{P}\mathsf{P}^{\mathsf{T}} = I_L$. This clustering-based state aggregation represents the aggregation

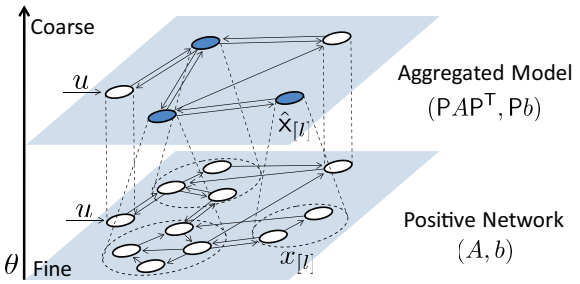

Fig. 2. Depiction of Clustering-based State Aggregation.

of the original state $x_{[l]} \in \mathbb{R}^{|\mathcal{I}_{[l]}|}$ into the aggregated states $\hat{\mathsf{x}}_{[l]} \in \mathbb{R}$ under the aggregation weight of $\mathsf{p}_{[l]} \in \mathbb{R}^{1\times|\mathcal{I}_{[l]}|}_+$; see Fig. 2. Let us consider recovering $x_{[l]} \in \mathbb{R}^{|\mathcal{I}_{[l]}|}$ from the aggregated states by $\mathsf{p}_{[l]}^\mathsf{T}\hat{\mathsf{x}}_{[l]} = (\mathsf{p}_{[l]}^\mathsf{T}\mathsf{p}_{[l]})x_{[l]} \in \mathbb{R}^{|\mathcal{I}_{[l]}|}$. According to this embedding, the transfer functions of the positive network and the aggregated model are defined by

$$\begin{aligned} g(s) &= (sI_n - A)^{-1}b \\ \hat{\mathsf{g}}(s) &= \mathsf{P}^\mathsf{T}(sI_L - \mathsf{P}A\mathsf{P}^\mathsf{T})^{-1}\mathsf{P}b, \end{aligned} \tag{5}$$

respectively. In this notation, the state aggregation problem to be considered is formulated as follows:

Problem 1: Let (A, b) be a positive network. Given a constant $\epsilon \in \mathbb{R}_+$, find an aggregation matrix P in (4) such that the aggregated model $(\mathsf{P}A\mathsf{P}^\mathsf{T}, \mathsf{P}b)$ is a positive network and satisfies $\|g(s) - \hat{\mathsf{g}}(s)\|_{\mathcal{H}_2} \le \epsilon$.

III. Main Results

A. Exact Aggregation

In what follows, we only consider the case of $\lambda_F(A) = 0$ because the results derived here is almost straightforwardly applied to stable positive networks. In this situation, we are required to take care of the stability of the error system $g - \hat{\mathsf{g}}$ in Problem 1. To ensure the stability of the error system, we provide the following form of the aggregation weight:

Lemma 3: Given a positive network (A, b), assume $A \in \mathcal{M}_n^\dagger$. For any cluster set $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$, if

$$\mathsf{p}_{[l]} := \frac{v_l(A)e^n_{\mathcal{I}_{[l]}}}{\|v_l(A)e^n_{\mathcal{I}_{[l]}}\|} \in \mathbb{R}^{1\times|\mathcal{I}_{[l]}|}_+, \tag{6}$$

then the error system $g - \hat{\mathsf{g}}$ associated with P in (4) is stable.

Proof: Consider a matrix $\mathsf{q}_{[l]} \in \mathbb{R}^{(|\mathcal{I}_{[l]}|-1)\times|\mathcal{I}_{[l]}|}$ such that $[\mathsf{p}_{[l]}^\mathsf{T}, \mathsf{q}_{[l]}^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^{|\mathcal{I}_{[l]}|\times|\mathcal{I}_{[l]}|}$ is unitary. Define Q by replacing $\mathsf{p}_{[l]}$ in (4) with $\mathsf{q}_{[l]}$ for each $l \in \mathbb{L}$, where we allow empty $\mathsf{q}_{[l]}$ if $|\mathcal{I}_{[l]}| = 1$. Considering the coordinate transformation of the error system by the unitary matrix $[\mathsf{P}^\mathsf{T}, \mathsf{Q}^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^{n\times n}$, we have

$$g(s) - \hat{\mathsf{g}}(s) = \Xi(s)\mathsf{Q}^\mathsf{T}\mathsf{Q}g(s) \tag{7}$$

where $\Xi(s) := \mathsf{P}^\mathsf{T}(sI_L - \mathsf{P}A\mathsf{P}^\mathsf{T})^{-1}\mathsf{P}A + I_n$. To prove the claim, it suffices to show that Ξ and $\mathsf{Q}g$ are both stable, instead marginally stable.

First, note that the definition of $\mathsf{p}_{[l]}$ in (6) guarantees $\mathrm{span}(v_l^\mathsf{T}(A)) \subset \mathrm{span}(\mathsf{P}^\mathsf{T})$. This provides

$$v_l(A)\mathsf{P}^\mathsf{T}\mathsf{P} = v_l(A), \quad \mathsf{P}^\mathsf{T}\mathsf{P}v_r(A) = v_r(A), \tag{8}$$

where the property of $A \in \mathcal{M}_n^\dagger$ ensures the second equality. Using (8), we have $\mathsf{Q}v_r(A) = \mathsf{Q}\mathsf{P}^\mathsf{T}\mathsf{P}v_r(A) = 0$, which implies that the eigenspace of (A, Q) associated with $\lambda_F(A) = 0$ is unobservable. Thus, all the poles of $\mathsf{Q}g$ are in the open left-half plane.

Next, we prove the stability of Ξ. Since the positivity of $\mathsf{p}_{[l]} \in \mathbb{R}^{1\times|\mathcal{I}_{[l]}|}_+$ holds for all $l \in \mathbb{L}$, the off-diagonal entries of $\mathsf{P}A\mathsf{P}^\mathsf{T}$ are all non-negative. Furthermore, using (8), we have $\lambda_F(\mathsf{P}A\mathsf{P}^\mathsf{T}) = 0$ and $v_l(\mathsf{P}A\mathsf{P}^\mathsf{T}) = v_l(A)\mathsf{P}^\mathsf{T}$, which ensure the marginal stability of $\mathsf{P}A\mathsf{P}^\mathsf{T} \in \mathcal{M}_L$. In addition, (8) provides $v_l(\mathsf{P}A\mathsf{P}^\mathsf{T})\mathsf{P}A = v_l(A)A = 0$, which implies that the eigenspace of $(\mathsf{P}A\mathsf{P}^\mathsf{T}, \mathsf{P}A)$ associated with $\lambda_F(\mathsf{P}A\mathsf{P}^\mathsf{T}) = 0$ is uncontrollable. Hence, all the poles of Ξ are as well in the open left-half plane. ■

Lemma 3 provides a suitable aggregation weight $\mathsf{p}_{[l]}$ to guarantee the stability of the error system. This is based on the fact that the eigenspace of (A, b) associated with $\lambda_F(A) = 0$ is exactly preserved into that of $(\mathsf{P}A\mathsf{P}^\mathsf{T}, \mathsf{P}b)$. Next, let us consider a simple situation under which a cluster is exactly redundant in the following sense:

Definition 3: Let (A, b) be a positive network. A cluster $\mathcal{I}_{[l]}$ is said to be *redundant* if there exists a scaler rational function $\hat{\mathsf{g}}_{[l]}$ such that

$$(e^n_{\mathcal{I}_{[l]}})^\mathsf{T}g(s) = \mathsf{p}_{[l]}^\mathsf{T}\hat{\mathsf{g}}_{[l]}(s), \tag{9}$$

where $\mathsf{p}_{[l]}$ is defined in (6).

This redundancy represents the linear dependence of the transfer functions corresponding to $\mathcal{I}_{[l]}$, namely the uncontrollability of the cluster state $x_{[l]} = (e^n_{\mathcal{I}_{[l]}})^\mathsf{T}x$. In order to characterize (9) in an algebraic manner, we introduce a projected controllability gramian of (A, b) as shown in the following lemma:

Lemma 4: Given a positive network (A, b), let $W \in \mathbb{R}^{(n-1)\times n}$ such that $[v_l^\mathsf{T}(A), W^\mathsf{T}] \in \mathbb{R}^{n\times n}$ is unitary. Define $\Phi_W := \int_0^\infty e^{WAW^\mathsf{T}\tau}Wb(e^{WAW^\mathsf{T}\tau}Wb)^\mathsf{T}d\tau$. Then, the positive semi-definite matrix

$$\Phi := W^\mathsf{T}\Phi_W W \in \mathbb{R}^{n\times n} \tag{10}$$

is independent of the choice of W.

The positive semi-definite matrix Φ in (10) represents the controllability gramian of the stable state-space of (A, b). The positive semi-definite matrix Φ algebraically characterizes the redundancy of clusters as follows:

Theorem 1: Given a positive network (A, b), assume $A \in \mathcal{M}_n^\dagger$. Define $\Phi \in \mathbb{R}^{n\times n}$ in (10). A cluster $\mathcal{I}_{[l]}$ is redundant if and only if there exists a row vector $\hat{\phi}_{[l]} \in \mathbb{R}^{1\times n}$ such that

$$(e^n_{\mathcal{I}_{[l]}})^\mathsf{T}\Phi_{\frac{1}{2}} = \mathsf{p}_{[l]}^\mathsf{T}\hat{\phi}_{[l]} \tag{11}$$

where $\mathsf{p}_{[l]}$ is defined in (6) and $\Phi_{\frac{1}{2}}$ denotes a Cholesky factor such that $\Phi = \Phi_{\frac{1}{2}}\Phi_{\frac{1}{2}}^\mathsf{T}$. Furthermore, if all clusters are redundant, then the aggregated model $(\mathsf{P}A\mathsf{P}^\mathsf{T}, \mathsf{P}b)$ associated with P in (4) is a positive network and satisfies

$$g(s) = \hat{\mathsf{g}}(s). \tag{12}$$

Proof: [proof of (11) $\Leftrightarrow$ (9)] Note that (11) holds if and only if there exists $\mathsf{q}_{[l]} \in \mathbb{R}^{(|\mathcal{I}_{[l]}|-1)\times|\mathcal{I}_{[l]}|}$ such that $[\mathsf{p}_{[l]}^{\mathsf{T}}, \mathsf{q}_{[l]}^{\mathsf{T}}]^{\mathsf{T}}$ is unitary and $\mathsf{q}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}\Phi_{\frac{1}{2}} = 0$ holds. Furthermore, it is also equivalent to

$$\mathsf{q}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}\Phi_{\frac{1}{2}}\{\mathsf{q}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}\Phi_{\frac{1}{2}}\}^{\mathsf{T}} = 0.$$

Thus, we have

$$\begin{aligned}
&\int_0^\infty \mathsf{q}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}e^{A\tau}W^{\mathsf{T}}Wbb^{\mathsf{T}}W^{\mathsf{T}}We^{A^{\mathsf{T}}\tau}e^n_{\mathcal{I}_{[l]}}\mathsf{q}_{[l]}^{\mathsf{T}}d\tau = 0\\
&\Leftrightarrow\ \mathsf{q}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}e^{At}W^{\mathsf{T}}Wb = 0,\ \forall t \in [0,\infty)\\
&\Leftrightarrow\ \mathsf{q}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}(sI_n - A)^{-1}W^{\mathsf{T}}Wb = 0.
\end{aligned}$$

In what follows, let us prove

$$\mathsf{q}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}(sI_n - A)^{-1}W^{\mathsf{T}}Wb = \mathsf{q}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}g(s). \quad (13)$$

Note that (9) is equivalent to the existence of $\mathsf{q}_{[l]}$ such that (13) is equal to 0. Here, $v_l(A)W^{\mathsf{T}}Wb = 0$ follows from $[v_l^{\mathsf{T}}(A), W^{\mathsf{T}}]$ being unitary and

$$\mathsf{q}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}v_r(A) = \mathsf{q}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}\mathsf{P}^{\mathsf{T}}\mathsf{P}v_r(A) = 0 \quad (14)$$

follows from (8). These imply that the eigenspaces of $(A, W^{\mathsf{T}}Wb)$ and $(A, \mathsf{q}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}})$ associated with $\lambda_F(A) = 0$ are uncontrollable and unobservable, respectively. Hence, (13) follows from the elimination of the uncontrollable and unobservable state-space.

[proof of (12)] The redundancy of all clusters implies that $\mathsf{Q}g = 0$ in (7). Hence, (12) follows. ■

Theorem 1 implies that the redundancy of a cluster $\mathcal{I}_{[l]}$ is characterized by the linear dependence among the row vectors of $\Phi_{\frac{1}{2}}$. More specifically, the behavior of x_i is algebraically translated into the ith row vector of $\Phi_{\frac{1}{2}}$. Note, however, that the redundancy (9) or equivalently (11) is restrictive in general. This is because, it represents a kind of local uncontrollability such that the controllable subspace of $x_{[l]} = (e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}x \in \mathbb{R}^{|\mathcal{I}_{[l]}|}$ is one-dimensional.

B. Redundancy Relaxation

In what follows, aiming at significant order reduction, we relax (9) through its equivalent characterization of (11). To this end, we introduce the following relaxation of the redundancy:

Definition 4: Let (A, b) be a positive network. Define $\Phi \in \mathbb{R}^{n\times n}$ in (10). A cluster $\mathcal{I}_{[l]}$ is said to be *θ-redundant* if there exists a row vector $\hat{\phi}_{[l]} \in \mathbb{R}^{1\times n}$ such that

$$\left\|(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}\Phi_{\frac{1}{2}} - \mathsf{p}_{[l]}^{\mathsf{T}}\hat{\phi}_{[l]}\right\|_F \le |\mathcal{I}_{[l]}|^{\frac{1}{2}}\theta, \quad \theta \in \mathbb{R}_+ \quad (15)$$

where $\mathsf{p}_{[l]}$ is defined in (6) and $\Phi_{\frac{1}{2}}$ denotes a Cholesky factor such that $\Phi = \Phi_{\frac{1}{2}}\Phi_{\frac{1}{2}}^{\mathsf{T}}$.

In this definition, the constant θ represents the degree of the redundancy (the normalization $|\mathcal{I}_{[l]}|^{\frac{1}{2}}$ is introduced for a technical reason). Clearly, (15) includes (11), and is equivalent to (11) if $\theta = 0$. Next, we propose to construct a cluster set $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ such that all clusters are θ-redundant. This notion yields an aggregated model as shown in the following theorem:

Theorem 2: Given a positive network (A, b), assume $A \in \mathcal{M}_n^\dagger$. Define $\Phi \in \mathbb{R}^{n\times n}$ in (10). If all cluster $\mathcal{I}_{[l]}$ are θ-redundant, then the aggregated model $(\mathsf{P}A\mathsf{P}^{\mathsf{T}}, \mathsf{P}b)$ associated with P in (4) is a positive network and satisfies

$$\|g(s) - \hat{\mathsf{g}}(s)\|_{\mathcal{H}_2} \le \sqrt{\alpha}\,\|\Xi(s)\|_{\mathcal{H}_\infty}\theta \quad (16)$$

where $\alpha := \sum_{l=1}^L |\mathcal{I}_{[l]}|(|\mathcal{I}_{[l]}| - 1)$ and Ξ is defined in (7).

Proof: First, $\|g - \hat{\mathsf{g}}\|_{\mathcal{H}_2} \le \|\Xi\|_{\mathcal{H}_\infty}\|\mathsf{Q}^{\mathsf{T}}\mathsf{Q}g\|_{\mathcal{H}_2}$ follows from (7). Using (13) and the fact that Φ is the controllability gramian of $(A, W^{\mathsf{T}}Wb)$, we have

$$\begin{aligned}
\|\mathsf{Q}^{\mathsf{T}}\mathsf{Q}g(s)\|_{\mathcal{H}_2} &= \|\mathsf{Q}^{\mathsf{T}}\mathsf{Q}(sI_n - A)^{-1}W^{\mathsf{T}}Wb\|_{\mathcal{H}_2}\\
&= (\mathrm{tr}[\mathsf{Q}^{\mathsf{T}}\mathsf{Q}\Phi\mathsf{Q}^{\mathsf{T}}\mathsf{Q}])^{\frac{1}{2}} = \|\mathsf{Q}\Phi_{\frac{1}{2}}\|_F.
\end{aligned}$$

Define $\Delta_{[l]} := (e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}\Phi_{\frac{1}{2}} - \mathsf{p}_{[l]}^{\mathsf{T}}\hat{\phi}_{[l]}$. Since $[\mathsf{p}_{[l]}^{\mathsf{T}}, \mathsf{q}_{[l]}^{\mathsf{T}}]^{\mathsf{T}}$ is unitary, we have $\mathsf{q}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}\Phi_{\frac{1}{2}} = \mathsf{q}_{[l]}\Delta_{[l]}$, where we allow empty $\mathsf{q}_{[l]}$ if $|\mathcal{I}_{[l]}| = 1$. The θ-redundancy implies $\|\Delta_{[l]}\|_F \le |\mathcal{I}_{[l]}|^{\frac{1}{2}}\theta$. Thus, we have

$$\|\mathsf{Q}\Phi_{\frac{1}{2}}\|_F^2 \le \sum_{l=1}^L \|\mathsf{q}_{[l]}\|_F^2\|\Delta_{[l]}\|_F^2 \le \sum_{l=1}^L |\mathcal{I}_{[l]}|(|\mathcal{I}_{[l]}| - 1)\theta^2.$$

Hence, (16) follows. ■

This theorem shows a linear dependence between the degree of the redundancy θ in (15) and the aggregation error in terms of $\mathcal{H}_2$-norm. This implies that we can use θ as a design parameter to regulate the coarseness of resultant aggregated models.

C. Cluster Construction

We give an algorithm for cluster determination. When we construct a cluster, we first choose an index i_0, then find indices satisfying

$$\left\|\phi_i - v_i v_{i_0}^{-1}\phi_{i_0}\right\|_F \le \theta, \quad \forall i \in \mathcal{I}_{[l]} \quad (17)$$

where $\phi_i \in \mathbb{R}^{1\times n}$ denotes the ith row vector of $\Phi_{\frac{1}{2}}$ and $v_i \in \mathbb{R}$ denotes the ith entry of $v_l(A)$. The condition (17) is a sufficient condition of (15), where we take $\hat{\phi}_{[l]} = \|v_l(A)e^n_{\mathcal{I}_{[l]}}\|v_{i_0}^{-1}\phi_{i_0}$. Based on (17), an algorithm to make clusters is constructed as follows: Suppose that, at the lth step, θ-redundant clusters $\{\mathcal{I}_{[1]}, \ldots, \mathcal{I}_{[l]}\} \neq \{1, \ldots, n\}$ are constructed. At the $(l+1)$th step, a subsequent cluster $\mathcal{I}_{[l+1]}$ is constructed by the procedure of

(i) choose $i_0 \in \{1, \ldots, n\}\backslash\{\mathcal{I}_{[1]}, \ldots, \mathcal{I}_{[l]}\}$, not belonging to any cluster, and set $\mathcal{I}_{[l+1]} = \{i_0\}$
(ii) for each $i \in \{1, \ldots, n\}\backslash\{\mathcal{I}_{[1]}, \ldots, \mathcal{I}_{[l]}\}$, calculate the norm in (17)
(iii) if the norm is less than θ, update $\mathcal{I}_{[l+1]} \leftarrow \{\mathcal{I}_{[l+1]}, i\}$.

Repeating this procedure yields a cluster set $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ satisfying the θ-redundancy in (15). In this construction algorithm, degrees of freedom for the choice of i_0 in step (i) and the choice of $\hat{\phi}_{[l]}$ remain.

D. Generalization to a Class of Reducible Systems

In this subsection, for marginally stable positive networks, we consider relaxing the irreducibility assumed in Definition 1. Note that for reducible A, we cannot apply the diagonal transformation in Lemma 2 because the left and right Frobenius eigenvectors of A have possibly zero entries. This spoils the property in (8), fundamentally used in Section III-A. Conversely, if once (8) is guaranteed, the assumption of $A \in \mathcal{M}_n^\dagger$ is not required in Theorem 2.

Without loss of generality, let us assume that reducible A is in the form of

$$A = \begin{bmatrix} A_{1,1} & & \\ \vdots & \ddots & \\ A_{K,1} & \cdots & A_{K,K} \end{bmatrix} \in \mathbb{R}^{n\times n}, \quad n = \sum_{k=1}^{K} n_k \quad (18)$$

where the diagonal blocks $A_{k,k} \in \mathcal{M}_{n_k}$ are irreducible for all $k \in \{1,\dots,K\}$. In addition, we assume that

- $A_{K,K} \in \mathcal{M}_{n_K}^\dagger$, without loss of generality
- $\lambda_F(A) = 0$ is simple
- the entries of the left eigenvector associated with $\lambda_F(A)$ are all positive, i.e., $v_l(A) \in \mathbb{R}_+^{1\times n}$.

Though the second and third items are not necessarily trivial for reducible A, chemical reaction systems described by the chemical master equation, in fact, satisfy these assumptions (see Section IV for details). Under these assumptions, it is readily verified that $\lambda_F(A) = \lambda_F(A_{K,K}) = 0$. In addition, $v_l(A)$ and $v_r(A)$ are in the form of $v_l(A) = [v_+, v_l(A_{K,K})]$ and $v_r(A) = [0, v_l(A_{K,K})]^\mathsf{T}$ where $v_+ \in \mathbb{R}_+^{1\times(n-n_K)}$ denotes some positive row vector.

We construct θ-redundant clusters imposing a particular constraint so as to guarantee (8). Denote the index set corresponding to the kth block by $\mathcal{N}_k$, i.e., $A_{k,k} = (e^n_{\mathcal{N}_k})^\mathsf{T} A e^n_{\mathcal{N}_k}$ holds for each $k \in \{1,\dots,K\}$. Let us construct clusters separately within $\mathcal{N}_{1:K-1} := \{\mathcal{N}_1,\dots,\mathcal{N}_{K-1}\}$ and $\mathcal{N}_K$. More specifically, we make $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ such that

$$\bigcup_{l=1}^{l_0} \mathcal{I}_{[l]} = \mathcal{N}_{1:K-1}, \quad \bigcup_{l=l_0+1}^{L} \mathcal{I}_{[l]} = \mathcal{N}_K \quad (19)$$

for some $l_0 \in \mathbb{L}$. Such a cluster set assures that the aggregation matrix is in the form of $\mathsf{P} = \mathrm{Diag}(\mathsf{P}_{1:K-1}, \mathsf{P}_K)$ where $\mathsf{P}_{1:K-1} \in \mathbb{R}^{l_0\times\sum_{k=1}^{K-1} n_k}$ and $\mathsf{P}_K \in \mathbb{R}^{(L-l_0)\times n_K}$. Since $v_l(A_{K,K})\mathsf{P}_K^\mathsf{T}\mathsf{P}_K = v_l(A_{K,K})$ holds, this ensures (8). Thus, P associated with a θ-redundant cluster set $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ satisfying (19) provides an aggregated model satisfying (16).

IV. Application to Reduction of Chemical Master Equation

A. Chemical Master Equation

In this section, we deal with the following chemical reaction system, called the Michaelis-Menten system composed of the 4 kinds of molecules $S_1,\dots,S_4$:

$$\begin{aligned} S_1 + S_2 &\underset{c_2}{\overset{c_1}{\rightleftarrows}} S_3 \\ S_3 &\overset{c_3}{\to} S_4 + S_2, \end{aligned} \quad (20)$$

where $c_1,\dots,c_3 \in \mathbb{R}_{0+}$ denote the reaction rate constants. Let us consider the situation where N_0 molecules of each S_1 and S_2 are present at the initial time. In this situation, all realizable distributions of the numbers of the molecules are enumerated by the $n = n(N_0) := (N_0^2 + 3N_0 + 2)/2$ kinds of vectors, i.e.

$$\begin{bmatrix} N_0 \\ N_0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} N_0-1 \\ N_0-1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} N_0-2 \\ N_0-2 \\ 2 \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ N_0 \\ 0 \\ N_0 \end{bmatrix} \in \mathbb{N}^4,$$

where the kth entry denotes the number of S_k for each $k \in \{1,\dots,4\}$.

Hereafter, denote each realizable distribution by $\xi_i \in \mathbb{N}^4$ for $i \in \{1,\dots,n\}$, and the realization probability of ξ_i at the time t by $x_i(t) \in [0,1]$. Under this denotation, the time evolution of (20) is represented by the chemical master equation (CME)

$$\dot{x} = Ax, \quad x(0) = [1,0,\dots,0]^\mathsf{T} \quad (21)$$

where $x := [x_1,\dots x_n]^\mathsf{T} \in [0,1]^n$ and $A \in \mathbb{R}^{n\times n}$ is marginally stable and Metzler (see [11] for details). Note that the initial condition $x(0)$ represents that the realization probability $x_1(t)$ of $\xi_1 := [N_0, N_0, 0, 0]^\mathsf{T}$ is 1 at $t = 0$. Furthermore, the sum of all states constantly remains to be 1 for all $t \in [0,\infty)$ because each x_i represents the realization probability of the distribution ξ_i. This implies that A in (21) satisfies $\lambda_F(A) = 0$ and $v_l(A) = [1,\dots,1]$. Moreover, due to the irreversibility of the the second reaction in (20), A is reducible and satisfies $v_r(A) = \lim_{t\to\infty} x(t) = [0,\dots,0,1]^\mathsf{T}$. Fig. 3 depicts the transition diagram of the distributions of (20) in the CME expression, where the pair of $(x_i;\xi_i)$, i.e., the realization probability and the realizable distribution, is assigned at each node i. In this figure, the horizontal and vertical transitions correspond to the first and the second reaction in (20), respectively.

Remark 1: In general, the transition diagrams of chemical reactions composed of κ kinds of molecules are described by a subgraph of the multidimensional lattice $\mathbb{N}^\kappa$. Furthermore, A is marginally stable and Metzler, and is uniquely determined from chemical reactions to be considered. In addition, it satisfies that $\lambda_F(A) = 0$ is simple, $v_l(A) = [1,\dots,1]$ and $v_r(A) = \lim_{t\to\infty} x(t)$. It should be emphasized that since the system dimension tends to be considerably large ($O(N_0^\kappa)$), the analytical/numerical treatments of CMEs are not necessarily easy [11]. Thus, toward practical analyses, we are required to construct a *mesoscopic model* that describes essential behavior of the microscopic expression in (21).

B. $\mathcal{H}_2$-aggregation of Chemical Master Equation

In what follows, we show the efficiency of the proposed method through the reduction of the CME in (21). For (20), let the reaction rate constants be $c_1 = 1$, $c_2 = 0.1$ and $c_3 = 3$. Furthermore, let the initial number of molecules be $N_0 = 140$, which yields the $n = 10011$th dimensional CME.

From Fig. 4, which shows the resultant order of aggregated models versus the value of θ, we can see that the order of

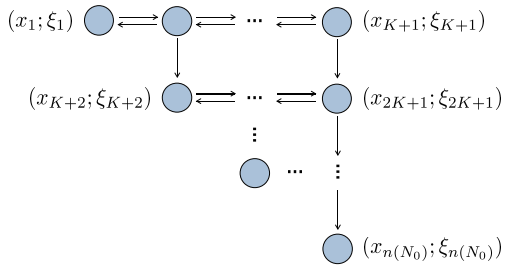

Fig. 3. State Transition Diagram of Chemical Master Equation ($\xi_1 = [N_0, N_0, 0, 0]^\mathsf{T}, \ldots, \xi_{n(N_0)} = [0, N_0, 0, N_0]^\mathsf{T}$).

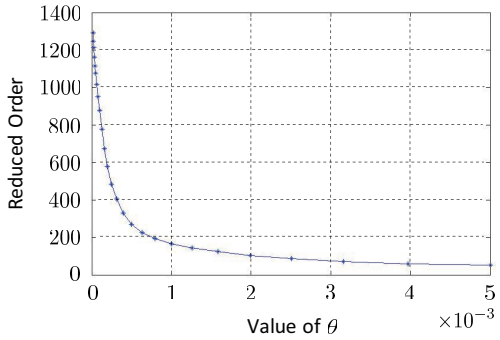

Fig. 4. Order of Resulting Models versus Values of θ.

aggregated models decreases as θ increases. This means that the θ-redundancy appropriately captures the coarseness of aggregated models. In particular, when we take $\theta = 5\times10^{-5}$, the order of the aggregated model is $L = 1077$ and the resultant relative error is $\|g - \mathrm{g}\|_{\mathcal{H}_2}/\|g\|_{\mathcal{H}_2} = 0.0239$.

Furthermore, Fig. 5 shows the state trajectory of the original system $x \in \mathbb{R}^n$ (the solid line, $n = 10011$th order) and that of the aggregated model $\mathsf{P}^\mathsf{T}\hat{\mathrm{x}} \in \mathbb{R}^n$ (the broken line, $L = 1077$th order). We can see from this figure that the behavior of the CME in (21) is well approximated by that of the aggregated model. In addition, the aggregated model not only approximates the behavior of (21), but also preserves specific properties of the CME, namely it is equipped with the following properties:

- the positivity of the system (non-negativity of the state trajectory) is preserved
- the sum of all states constantly remains to be 1
- the steady-state distribution is exactly preserved
- each state of the aggregated model represents averaged states of the original CME

where the second to fourth items are assured by (8).

V. Conclusion

In this paper, we have proposed a model reduction method for positive systems evolving on directed networks, called positive networks. In this method, we construct a set of clusters based on a kind of local controllability of the state-space. Aggregating the cluster set under suitable weighting, we obtain a reduced model that preserves interconnection topology among the clusters as well as some specific properties, such as stability, steady-state characteristic and system positivity. Furthermore, we have derived an $\mathcal{H}_2$-error bound of the state discrepancy caused by the aggregation. In addition, applying the proposed method to chemical reaction

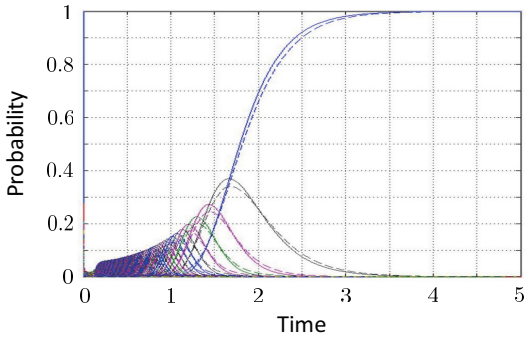

Fig. 5. State Trajectory of Original System (the solid line, 10011th order) and Aggregated Model (the broken line, 1077th order).

systems described by the chemical master equation, we have obtained a mesoscopic model that well approximates the original system as well as preserves specific properties as the chemical master equation.

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 Program for World-Leading Innovative R&D on Science and Technology (FIRST Program)," initiated by the Council for Science and Technology Policy (CSTP). This research is also supported by Grant-in-Aid for JSPS Fellows. The authors would like to thank D. Cheong from Tokyo Institute of Technology for fruitful discussion.

References

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

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

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

[4] 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.

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

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

[7] 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.

[8] T. Ishizaki, K. Kashima, J. Imura, and K. Aihara, "Reaction-diffusion clustering of single-input dynamical networks," in *50th IEEE Conference on Decision and Control, and European Control Conference*, 2011, pp. 7837–7842.

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

[10] L. Benvenuti and L. Farina, "A tutorial on the positive realization problem," *IEEE Transactions on Automatic Control*, vol. 49-5, pp. 651–664, 2004.

[11] B. Munsky and M. Khammash, "The finite state projection approach for the analysis of stochastic noise in gene networks," *IEEE Transactions on Automatic Control*, vol. 53, Special Issue, pp. 201–214, 2008.