Clustered Model Reduction of Positive Directed Networks ⋆

Takayuki Ishizaki [a], Kenji Kashima [b], Antoine Girard [c], Jun-ichi Imura [a], Luonan Chen [d], and Kazuyuki Aihara [e]

[a] *Graduate School of Information Science and Engineering, Tokyo Institute of Technology; 2-12-1, Meguro ward, Tokyo*

[b] *Graduate School of Informatics, Kyoto University; Yoshida Honmachi, Sakyo-ku, Kyoto, 606-8501, Japan*

[c] *Laboratory Jean Kuntzmann, University of Grenoble; B.P. 53, 38041 Grenoble, France*

[d] *Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences; 500 Cao Bao Road, Shanghai 200233*

[e] *Institute of Industrial Science, University of Tokyo; 4-6-1 Komaba, Meguro ward, Tokyo*

Abstract

This paper proposes a clustered model reduction method for semistable positive linear systems evolving over directed networks. In this method, we construct a set of clusters, i.e., disjoint sets of state variables, based on a notion of cluster reducibility, defined as the uncontrollability of local states. By aggregating the reducible clusters with aggregation coefficients associated with the Frobenius eigenvector, we obtain an approximate model that preserves not only a network structure among clusters, but also several fundamental properties, such as semistability, positivity, and steady state characteristics. Furthermore, it is found that the cluster reducibility can be characterized for semistable systems based on a projected controllability Gramian that leads to an a priori $\mathcal{H}_2$-error bound of the state discrepancy caused by aggregation. The efficiency of the proposed method is demonstrated through an illustrative example of enzyme-catalyzed reaction systems described by a chemical master equation. This captures the time evolution of chemical reaction systems in terms of a set of ordinary differential equations.

Key words: Positive linear systems; Model reduction; Network clustering; Chemical master equations.

1 Introduction

Many dynamical systems of interest to control community are inherently constructed from subsystem interconnections. Examples of such interconnected systems include power grids, transportation networks and so forth; see [3] for an overview. Since the network structure of such systems is often complex and large-scale, it is crucial to develop an approximation method that enables us to reduce their complexity (dimension). In addition, it is more desirable to preserve some particular properties of these systems, such as a network structure, stability, and positivity, throughout the approximation. This kind of structure-preserving model reduction has the potential to significantly simplify to analyse large-scale systems while capturing their essential properties of interest.

A number of model reduction methods can be found in the literature [1]. For instance, model reduction methods inspired by principal component analysis, such as the balanced truncation [4] and the Hankel norm approximation [15], are well known. A major advantage of these methods is the availability of an error bound in terms of the $\mathcal{H}_\infty$-norm or Hankel norm. Furthermore, the class of moment matching methods, including the Krylov subspace methods, is also well known [8]. This class of methods aims to suppress discrepancies in the system behavior for specific input signals, and has the advantage of a computationally efficient implementation. However, unlike the former class of methods, a priori error bounds have not yet been derived. For these existing model reduction, a systematic procedure is provided. However, they have a drawback in terms of their application to network system: the network structure of systems, i.e., the interconnection topology among state variables or subsystems, is destroyed through the approximation. This is because each state of the resultant approximants is constructed by a linear combination of all the original states. Therefore, to practically approximate a network system,

⋆ This paper was not presented at any IFAC meeting. This research is supported in part by the Aihara Project, the FIRST program from JSPS, initiated by CSTP, and by Japan Science and Technology Agency, CREST. Corresponding author T. Ishizaki Tel. & Fax +81-3-5734-2646.

Email addresses: ishizaki@mei.titech.ac.jp (Takayuki Ishizaki), kashima@amp.i.kyoto-u.ac.jp (Kenji Kashima), Antoine.Girard@imag.fr (Antoine Girard), imura@mei.titech.ac.jp (Jun-ichi Imura), lnchen@sibs.ac.cn (Luonan Chen), aihara@sat.t.u-tokyo.ac.jp (Kazuyuki Aihara).

it is crucial to develop a model reduction method that explicitly preserves the network structure of the system.

One approach to network structure-preserving model reduction can be an extension of structure-preserving model reduction methods found in the literature. For example, [18] addresses a model reduction problem that considers the preservation of the second-order structure. However, this problem is not formulated on the premise of network structure preservation. In contrast to the existing approach, a clustered model reduction method has been developed for stable systems evolving over undirected, or bidirectional networks [12]. In this method, by focusing on the symmetry of system matrices, we have introduced a system transformation, called positive tridiagonalization, to characterize the cluster reducibility, defined as the uncontrollability of disjoint subsets, or clusters, of state variables. The aggregation of reducible clusters yields an approximate model that preserves the network structure among clusters and the stability of systems, and provides an error bound in terms of the $\mathcal{H}_\infty$-norm. However, the applicability of this clustered model reduction is rather restricted because both stability preservation and reducibility characterization are heavily reliant on the symmetry of the system matrices. From a practical point of view, it is crucial to improve the applicability of our clustered model reduction framework.

One major difficulty confronted by network structure-preserving model reduction involves preserving the stability of the original system in its approximants. To enable the systematic development of clustered model reduction, it is important to clarify the class of systems to which it can reasonably be applied. In [5], it has been found that stability analyses can be tractably performed for a class of systems admitting a positive property, called (internally) positive systems. More specifically, the stability of positive systems can be characterized by an eigenpair, called the Frobenius eigenvalue and eigenvector. In fact, clustered model reduction has good compatibility with the approximation of positive systems because, as long as we make the aggregation coefficients non-negative, the positivity property of systems can be preserved in its approximants. In this paper, we use this compatibility to show that the semistability of positive systems can be preserved by a selection of aggregation coefficients specified by the Frobenius eigenvector. Moreover, we derive an alternative characterization of cluster reducibility based on a projected controllability Gramian. Owing to this development, we can apply clustered model reduction to semistable positive systems, called positive directed networks, involving compartmental systems, and Markovian processes [5].

To demonstrate the improved applicability, we provide an illustrative example of a chemical master equation (CME) compatible with enzyme-catalyzed reaction systems. It is known that CMEs belong to a class of Markovian processes [9,16], which can be regarded as a semistable positive directed network. Since the dimension of CMEs tends to be large, they are not necessarily analytically or numerically tractable. To overcome this difficulty, the proposed clustered model reduction method produces an aggregated model that preserves several fundamental properties as Markovian processes. A preliminary version of this paper was published in [11]. In comparison with it, this paper provides detailed proofs and explanations for our theoretical results.

The remainder of this paper is structured as follows: In Section 2, we first formulate a clustered model reduction problem for positive directed networks. In Section 3, we characterize the cluster reducibility using a projected controllability Gramian, and develop a clustered model reduction method. Section 4 demonstrates the efficiency of the proposed method through an illustrative example of CMEs. Finally, concluding remarks are provided in Section 5.

Notation $\mathbb{R}$: the set of real numbers, $\mathbb{R}_{>0}$ ($\mathbb{R}_{\geq 0}$) : the set of positive (non-negative) real numbers, $\mathbb{N}$: the set of non-negative integers, I_n: the n-dimensional identity matrix, $|\mathcal{I}|$: the cardinality of a set $\mathcal{I}$, $\mathrm{im}(M)$: the image of a matrix M, $\|M\|_{\mathrm{F}}$: the Frobenius norm of a matrix M, $\mathrm{diag}(v)$: the diagonal matrix having a vector v on its diagonal, $\mathrm{Diag}(M_1, \ldots, M_n)$: the block diagonal matrix having matrices $M_1, \ldots, M_n$ on its block diagonal.

For $\mathcal{I} \subseteq \{1, \ldots, n\}$, let $e^n_{\mathcal{I}} \in \mathbb{R}^{n \times |\mathcal{I}|}$ denote the matrix composed of the column vectors of I_n compatible with $\mathcal{I}$. A square matrix M (respectively, a transfer matrix G) is said to be *semistable* if all eigenvalues of M (poles of G) are in the closed left-half plane, and all eigenvalues (poles) with zero real value are simple roots. A square matrix 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. Furthermore, M is said to be *Metzler* if the off-diagonal entries of M are all non-negative. The positive (negative) semidefiniteness of $M = M^{\mathsf{T}} \in \mathbb{R}^{n \times n}$ is denoted by $M \succeq O_n$ ($M \preceq O_n$). Its positive (negative) definiteness is denoted similarly. The $\mathcal{H}_\infty$-norm of a stable proper transfer matrix G and the $\mathcal{H}_2$-norm of a stable strictly proper transfer matrix G are denoted by $\|G\|_{\mathcal{H}_\infty}$ and $\|G\|_{\mathcal{H}_2}$.

2 Problem Formulation

2.1 Preliminaries

In this paper, we deal with a class of positive linear systems evolving over directed networks. We denote a set of irreducible Metzler matrices by

$$\mathbb{M}_n := \{M \in \mathbb{R}^{n \times n} : \text{irreducible, Metzler}\}. \quad (1)$$

In this notation, we define the following class of positive systems:

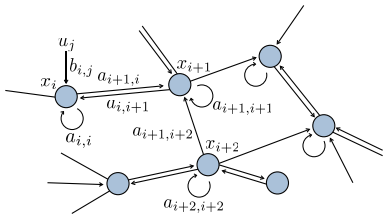

Fig. 1. Depiction of positive directed networks.

Definition 1 A linear system

$$\Sigma : \dot{x} = Ax + Bu \tag{2}$$

is said to be a *positive directed network* if $A \in \mathbb{M}_n$ and $B \in \mathbb{R}^{n\times m}_{\geq 0}$.

This class of systems includes spatially-discrete reaction-diffusion systems, electrical circuit networks, continuous-time Markovian processes, and so forth. Their state trajectory does not escape from the non-negative orthant $\mathbb{R}^n_{\geq 0}$ under non-negative input signals and initial conditions. Such systems having the non-negative property often appear in science and engineering [5]. With the notation of $A = \{a_{i,j}\}$ and $B = \{b_{i,j}\}$, Fig. 1 depicts the interconnection topology (network structure) of positive directed networks. Note that the irreducibility of $A \in \mathbb{M}_n$ assumed in (1) coincides with the strong connectivity of networks, which can be relaxed under a suitable situation; see Section 3.4 for details.

In matrix theory, the dominant eigenvalue of a Metzler matrix $A \in \mathbb{M}_n$ is called the Frobenius eigenvalue of A. An eigenvalue λ of A is said to be *dominant* if it satisfies $\mathrm{Re}\,\lambda = \max\{\mathrm{Re}\,\tilde{\lambda} : \tilde{\lambda} \in \Lambda(A)\}$, where $\mathrm{Re}\,\lambda$ denotes the real part of λ, and $\Lambda(A)$ denotes the spectrum of A. In this paper, its Frobenius eigenvalue is denoted by $\lambda_{\mathrm{F}}(A)$, and the associated left and right eigenvectors, called the left and right Frobenius eigenvectors, are denoted by $v_{\mathrm{L}}(A) \in \mathbb{R}^{1\times n}$ and $v_{\mathrm{R}}(A) \in \mathbb{R}^n$, respectively. It is known that $\lambda_{\mathrm{F}}(A)$ is real and unique, and all entries of $v_{\mathrm{L}}(A)$ and $v_{\mathrm{R}}(A)$ are non-negative; see Theorem 11 in [5]. Without loss of generality, we can assume that the left and right Frobenius eigenvectors have the unit norm, namely $\|v_{\mathrm{L}}(A)\| = \|v_{\mathrm{R}}(A)\| = 1$. In this notation, several properties of a semistable Metzler matrix $A \in \mathbb{M}_n$ are shown as follows:

Lemma 1 Let $A \in \mathbb{M}_n$ be given, and assume that it is semistable. Then, it follows that

(i) $\lambda_{\mathrm{F}}(A)$ has an algebraic multiplicity equal to 1,
(ii) all entries of both $v_{\mathrm{L}}(A)$ and $v_{\mathrm{R}}(A)$, which are uniquely determined, are positive, and
(iii) no eigenvalue of A is on the imaginary axis except for the origin.

PROOF. Since $A \in \mathbb{M}_n$ is irreducible, the first and second claims are assured by Theorem 17 in [5]. Let us prove the third claim by contradiction. Suppose that the claim does not hold, i.e., A has eigenvalues on the imaginary axis away from the origin. Since $\lambda_{\mathrm{F}}(A)$ is the unique largest real eigenvalue of A, the supposition requires that $\lambda_{\mathrm{F}}(A)$ is on the positive real axis. However, this requirement contradicts the assumption of semistability. Hence, the third claim is verified. □

Lemma 1 ensures that, if $\lambda_{\mathrm{F}}(A) = 0$, the corresponding left and right eigenspaces are necessarily one-dimensional owing to the uniqueness, and they are spanned by vectors having positive entries. Based on this fact, let us define a special class of matrices in $\mathbb{M}_n$ as

$$\mathbb{M}^{\dagger}_n := \{M \in \mathbb{M}_n : v_{\mathrm{R}}(M) = v^{\mathsf{T}}_{\mathrm{L}}(M)\},$$

which means that $A \in \mathbb{M}^{\dagger}_n$ has the same left and right Frobenius eigenvectors up to their transpose. As shown in the following lemma, any $A \in \mathbb{M}_n$ is diagonally similar to a matrix in $\mathbb{M}^{\dagger}_n$:

Lemma 2 For any $A \in \mathbb{M}_n$, $D^{-\frac{1}{2}}AD^{\frac{1}{2}} \in \mathbb{M}^{\dagger}_n$ with

$$D := \left\{\mathrm{diag}\left(v^{\mathsf{T}}_{\mathrm{L}}(A)\right)\right\}^{-1} \mathrm{diag}(v_{\mathrm{R}}(A)). \tag{3}$$

PROOF. From the second property shown in Lemma 1, we see that all diagonal entries of D are positive. Hence, all off-diagonal entries of $D^{-\frac{1}{2}}AD^{\frac{1}{2}}$ are non-negative, i.e., $D^{-\frac{1}{2}}AD^{\frac{1}{2}}$ is Metzler. From direct calculation, we can verify that both left and right Frobenius eigenvectors of $D^{-\frac{1}{2}}AD^{\frac{1}{2}}$ coincide with $v_{\mathrm{F}} := \{v_{\mathrm{R}}(A) * v^{\mathsf{T}}_{\mathrm{L}}(A)\}^{1/2} \in \mathbb{R}^n_{>0}$ up to its transpose, where both the square root $\{\cdot\}^{1/2}$ and multiplication $*$ are performed in an entry-wise manner. Hence, the claim follows. □

Lemma 2 shows that any $A \in \mathbb{M}_n$ is diagonally similar to $D^{-\frac{1}{2}}AD^{\frac{1}{2}} \in \mathbb{M}^{\dagger}_n$. Owing to D being diagonal, the transformed system matrix $D^{-\frac{1}{2}}AD^{\frac{1}{2}}$ has the same network structure (i.e., Boolean structure) as that of the original system matrix A. Thus, without loss of generality, we can assume that $A \in \mathbb{M}^{\dagger}_n$ holds for any positive directed network Σ. In the rest of this paper, let $v_{\mathrm{F}} \in \mathbb{R}^n_{>0}$ denote the Frobenius eigenvector of $A \in \mathbb{M}^{\dagger}_n$. One major benefit of this coincidence of the left and right Frobenius eigenvectors is that the semistability of $A \in \mathbb{M}^{\dagger}_n$ can be proved by the quadratic Lyapunov function associated with the identity matrix I_n, namely

$$A + A^{\mathsf{T}} \preceq \mathcal{O}_n, \quad A \in \mathbb{M}^{\dagger}_n. \tag{4}$$

This is proved by the fact that, for any semistable $A \in \mathbb{M}^{\dagger}_n$, v_{F} and $v^{\mathsf{T}}_{\mathrm{F}}$ are the right and left Frobenius eigenvectors of $A + A^{\mathsf{T}} \in \mathbb{M}^{\dagger}_n$ associated with $\lambda_{\mathrm{F}}(A + A^{\mathsf{T}}) = 2\lambda_{\mathrm{F}}(A) \leq 0$. The negative semidefiniteness of A shown in (4) has good compatibility with the semistability-preserving model reduction based on orthogonal projection, as shown in the following lemma:

Lemma 3 Let $P \in \mathbb{R}^{L\times n}$ be such that $PP^{\mathsf{T}} = I_L$ with $L \le n$. If $A \in \mathbb{M}_n^\dagger$ is semistable, then PAP^{T} is also semistable. In particular, if $A \in \mathbb{M}_n^\dagger$ is stable, then PAP^{T} is stable.

PROOF. Since any principal submatrix of a negative semidefinite matrix is also negative semidefinite, it follows from (4) that $P(A + A^{\mathsf{T}})P^{\mathsf{T}} \preceq \mathcal{O}_L$. Hence, the semistability of PAP^{T} is proved. The same argument for a stable $A \in \mathbb{M}_n^\dagger$ ensures the stability of PAP^{T}. □

Lemma 3 guarantees that the semistability of $A \in \mathbb{M}_n^\dagger$ is preserved in the projected matrix PAP^{T}. Note that PAP^{T} is not necessarily semistable for a general semistable matrix A. Furthermore, the condition $PP^{\mathsf{T}} = I_L$ indicates that all column vectors of P^{T} are orthonormal. Thus, we see that the assumption of $A \in \mathbb{M}_n^\dagger$ has good compatibility with model reduction based on orthogonal projection preserving the semistability of systems. However, note that PAP^{T} is not necessarily Metzler even if $A \in \mathbb{M}_n^\dagger$. Moreover, if P is a dense matrix, then so is PAP^{T}. This implies that the network structure of the original Σ is destroyed through the orthogonal projection. Therefore, to preserve the positivity as well as the network structure of Σ, we need to impose a specific structure on $P \in \mathbb{R}^{L\times n}$.

2.2 Clustered Model Reduction Problem for Positive Directed Networks

In this subsection, we formulate the problem of clustered model reduction for positive directed networks. To this end, we first introduce the following notion of network clustering [12]:

Definition 2 Let $\mathbb{L} := \{1, \dots, L\}$. The family of an index set $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ is called a *cluster set*, each of whose elements is referred to as a cluster, if each element $\mathcal{I}_{[l]}$ is a disjoint subset of $\{1, \dots, n\}$ and satisfies $\bigcup_{l\in\mathbb{L}} \mathcal{I}_{[l]} = \{1, \dots, n\}$. Furthermore, an *aggregation matrix* compatible with $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ is defined by

$$P := \mathrm{Diag}(p_{[1]}, \dots, p_{[L]})\Pi \in \mathbb{R}^{L\times n} \tag{5}$$

where $p_{[l]} \in \mathbb{R}^{1\times|\mathcal{I}_{[l]}|}$ such that $\|p_{[l]}\| = 1$, and the permutation matrix is defined as $\Pi := [e^n_{\mathcal{I}_{[1]}}, \dots, e^n_{\mathcal{I}_{[L]}}]^{\mathsf{T}}$ for $e^n_{\mathcal{I}_{[l]}} \in \mathbb{R}^{n\times|\mathcal{I}_{[l]}|}$.

Based on Definition 2, we define the aggregated model of Σ in (2) by

$$\hat{\Sigma}_P : \begin{cases} \dot{\xi} = PAP^{\mathsf{T}}\xi + PBu \\ \hat{x} = P^{\mathsf{T}}\xi. \end{cases} \tag{6}$$

Each state of $\hat{\Sigma}_P$, denoted by $\xi_l \in \mathbb{R}$, is an approximant of the clustered states given by $x_{[l]} := (e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}x \in \mathbb{R}^{|\mathcal{I}_{[l]}|}$. To see this more specifically, let us consider an example in which $\mathcal{I}_{[l]} = \{1, 2, 3\}$ and $p_{[l]} = [1, 1, 1]/\sqrt{3}$. In this case, as long as the approximation is good, i.e., the trajectory of $\hat{x} \in \mathbb{R}^n$ is close to that of $x \in \mathbb{R}^n$, it follows that $p_{[l]}^{\mathsf{T}}\xi_l =: \hat{x}_{[l]} \simeq x_{[l]}$, namely

$$\frac{1}{\sqrt{3}}[1 \;\; 1 \;\; 1]^{\mathsf{T}}\,\xi_l(t) \simeq [x_1(t) \;\; x_2(t) \;\; x_3(t)]^{\mathsf{T}},$$

where x_i denotes the ith element of x. As shown in this relation, the trajectory of each state of $\hat{\Sigma}_P$ aims to trace the trajectory of a kind of centroid compatible with the clustered states of Σ. In the following, the aggregated model $\hat{\Sigma}_P$ in (6) is said to be *positive* if $PAP^{\mathsf{T}} \in \mathbb{M}_L$, $PB \in \mathbb{R}^{L\times m}_{\ge 0}$, and $P^{\mathsf{T}} \in \mathbb{R}^{n\times L}_{\ge 0}$. Then, we formulate the following problem of clustered model reduction:

Problem Let a semistable positive directed network Σ in (2) be given, and suppose that $A \in \mathbb{M}_n^\dagger$. Given a constant $\epsilon \ge 0$, find an aggregation matrix P in (5) such that the aggregated model $\hat{\Sigma}_P$ in (6) is semistable and positive, and satisfies

$$\|g(s) - \hat{g}(s)\|_{\mathcal{H}_2} \le \epsilon, \tag{7}$$

where the transfer matrices of Σ and $\hat{\Sigma}_P$ are defined by

$$g(s) := (sI_n - A)^{-1}B, \quad \hat{g}(s) := P^{\mathsf{T}}(sI_L - PAP^{\mathsf{T}})^{-1}PB.$$

In model reduction based on the balanced realization [4,15] as well as the positivity-preserving model reduction based on linear matrix inequalities [14], each state of the reduced model is usually obtained as a linear combination of *all* states of the original system, i.e., the transformation matrix is a dense matrix. This clearly contrasts with our problem formulation, because P in (5) is structured as the product of a block-diagonal matrix and a permutation matrix.

3 Clustered Model Reduction Theory

3.1 Exact Clustered Model Reduction

In the rest of this paper, unless otherwise stated, we assume that g in (7) is semistable but not stable, or equivalently, $\lambda_{\mathrm{F}}(A) = 0$ and $v_{\mathrm{F}}^{\mathsf{T}}B \ne 0$. Similar results can be straightforwardly obtained for stable systems. Note that, if $\lambda_{\mathrm{F}}(A) = 0$, guaranteeing the stability of error systems is more challenging than the model reduction of stable systems. This is because the semistability of $\hat{g}$, which can be guaranteed by Lemma 3, does not imply the stability of $g - \hat{g}$ in (7). Note that we can parametrize P in (5) by an aggregation coefficient vector $p \in \mathbb{R}^n$ and a cluster set $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ via

$$p_{[l]} = \|(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}p\|^{-1}\{(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}p\}^{\mathsf{T}}, \quad l \in \mathbb{L}. \tag{8}$$

In the following, for simplicity, we do not care about the scale of p because any multiple of p yields the same P for a fixed $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$. Then, we obtain the following result to prove the stability of error systems:

Lemma 4 Let a semistable positive directed network Σ in (2) be given, and suppose that $A \in \mathbb{M}_n^\dagger$. Consider an aggregation matrix P in (5) with the parameterization of (8), and define g and $\hat{g}$ as in (7). Then, the error system $g - \hat{g}$ is stable for any $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ if and only if $p = v_{\rm F} \in \mathbb{R}^n_{>0}$.

PROOF. As shown in the proof of Theorem 4 in [12], the error system can be in the cascaded form of

$$g(s) - \hat{g}(s) = \Xi(s)\overline{P}^{\mathsf{T}}\overline{P}g(s), \tag{9}$$

where $\Xi(s) := P^{\mathsf{T}}(sI_L - PAP^{\mathsf{T}})^{-1}PA + I_n$. To prove the sufficiency, we first show that both Ξ and $\overline{P}g$ are stable for any $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ if $p = v_{\rm F}$. Note that

$$v_{\rm F} \in \mathrm{im}(P^{\mathsf{T}}) \tag{10}$$

follows for any $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$. Noting that $P^{\mathsf{T}}P$ is the orthogonal projection matrix onto $\mathrm{im}(P^{\mathsf{T}})$, we see that $P^{\mathsf{T}}Pv_{\rm F} = v_{\rm F}$, or equivalently, $\overline{P}v_{\rm F} = 0$. This implies that the eigenspace of $(A, \overline{P})$ associated with $\lambda_{\rm F}(A) = 0$ is unobservable. Hence, all poles of $\overline{P}g$ are in the open left-half plane. Next, we prove the stability of Ξ, whose semistability is ensured by Lemma 3. From (10), it follows that $(Pv_{\rm F})^{\mathsf{T}}$ is the left Frobenius eigenvector of PAP^{T}. Thus, we have $(Pv_{\rm F})^{\mathsf{T}}PA = v_{\rm F}^{\mathsf{T}}A = 0$, which implies that the left eigenspace of (PAP^{T}, PA) associated with $\lambda_{\rm F}(PAP^{\mathsf{T}}) = 0$ is uncontrollable. Hence, all poles of Ξ are also in the open left-half plane.

Next, we prove the necessity by contraposition. Let us consider the case where all states merge into a single cluster, i.e., $\mathcal{I}_{[1]} = \{1, \ldots, n\}$. If $p \neq v_{\rm F}$, we can see that (10) does not hold because $P^{\mathsf{T}} = p$. Thus, we have $\overline{P}v_{\rm F} \neq 0$, which implies that $\overline{P}g$ has a pole on the origin. On the other hand, it follows from the negative semidefiniteness of A shown in (4) that $PAP^{\mathsf{T}} = p^{\mathsf{T}}Ap < 0$, whose strict inequality comes from $p \neq v_{\rm F}$. This implies that Ξ in (9), which is a first order transfer matrix, is stable and its zeros are not on the origin. Hence, the error system in (9) is not stable. □

Lemma 4 shows that the selection of $p = v_{\rm F}$ is necessary and sufficient to guarantee the stability of error systems for any $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$. This is ensured by the fact that the eigenspace of Σ associated with $\lambda_{\rm F}(A) = 0$ is *exactly* preserved in that of $\hat{\Sigma}_P$. Note that $v_{\rm F}$ corresponds to the steady state of Σ, namely, when $u = 0$, it follows that $\lim_{t\to\infty} x(t) = v_{\rm F}$, where we drop the normalization of $v_{\rm F}$. This means that, if we consider an example in which the original Σ has consensus-type dynamics with a symmetric A, the consensus property can be preserved in the aggregated model $\hat{\Sigma}_P$. Thus, we can see that this selection of p reflects the magnitude of the state behavior in a steady state.

Next, we introduce the notion of *cluster reducibility*, which is defined as a kind of uncontrollability of local states. Let us consider the transfer function from the input u to the ith state x_i, denoted by g_i. In this notation, $g_i = g_j$ for $i \neq j$ implies that the ith and jth states have exactly the same behavior for all input signals. Thus, these states would be aggregated into a scalar state under the identical aggregation coefficients. However, note that, as shown in Lemma 4, the aggregation coefficients should comply with the corresponding elements of $v_{\rm F}$ to guarantee the stability of error systems. Therefore, only if $v_j g_i = v_i g_j$, or equivalently, if there exists some $g^\star$ such that

$$\begin{bmatrix} g_i(s) \\ g_j(s) \end{bmatrix} = \begin{bmatrix} v_i \\ v_j \end{bmatrix} g^\star(s),$$

where v_i denotes the ith element of $v_{\rm F}$, we perform the aggregation of x_i and x_j while guaranteeing the stability of error systems. Based on this observation, we define the following cluster reducibility:

Definition 3 Let a semistable positive directed network Σ in (2) be given, and define g as in (7). A cluster $\mathcal{I}_{[l]}$ is said to be *reducible* if there exists a row vector $g^\star_{[l]}$ of rational functions such that

$$(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}g(s) = p^{\mathsf{T}}_{[l]}g^\star_{[l]}(s), \tag{11}$$

where $p_{[l]}$ is defined as in (8) with $p = v_{\rm F} \in \mathbb{R}^n_{>0}$.

Let us consider the characterization of this cluster reducibility in an algebraic manner. For linear systems, the controllability Gramian is often used to analyze the controllability of stable systems. However, note that the usual controllability Gramian cannot be defined for semistable positive directed networks because they are not stable. Therefore, we consider using a projected controllability Gramian as follows:

Theorem 1 Let a semistable positive directed network Σ in (2) be given, and suppose that $A \in \mathbb{M}_n^\dagger$. Let $\overline{V} \in \mathbb{R}^{n\times(n-1)}$ be such that $[v_{\rm F}, \overline{V}] \in \mathbb{R}^{n\times n}$ is unitary, and define $\hat{\Phi} = \hat{\Phi}^{\mathsf{T}} \succeq \mathcal{O}_{n-1}$ such that

$$\overline{V}^{\mathsf{T}}A\overline{V}\hat{\Phi} + \hat{\Phi}\overline{V}^{\mathsf{T}}A^{\mathsf{T}}\overline{V} + \overline{V}^{\mathsf{T}}BB^{\mathsf{T}}\overline{V} = 0. \tag{12}$$

Furthermore, for $\Phi := \overline{V}\hat{\Phi}\overline{V}^{\mathsf{T}} \succeq \mathcal{O}_n$, let $\Phi_{\frac{1}{2}}$ denote a Cholesky factor satisfying $\Phi = \Phi_{\frac{1}{2}}\Phi_{\frac{1}{2}}^{\mathsf{T}}$. Then, a cluster $\mathcal{I}_{[l]}$ is reducible if and only if there exists a row vector $\phi^\star_{[l]} \in \mathbb{R}^{1\times n}$ such that

$$(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}}\Phi_{\frac{1}{2}} = p^{\mathsf{T}}_{[l]}\phi^\star_{[l]}, \tag{13}$$

where $p_{[l]}$ is defined as in (8) with $p = v_{\rm F} \in \mathbb{R}^n_{>0}$. In addition, if all clusters are reducible, then the aggregated model $\hat{\Sigma}_P$ in (6) is semistable and positive, and satisfies $g = \hat{g}$, where g and $\hat{g}$ are defined as in (7).

PROOF. Note that (11) holds if and only if there exists $\overline{p}_{[l]} \in \mathbb{R}^{(|\mathcal{I}_{[l]}|-1)\times|\mathcal{I}_{[l]}|}$ such that $[p_{[l]}^{\mathsf{T}}, \overline{p}_{[l]}^{\mathsf{T}}]^{\mathsf{T}}$ is unitary and

$$\overline{p}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} g(s) = 0. \tag{14}$$

Similarly, (13) is equivalent to the existence of $\overline{p}_{[l]}$ such that

$$\overline{p}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} \Phi_{\frac{1}{2}} = 0. \tag{15}$$

Therefore, in the following, we prove the equivalence between (14) and (15). Note that

$$\overline{p}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} v_{\rm F} = \overline{p}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} P^{\mathsf{T}} P v_{\rm F} = 0 \tag{16}$$

follows from (10) with $p = v_{\rm F}$. Thus, the unitary transformation of the left-hand side of (14) by $[v_{\rm F}, \overline{V}]$ leads to the equality of

$$\overline{p}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} g(s) = \overline{p}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} \overline{V}(sI_{n-1} - \overline{V}^{\mathsf{T}} A \overline{V})^{-1} \overline{V}^{\mathsf{T}} B.$$

Since $\hat{\Phi}$ is the controllability Gramian associated with $(\overline{V}^{\mathsf{T}} A \overline{V}, \overline{V}^{\mathsf{T}} B)$, we have

$$\|\overline{p}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} g(s)\|_{\mathcal{H}_2} = \|\overline{p}_{[l]}(e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} \Phi_{\frac{1}{2}}\|_{\rm F}. \tag{17}$$

Hence, (14) is equivalent to (15). Clearly, $g = \hat{g}$ is ensured by the fact that the reducibility of all clusters with (8) implies that $\overline{P} g = 0$ in (9). Finally, we have $PAP^{\mathsf{T}} \in \mathbb{M}_L^{\dagger}$ and $PB \in \mathbb{R}^{L\times m}_{\geq 0}$ from the positivity of all $p_{[l]}$ with the fact that $Pv_{\rm F}$ and $(Pv_{\rm F})^{\mathsf{T}}$ are the right and left Frobenius eigenvectors of PAP^{T}, respectively. Hence, $\hat{\Sigma}_P$ is semistable and positive. □

Theorem 1 shows that the cluster reducibility can be characterized by linear dependence among the corresponding row vectors of $\Phi_{\frac{1}{2}}$. Note that Φ contains information on the degree of controllability with respect to the stable subspace of Σ. More specifically, it reflects the magnitude of the state behavior when we apply input signals that do not excite the eigenspace associated with $\lambda_{\rm F}(A) = 0$. As shown in this theorem, the aggregation of any reducible cluster causes no approximation error. Even though such an exact aggregation is indeed desirable, the reduction of system dimensions should be restrictive.

3.2 Clustered Model Reduction in Terms of the $\mathcal{H}_2$-Norm

To perform an approximation error analysis in terms of the $\mathcal{H}_2$-norm, we define a relaxed notion of cluster reducibility as follows:

Definition 4 Let a semistable positive directed network Σ in (2) be given. Under the same notation as that in Theorem 1, a cluster $\mathcal{I}_{[l]}$ is said to be *θ-reducible* if there exists $\phi^{\star}_{[l]} \in \mathbb{R}^{1\times n}$ such that

$$\left\| (e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} \Phi_{\frac{1}{2}} - p_{[l]}^{\mathsf{T}} \phi^{\star}_{[l]} \right\|_{\rm F} \leq \sqrt{|\mathcal{I}_{[l]}|}\,\theta, \quad \theta \geq 0. \tag{18}$$

In Definition 4, the constant θ represents a quantitative index of cluster reducibility, where the scaling by $\sqrt{|\mathcal{I}_{[l]}|}$ is introduced for technical reasons. Clearly, the relaxed condition in (18) includes the exact condition in (13), and they are equivalent if $\theta = 0$. This inclusion implies that, in the construction of clusters, we find a set of states that behave similarly for all input signals. Intuitively, giving a small θ, we can expect to obtain $\hat{\Sigma}_P$ that is a good approximant of the original Σ. Then, we can obtain the following result:

Theorem 2 Let a semistable positive directed network Σ in (2) be given, and suppose that $A \in \mathbb{M}_n^{\dagger}$. Under the same notation as that in Theorem 1, let $\gamma > 0$ be such that

$$\overline{V}^{\mathsf{T}} A \overline{V} + \overline{V}^{\mathsf{T}} A^{\mathsf{T}} \overline{V} + \gamma^{-1}(\overline{V}^{\mathsf{T}} A A^{\mathsf{T}} \overline{V} + I_{n-1}) \prec \mathcal{O}_{n-1}. \tag{19}$$

If all clusters are θ-reducible, then the aggregated model $\hat{\Sigma}_P$ in (6) is semistable and positive, and satisfies

$$\|g(s) - \hat{g}(s)\|_{\mathcal{H}_2} \leq \gamma \sqrt{\textstyle\sum_{l=1}^{L} |\mathcal{I}_{[l]}|(|\mathcal{I}_{[l]}| - 1)}\;\theta. \tag{20}$$

PROOF. By (9), $\|g - \hat{g}\|_{\mathcal{H}_2} \leq \|\Xi\|_{\mathcal{H}_\infty} \|\overline{P}^{\mathsf{T}} \overline{P} g\|_{\mathcal{H}_2}$ holds, where both Ξ and $\overline{P} g$ are stable as shown in the proof of Lemma 4. Since the Riccati inequality

$$PA^{\mathsf{T}} P^{\mathsf{T}} X + XPAP^{\mathsf{T}} + \frac{1}{\mu}(XPAA^{\mathsf{T}} P^{\mathsf{T}} X + I_L) \prec \mathcal{O}_L$$

associated with Ξ is identical to that associated with

$$\hat{\Xi}(s) := (sI_L - PAP^{\mathsf{T}})^{-1} PA,$$

we verify the equality of $\|\Xi\|_{\mathcal{H}_\infty} = \|\hat{\Xi}\|_{\mathcal{H}_\infty}$. Thus, in the following, we analyze $\|\hat{\Xi}\|_{\mathcal{H}_\infty}$. For $\tilde{V} \in \mathbb{R}^{L\times(L-1)}$ such that $[Pv_{\rm F}, \tilde{V}]$ is unitary, we can rewrite $\hat{\Xi}$ as

$$\hat{\Xi}(s) = \tilde{V}(sI_{L-1} - \tilde{V}^{\mathsf{T}} PAP^{\mathsf{T}} \tilde{V})^{-1} \tilde{V}^{\mathsf{T}} PA,$$

where $\tilde{V}^{\mathsf{T}} PAP^{\mathsf{T}} \tilde{V}$ is stable. Based on this representation, let us prove that $\|\hat{\Xi}\|_{\mathcal{H}_\infty} < \gamma$ by virtue of (19). To this end, we first prove that there exists some $\gamma > 0$ such that (19) holds. From $A \in \mathbb{M}_n^{\dagger}$, it follows that

$$A + A^{\mathsf{T}} \in \mathbb{M}_n^{\dagger}, \quad \lambda_{\rm F}(A + A^{\mathsf{T}}) = 0, \quad (A + A^{\mathsf{T}}) v_{\rm F} = 0.$$

Thus, $X := \overline{V}^{\mathsf{T}} A \overline{V} + \overline{V}^{\mathsf{T}} A^{\mathsf{T}} \overline{V}$ is negative definite. Note that $X \preceq -\lambda_1 I_{n-1}$ holds for λ_1 denoting the minimal eigenvalue of $-X$. Similarly, for λ_2 denoting the maximal eigenvalue of the positive definite $Y := \overline{V}^{\mathsf{T}} A A^{\mathsf{T}} \overline{V} + I_{n-1}$, we have $Y \preceq \lambda_2 I_{n-1}$. Therefore, the left-hand side of (19) satisfies

$$X + \gamma^{-1} Y \preceq (-\lambda_1 + \gamma^{-1}\lambda_2) I_{n-1}.$$

Hence, (19) is satisfied for any $\gamma > \lambda_2/\lambda_1$. Multiplying (19) by $\tilde{V}^{\mathsf{T}} P \overline{V}$ and $(\tilde{V}^{\mathsf{T}} P \overline{V})^{\mathsf{T}}$ from the left and right, respectively, we have

$$\begin{aligned} &\tilde{V}^{\mathsf{T}} P A P^{\mathsf{T}} \tilde{V} + \tilde{V}^{\mathsf{T}} P A^{\mathsf{T}} P^{\mathsf{T}} \tilde{V} \\ &\quad + \gamma^{-1} (\tilde{V}^{\mathsf{T}} P A A^{\mathsf{T}} P^{\mathsf{T}} \tilde{V} + I_{L-1}) \prec \mathcal{O}_{L-1}, \end{aligned} \tag{21}$$

where we have used $\overline{V}\,\overline{V}^{\mathsf{T}} A = A \overline{V}\,\overline{V}^{\mathsf{T}} = A$ and

$$\tilde{V}^{\mathsf{T}} P \overline{V}\,\overline{V}^{\mathsf{T}} P^{\mathsf{T}} \tilde{V} = \tilde{V}^{\mathsf{T}} P (I_n - v_{\mathrm{F}} v_{\mathrm{F}}^{\mathsf{T}}) P^{\mathsf{T}} \tilde{V} = I_{L-1}.$$

Thus, the bounded real lemma ensures $\|\hat{\Xi}\|_{\mathcal{H}_\infty} < \gamma$ by (21). Next, we evaluate $\|\overline{P}^{\mathsf{T}} \overline{P} g\|_{\mathcal{H}_2}$. From (17), it follows that $\|\overline{P}^{\mathsf{T}} \overline{P} g\|_{\mathcal{H}_2} = \|\overline{P} \Phi_{\frac{1}{2}}\|_{\mathrm{F}}$, where $\overline{P}$ can be formed by replacing each $p_{[l]}$ in (5) with $\overline{p}_{[l]}$. Since $[p_{[l]}^{\mathsf{T}}, \overline{p}_{[l]}^{\mathsf{T}}]^{\mathsf{T}}$ is unitary, we have

$$\overline{p}_{[l]} (e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} \Phi_{\frac{1}{2}} = \overline{p}_{[l]} \Delta_{[l]}, \quad \Delta_{[l]} := (e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} \Phi_{\frac{1}{2}} - p_{[l]}^{\mathsf{T}} \phi^\star_{[l]},$$

where $\overline{p}_{[l]}$ is allowed to be empty if $|\mathcal{I}_{[l]}| = 1$. The θ-reducibility in Definition 4 implies $\|\Delta_{[l]}\|_{\mathrm{F}} \le \sqrt{|\mathcal{I}_{[l]}|}\,\theta$. Thus, we have

$$\|\overline{P} \Phi_{\frac{1}{2}}\|_{\mathrm{F}}^2 \le \sum_{l=1}^{L} \|\overline{p}_{[l]}\|_{\mathrm{F}}^2 \|\Delta_{[l]}\|_{\mathrm{F}}^2 \le \sum_{l=1}^{L} |\mathcal{I}_{[l]}| (|\mathcal{I}_{[l]}| - 1) \theta^2,$$

where we have used $\|\overline{p}_{[l]}\|_{\mathrm{F}}^2 = \mathrm{tr}(\overline{p}_{[l]} \overline{p}_{[l]}^{\mathsf{T}}) = |\mathcal{I}_{[l]}| - 1$. Hence, (20) follows. Finally, the positivity of $\hat{\Sigma}_P$ follows from the positivity of all $p_{[l]}$. □

Theorem 2 shows that there is a linear dependence between θ and the approximation error. Thus, we can use θ as a design parameter to regulate the approximation quality of resultant aggregated models. Note that the clustered model reduction method can generally provide better approximation than a positivity-preserving model reduction method performed by the simple truncation of state variables having a small influence on the input-to-output mapping [17]. This is because the clustered model reduction can aggregate the states not only having small behavior but also having similar behavior for input signals.

In model reduction theory, it is known that guaranteeing semistability preservation and performing an error analysis are generally challenging problems, especially in structure-preserving model reduction [19]. Our success stems from the clarification that

- there exists a diagonal similarity transformation to make a Metzler matrix negative semidefinite, as shown in Lemma 2,
- the stability of error systems can be guaranteed by giving aggregation coefficients associated with the Frobenius eigenvector, as shown in Lemma 4, and
- each row vector in the Cholesky factor of a projected controllability Gramian, compatible with the $\mathcal{H}_2$-analysis, algebraically captures the characteristics of state behavior for input signals, as shown in Theorems 1 and 2.

3.3 Systematic Implementation

In this subsection, we propose an algorithm for solving the clustered model reduction problem in Section 2.2 with a prescribed value of ϵ. To this end, we first provide a procedure to construct a set of θ-reducible clusters, on the premise that $\theta \ge 0$ is given and $\Phi_{\frac{1}{2}}$ is obtained. Assume that a set of clusters $\{\mathcal{I}_{[1]}, \ldots, \mathcal{I}_{[l-1]}\}$ has already been formed, and let $\mathcal{J} := \{1, \ldots, n\} \backslash \bigcup_{k=1}^{l-1} \mathcal{I}_{[k]}$ denote the set of state indices that do not belong to any cluster. When we make a new cluster $\mathcal{I}_{[l]}$, we first choose an index $j \in \mathcal{J}$, and then look for indices $i \in \mathcal{J}$ satisfying

$$\left\| \phi_i - v_i v_j^{-1} \phi_j \right\| \le \theta, \tag{22}$$

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_{F}. Note that $\|M\|_{\mathrm{F}}^2 = \sum_{i=1}^{m} \|M_i\|^2$ holds, where $M_i \in \mathbb{R}^{1 \times n}$ denotes the ith row vector of $M \in \mathbb{R}^{m \times n}$. Thus, if (22) holds for all $i \in \mathcal{I}_{[l]} \subseteq \mathcal{J}$, then

$$\left\| (e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} \Phi_{\frac{1}{2}} - p_{[l]}^{\mathsf{T}} \phi_j \right\|_{\mathrm{F}}^2 = \sum_{i \in \mathcal{I}_{[l]}} \|\phi_i - v_i v_j^{-1} \phi_j\|^2 \le |\mathcal{I}_{[l]}| \theta^2.$$

Hence, we see that the new cluster $\mathcal{I}_{[l]}$ is θ-reducible. On the basis of this, an algorithm producing a set of θ-reducible clusters is provided as follows:

(i) Initialize temporal variables as $l = 0$ and $\mathbb{L} = \emptyset$.
(ii) Repeat (iii)–(v) while $\bigcup_{k \in \mathbb{L}} \mathcal{I}_{[k]} \neq \{1, \ldots, n\}$.
(iii) Update the variables as $l \leftarrow l + 1$ and $\mathbb{L} \leftarrow \{\mathbb{L}, l\}$.
(iv) Choose $i \in \{1, \ldots, n\} \backslash \bigcup_{k \in \mathbb{L}} \mathcal{I}_{[k]}$ and set $\mathcal{I}_{[l]} = \{i\}$.
(v) For all $j \in \{1, \ldots, n\} \backslash \bigcup_{k \in \mathbb{L}} \mathcal{I}_{[k]}$, if (22) holds for i and j, then $\mathcal{I}_{[l]} \leftarrow \{\mathcal{I}_{[l]}, j\}$.

It should be noted that this algorithm does not necessarily produce a unique cluster set, namely a degree of freedom remains in the cluster construction. The explicit consideration on this remaining freedom would be a meaningful future work to pursue.

From the alternative expression in (22), we see that, when $\|v_j^{-1} \phi_j\| \gg \theta$ and $\|v_j^{-1} \phi_j\| \gg \|v_i^{-1} \phi_i\|$, the cluster

to which the jth state belongs should not merge with the ith state. From this observation, we can see that the magnitude of $\|v_j^{-1}\phi_j\|$ can be regarded as a kind of *uniqueness* of the corresponding state variable. Note that v_j corresponds to the magnitude of the state behavior in a steady state, and ϕ_j contains information on the degree of controllability with respect to the stable subspace. Thus, we see a tendency for a state variable not to be clustered if its controllability is relatively large while its steady state value is small.

Some discussion on the conservativeness of the error bound is in order. Indeed, even though Theorem 2 provides a theoretically reasonable strategy to find $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$, the error bound in (20) may become conservative, especially in a large-scale setting. This is because no information on P and $\overline{P}$ is taken into account to derive the upper bound of γ in (19). To compensate for this weakness, we propose an efficient method to calculate the approximation error by utilizing the *cascaded* form of the error system in (9). Let $\overline{V} \in \mathbb{R}^{n\times(n-1)}$ and $\tilde{V} \in \mathbb{R}^{L\times(L-1)}$ be such that $[v_{\mathrm{F}}, \overline{V}]$ and $[Pv_{\mathrm{F}}, \tilde{V}]$ are unitary, respectively. Then, it follows that

$$\begin{aligned}\overline{P}^{\mathsf{T}}\overline{P}g(s) &= \overline{P}^{\mathsf{T}}\overline{P}\,\overline{V}(sI_{n-1} - \overline{V}^{\mathsf{T}}A\overline{V})^{-1}\overline{V}^{\mathsf{T}}B\\ \Xi(s) &= P^{\mathsf{T}}\tilde{V}(sI_{L-1} - \tilde{V}^{\mathsf{T}}PAP^{\mathsf{T}}\tilde{V})^{-1}\tilde{V}^{\mathsf{T}}PA + I_n.\end{aligned}$$

Thus, the error system can be represented as

$$g(s) - \hat{g}(s) = \mathcal{C}(sI_{n+L-2} - \mathcal{A})^{-1}\mathcal{B}, \tag{23}$$

where

$$\mathcal{A} = \begin{bmatrix}\mathcal{A}_{11} & \mathcal{A}_{12}\\ 0 & \mathcal{A}_{22}\end{bmatrix},\ \mathcal{B} = \begin{bmatrix}0\\ \overline{V}^{\mathsf{T}}B\end{bmatrix},\ \mathcal{C} = \begin{bmatrix}P^{\mathsf{T}}\tilde{V} & \overline{P}^{\mathsf{T}}\overline{P}\,\overline{V}\end{bmatrix}$$

with $\mathcal{A}_{11} := \tilde{V}^{\mathsf{T}}PAP^{\mathsf{T}}\tilde{V}$, $\mathcal{A}_{12} := \tilde{V}^{\mathsf{T}}PA\overline{P}^{\mathsf{T}}\overline{P}\,\overline{V}$, and $\mathcal{A}_{22} := \overline{V}^{\mathsf{T}}A\overline{V}$. For a fixed P in (5), let $\mathcal{X} \succeq \mathcal{O}_{L-1}$ and $\mathcal{Y} \in \mathbb{R}^{(L-1)\times(n-1)}$ be solutions of the following Lyapunov and Sylvester equations:

$$\begin{aligned}&\mathcal{A}_{11}\mathcal{X} + \mathcal{X}\mathcal{A}_{11}^{\mathsf{T}} + \mathcal{A}_{12}\mathcal{Y}^{\mathsf{T}} + \mathcal{Y}\mathcal{A}_{12}^{\mathsf{T}} = 0,\\ &\mathcal{A}_{11}\mathcal{Y} + \mathcal{Y}\mathcal{A}_{22}^{\mathsf{T}} + \mathcal{A}_{12}\hat{\Phi} = 0,\end{aligned} \tag{24}$$

where $\hat{\Phi}$ is the solution of (12). Since the controllability Gramian for the realization of the error system in (23) is given by

$$\mathcal{W} := \begin{bmatrix}\mathcal{X} & \mathcal{Y}\\ \mathcal{Y}^{\mathsf{T}} & \hat{\Phi}\end{bmatrix},$$

the approximation error in terms of the $\mathcal{H}_2$-norm is calculated with

$$\|g(s) - \hat{g}(s)\|_{\mathcal{H}_2} = \sqrt{\mathrm{tr}(\mathcal{X}) + \mathrm{tr}(\overline{P}\Phi\overline{P}^{\mathsf{T}})}, \tag{25}$$

where we have used

$$\begin{aligned}&\mathrm{tr}(P^{\mathsf{T}}\tilde{V}\mathcal{X}\tilde{V}^{\mathsf{T}}P) = \mathrm{tr}(\tilde{V}^{\mathsf{T}}PP^{\mathsf{T}}\tilde{V}\mathcal{X}) = \mathrm{tr}(\mathcal{X}),\\ &\mathrm{tr}(\overline{P}^{\mathsf{T}}\overline{P}\overline{V}\hat{\Phi}\overline{V}^{\mathsf{T}}\overline{P}^{\mathsf{T}}\overline{P}) = \mathrm{tr}(\overline{P}\overline{P}^{\mathsf{T}}\overline{P}\Phi\overline{P}^{\mathsf{T}}) = \mathrm{tr}(\overline{P}\Phi\overline{P}^{\mathsf{T}})\end{aligned}$$

with $PP^{\mathsf{T}} = I_L$, $\tilde{V}^{\mathsf{T}}\tilde{V} = I_{L-1}$ and $\overline{P}\overline{P}^{\mathsf{T}} = I_{n-L}$. Note that $\hat{\Phi}$ in (24) is obtained in advance to find P because it is used to find Φ. Furthermore, since we only need to find the lower-dimensional solutions $\mathcal{X}$ and $\mathcal{Y}$ individually, the equations in (24) can be solved more efficiently than the Lyapunov equation with respect to the simple $(n+L)$-dimensional error system. Using this a posteriori error calculation, the following algorithm can be used to solve the clustered model reduction problem:

(a) Prescribe the admissible error $\epsilon \geq 0$.
(b) Find Φ and calculate its Cholesky factor $\Phi_{\frac{1}{2}} \in \mathbb{R}^{n\times n}$.
(c) For a fixed $\theta \geq 0$, construct a set of θ-reducible clusters $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$.
(d) Find $\mathcal{X} \in \mathbb{R}^{(L-1)\times(L-1)}$ by solving the Lyapunov and Sylvester equations in (24).
(e) Calculate the resultant approximation error by (25).
(f) If the approximation error is not less than ϵ, then return to (c) after setting a smaller $\theta \geq 0$.

We give a note on the computational effort required to compute solutions to the Lyapunov and Sylvester equations. Even though their computation may become time consuming in a large-scale setting, a number of effective methods for solving large-scale Lyapunov and Sylvester equations can be found in the literature. One widely used approach is an iterative calculation based on Krylov subspace methods. For example, [7,10] develop approximate solution algorithms by explicitly considering the low-rank nature of system matrices. Furthermore, we give an additional note on the initial choice of θ. We notice from (22) that its order of magnitude needs to be comparable with that of $\|\phi_i\|$ in the construction of a reasonable cluster set. This can be confirmed by the fact that, if $\|\phi_i\| \ll \theta$ for all $i \in \{1,\ldots,n\}$, then all states should merge into a single cluster. Finally, note that the posteriori error calculation in Steps (d)–(e) can be skipped when we are interested in the dimension of approximants. Such an approximation is effective when only limited memory storage is available.

3.4 Generalization to Reducible Systems

In this subsection, we consider relaxing the irreducibility condition of $A \in \mathbb{M}_n^{\dagger}$ assumed in Definition 1. Note that, for a reducible matrix A, we cannot apply the diagonal similarity transformation shown in Lemma 2, because its left and right Frobenius eigenvectors may have zero entries. This spoils the property in (10) that can be rephrased as

$$\mathrm{im}(W) \subseteq \mathrm{im}(P^{\mathsf{T}}), \quad W := \begin{bmatrix}v_{\mathrm{L}}^{\mathsf{T}}(A) & v_{\mathrm{R}}(A)\end{bmatrix}. \tag{26}$$

This relation is fundamental to the above arguments used to prove the stability of error systems. Conversely, if (26) is satisfied, the assumption of irreducibility is not required to prove the stability of error systems.

Without loss of generality, we can assume that a reducible matrix A is structured as

$$A = \begin{bmatrix} A_{1,1} & 0 & \cdots & 0 \\ A_{2,1} & A_{2,2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ A_{K,1} & A_{K,2} & \cdots & A_{K,K} \end{bmatrix} \in \mathbb{R}^{n\times n}, \quad n = \sum_{k=1}^{K} n_k,$$

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

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

Even though the second and third assumptions are not necessarily trivial for a reducible matrix A, we shall see in Section 4 that they are indeed satisfied by a CME derived from enzyme-catalyzed reaction systems. Then, it is readily verified that $\lambda_{\mathrm{F}}(A) = \lambda_{\mathrm{F}}(A_{K,K}) = 0$. In addition, since $A_{K,K} \in \mathbb{M}^{\dagger}_{n_K}$, the left and right Frobenius eigenvectors can be described as $v_{\mathrm{L}}(A) = [w \;\; v_{\mathrm{L}}(A_{K,K})]$ and $v_{\mathrm{R}}(A) = \begin{bmatrix} 0 & v_{\mathrm{L}}^{\mathsf{T}}(A_{K,K}) \end{bmatrix}^{\mathsf{T}}$ where $w \in \mathbb{R}^{1\times(n-n_K)}_{>0}$ denotes a positive row vector. To guarantee the inclusion in (26), we construct a set of θ-reducible clusters by imposing the following additional constraint: Let $\mathcal{N}_k$ denote an index set corresponding to the kth block of A, namely $A_{k,k} = (e^n_{\mathcal{N}_k})^{\mathsf{T}} A e^n_{\mathcal{N}_k}$ for $k \in \{1, \ldots, K\}$. In this notation, we consider forming clusters inside $\mathcal{N}_{1:K-1} := \bigcup_{i=1}^{K-1} \mathcal{N}_i$ and $\mathcal{N}_K$ separately. More specifically, we form $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ while imposing the additional constraint

$$\bigcup_{l=1}^{l_0} \mathcal{I}_{[l]} = \mathcal{N}_{1:K-1}, \qquad \bigcup_{l=l_0+1}^{L} \mathcal{I}_{[l]} = \mathcal{N}_K \tag{27}$$

with some $l_0 \in \mathbb{L}$. By giving each $p_{[l]}$ in (8) with $p = v_{\mathrm{L}}^{\mathsf{T}}(A)$, it turns out that the resultant aggregation matrix satisfies (26). This is verified by the fact that P is to be structured as $P = \mathrm{Diag}(P_{1:K-1}, P_K) \in \mathbb{R}^{L\times n}$ where $P_{1:K-1} \in \mathbb{R}^{l_0\times(n-n_K)}$, and $P_K \in \mathbb{R}^{(L-l_0)\times n_K}$ satisfies

$$v_{\mathrm{L}}(A_{K,K}) P_K^{\mathsf{T}} P_K = v_{\mathrm{L}}(A_{K,K}).$$

Hence, the stability of error systems is guaranteed by the aggregation matrix P associated with $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ satisfying (27).

Moreover, we can derive an $\mathcal{H}_2$-error bound similar to that in Theorem 2, if all clusters are θ-reducible. In this case, though it is difficult to analytically show an a priori bound of $\|\Xi\|_{\mathcal{H}_\infty}$, we instead obtain

$$\|g(s) - \hat{g}(s)\|_{\mathcal{H}_2} \leq \|\Xi(s)\|_{\mathcal{H}_\infty} \sqrt{\textstyle\sum_{l=1}^{L} |\mathcal{I}_{[l]}|(|\mathcal{I}_{[l]}| - 1)}\; \theta,$$

where Ξ is defined as in (9). This expression implies that, even for Σ with a reducible matrix A, the parameter θ can be used as a design criterion to regulate the approximation quality of the resultant $\hat{\Sigma}_P$.

3.5 On the Applicability of Proposed Clustered Model Reduction

For the clustered model reduction method developed in Sections 3.1–3.4, the existence of the diagonal similarity transformation in Lemma 2 is a key factor in guaranteeing the semistability of the resultant aggregated models. The existence of such a diagonal similarity transformation is closely related to the existence of diagonal Lyapunov functions. This can be explained as follows: Let $A \in \mathbb{R}^{n\times n}$ be such that

$$A^{\mathsf{T}} V + V A \preceq \mathcal{O}_n \tag{28}$$

for some $V = V^{\mathsf{T}} \succ \mathcal{O}_n$, which is used to prove the semistability of A. Let us consider the Cholesky factor $V_{\frac{1}{2}}$ such that $V = V_{\frac{1}{2}} V_{\frac{1}{2}}^{\mathsf{T}}$. Multiplying (28) by $V_{\frac{1}{2}}^{-1}$ and $V_{\frac{1}{2}}^{-\mathsf{T}}$ from the left and right, respectively, it follows that

$$\tilde{A}^{\mathsf{T}} + \tilde{A} \preceq \mathcal{O}_n, \quad \tilde{A} := V_{\frac{1}{2}}^{\mathsf{T}} A V_{\frac{1}{2}}^{-\mathsf{T}}. \tag{29}$$

Thus, any A satisfying (28) is similar to the negative semidefinite matrix $\tilde{A}$. In these specific system coordinates, we can always perform projection-based model reduction that preserves the semistability, as long as we use orthogonal projection. However, the transformation in (29) is generally unworkable in clustered model reduction. This is because the network structure of the original A is destroyed through the transformation, since $V_{\frac{1}{2}}$ generally turns out to be a dense matrix.

In view of this, we can see that the existence of diagonal Lyapunov functions is key to theoretically guaranteeing the semistability preservation in clustered model reduction. One particular class of systems that admit diagonal Lyapunov functions is those with a symmetric system matrix. For this class of systems, we have developed [12] a clustered model reduction method, in which an error bound in terms of the $\mathcal{H}_\infty$-norm is guaranteed and the symmetry of the system matrix is preserved on the basis of orthogonal projection. In this paper, we enhance the applicability of our clustered model reduction by using the fact that any positive directed network admits a diagonal Lyapunov function, and cluster reducibility can be characterized for semistable systems based on a projected controllability Gramian. Furthermore, we

have clarified that a specific condition on the aggregation matrices, like (26), is crucial in guaranteeing the stability of error systems. Note that several characterizations of systems admitting diagonal Lyapunov functions can be found in the literature [2,13]. In particular, [13] develops an effective method to determine their existence through a feasibility test for a linear programming problem. Based on these characterizations, one can address a clustered model reduction problem for a broader class of systems.

4 Illustrative Example

4.1 Chemical Master Equation

In this section, we consider the following chemical reaction system composed of four different molecules, denoted by S_i for $i \in \{1, \dots, 4\}$:

$$S_1 + S_2 \underset{c_2}{\overset{c_1}{\rightleftarrows}} S_3, \quad S_3 \overset{c_3}{\to} S_4 + S_2, \tag{30}$$

where $c_i \geq 0$ denote the reaction rate constants. This class of a chemical reaction system is known as the mechanism of enzyme-catalyzed reactions [9]. Let us consider the case in which the initial number of molecules of both S_1 and S_2 is N_0. In this case, all realizable distributions of the number of molecules are enumerated by the $n = n(N_0) := (N_0^2 + 3N_0 + 2)/2$ kinds of vectors

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

where the ith entry denotes the number of molecules of S_i for $i \in \{1, \dots, 4\}$. We denote each realizable distribution by $\xi^{(i)} \in \mathbb{N}^4$ for $i \in \{1, \dots, n\}$, and the realization probability of $\xi^{(i)}$ at time t by $x_i(t) \in [0, 1]$. In this notation, the time evolution of (30) is captured by the CME

$$\dot{x} = Ax, \quad x(0) = e_1^n, \tag{31}$$

where $x := [x_1, \dots, x_n]^\mathsf{T} \in [0, 1]^n$ and $A \in \mathbb{R}^{n \times n}$ is semistable and Metzler; see [9] for the derivation. Note that the initial condition $x(0)$ represents the situation where the realization probability $x_1(t)$ of $\xi^{(1)} := [N_0, N_0, 0, 0]^\mathsf{T}$ is equal to 1 at $t = 0$. Furthermore, it follows that $\sum_{i=1}^n x_i(t) = 1$ for all $t \geq 0$ because each x_i represents the realization probability of distribution $\xi^{(i)}$. Accordingly, A in (31) satisfies

$$\lambda_\mathrm{F}(A) = 0, \quad \begin{cases} v_\mathrm{L}(A) = [1, \dots, 1], \\ v_\mathrm{R}(A) = \lim_{t \to \infty} x(t) = e_n^n, \end{cases} \tag{32}$$

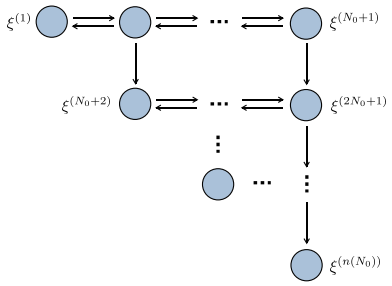

Fig. 2. Transition diagram of molecule number distributions.

where we drop the normalization of the Frobenius eigenvectors. In Fig. 2, where the distributions are enumerated as $\xi^{(1)} = [N_0, N_0, 0, 0]^\mathsf{T}, \dots, \xi^{(n(N_0))} = [0, N_0, 0, N_0]^\mathsf{T}$, we show the transition diagram of (30) in the CME expression. The horizontal transitions correspond to the first and second reactions, and the vertical transitions correspond to the third reaction in (30). Since the initial number of molecules is generally given as a value of the order of a few dozen or hundred, the dimension of (31) turns out to be somewhat large. Thus, even though the CME can capture the behavior of the chemical reaction in terms of a set of linear ordinary differential equation, it is not necessarily analytically or numerically tractable [9,16]. Therefore, to enable practical analyses, it is desirable to derive a *mesoscopic model* that approximates the essential properties of the microscopic expression in (31).

4.2 Clustered Model Reduction of the Chemical Master Equation

In this subsection, based on the fact that $(sI_n - A)^{-1}x(0)$ is equal to the Laplace transform of $e^{At}x(0)$, we approximate the CME in (31) described by Σ in (2) with $B = x(0)$. We apply the algorithm proposed in Section 3.3 to construct an aggregated model that satisfies a prescribed error precision. Let us prescribe the admissible error as $\epsilon = 7.6 \times 10^{-3}$. Against each value of θ, Fig. 3–(a) plots the resultant dimension of the aggregated model. Furthermore, in Fig. 3–(b), we plot the resultant approximation error, which is the value given by (25). These figures show that the dimension of the aggregated model increases, and the approximation error appropriately decreases, as θ decreases. This result confirms that the value of θ successfully captures the approximation quality of the aggregated models. By decreasing the value of θ according to Steps (c)–(f), we find that, when $\theta = 7.99 \times 10^{-6}$, the original 3321-dimensional CME is reduced to a 684-dimensional version, and the resultant approximation error is 6.2×10^{-3}, which is less than the prescribed ϵ. Fig. 3–(c) shows the trajectory of the original system $x \in \mathbb{R}^n$ (solid lines, 3321-dimensional) and that of the aggregated model $\hat{x} \in \mathbb{R}^n$ (dashed lines, 684-dimensional). We can see from this figure that the behavior of the CME in (31) is well approximated by that of the aggregated model.

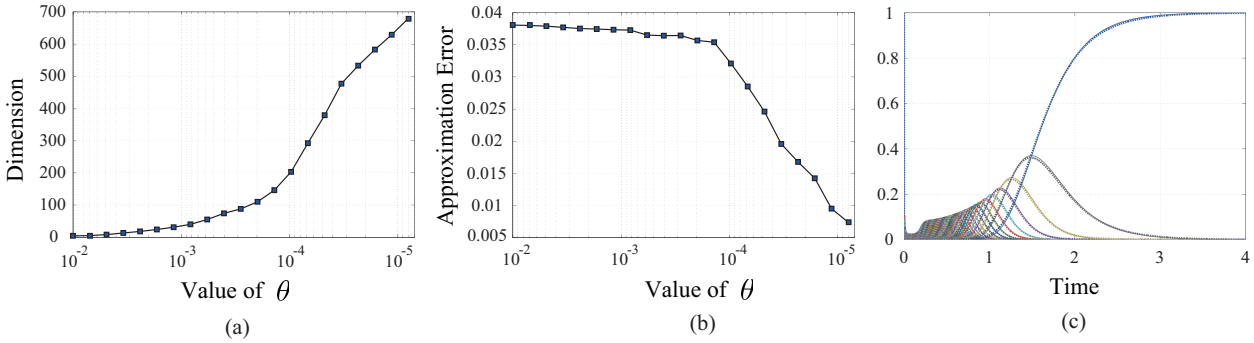

Fig. 3. (a) Dimension and (b) approximation error of aggregated models versus values of θ, (c) state trajectory of original system and aggregated model.

From the result above, we can conclude that the proposed clustered model reduction method works well even for reducible CMEs. In addition, the aggregated model not only approximates the behavior of (31), but also preserves particular properties of the CME. More specifically, the aggregated model has the following properties:

- the positivity of systems (the non-negativity of the state trajectory) is preserved,
- the sum of all states is equal to 1 for any $t \geq 0$,
- the steady state distribution is exactly preserved, and
- each state of the aggregated model tracks a centroid of a set of the original states,

where the second and third items are guaranteed by (26).

5 Concluding Remarks

In this paper, we have developed a clustered model reduction method for semistable positive systems evolving over directed networks, called positive directed networks. In comparison with our previous work [12], the applicability of clustered model reduction has been improved by the developments reported in this paper. This success stems from the clarification that the existence of diagonal Lyapunov functions is key to guaranteeing the semistability preservation in clustered model reduction, and cluster reducibility, defined as the uncontrollability of local states, can be characterized for semistable systems based on a projected controllability Gramian. The efficiency of the proposed method has been demonstrated through an illustrative example using CMEs, which are a type of Markovian process. The proposed method can systematically construct a mesoscopic model of CMEs that well approximates the behavior of the original system while preserving particular properties of the Markovian processes.

As for the optimality of approximation, although a necessary condition for the $\mathcal{H}_2$-optimality is derived for approximate models with a set of a priori fixed poles [6], an adequate solution to the optimal approximation problem has not yet been given even in unstructured model reduction. In view of this, it would be natural to say that the optimal approximation is harder to pursue in the clustered model reduction, where a non-negative sparse structure is imposed on aggregation matrices. This implies that structure-preserving model reduction is generally difficult to develop without the sacrifice of approximation accuracy.

References

[1] A. C. Antoulas. *Approximation of large-scale dynamical systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2005.

[2] G. Barker, A. Berman, and R. Plemmons. Positive diagonal solutions to the Lyapunov equations. *Linear and Multilinear Algebra*, 5(4):249–256, 1978.

[3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424(4):175–308, 2006.

[4] D. F. Enns. Model reduction with balanced realizations: An error bound and a frequency weighted generalization. In *Decision and Control (CDC), 1984 Proceedings of the 23rd IEEE Conference on*, volume 23, pages 127–132. IEEE, 1984.

[5] L. Farina and S. Rinaldi. *Positive linear systems: Theory and applications*. Wiley New York, 2000.

[6] S. Gugercin, A. C. Antoulas, and C. Beattie. H_2 model reduction for large-scale linear dynamical systems. *SIAM Journal on Matrix Analysis and Applications*, 30(2):609–638, 2008.

[7] S. Gugercin, D. C. Sorensen, and A. C. Antoulas. A modified low-rank Smith method for large-scale Lyapunov equations. *Numerical Algorithms*, 32(1):27–55, 2003.

[8] S. Gugercin and K. Willcox. Krylov projection framework for Fourier model reduction. *Automatica*, 44(1):209–215, 2008.

[9] D. J. Higham. Modeling and simulating chemical reactions. *SIAM review*, 50(2):347–368, 2008.

[10] D. Y. Hu and L. Reichel. Krylov-subspace methods for the Sylvester equation. *Linear Algebra and its Applications*, 172:283–313, 1992.

[11] T. Ishizaki, K. Kashima, A. Girard, J. Imura, L. Chen, and K. Aihara. Clustering-based $\mathcal{H}_2$-state aggregation of positive networks and its application to reduction of chemical master equations. In *Decision and Control (CDC), 2012 Proceedings of the 51st IEEE Conference on*, pages 4175–4180. IEEE, 2012.

[12] T. Ishizaki, K. Kashima, J. Imura, and K. Aihara. Model reduction and clusterization of large-scale bidirectional networks. *Automatic Control, IEEE Transactions on*, 59(1):48–63, 2014.

[13] H. K. Khalil. On the existence of positive diagonal P such that $PA + A^T P < 0$. *Automatic Control, IEEE Transactions on*, 27(1):181–184, 1982.

[14] P. Li, J. Lam, Z. Wang, and P. Date. Positivity-preserving H_∞ model reduction for positive systems. *Automatica*, 47(7):1504–1511, 2011.

[15] C.-A. Lin and T.-Y. Chiu. Model reduction via frequency weighted balanced realization. In *American Control Conference (ACC), 1990*, pages 2069–2070. IEEE, 1990.

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

[17] A. Sootla and A. Rantzer. Scalable positivity preserving model reduction using linear energy functions. In *Decision and Control (CDC), 2012 Proceedings of the 51st IEEE Conference on*, pages 4285–4290. IEEE, 2012.

[18] A. van der Schaft and R. Polyuga. Structure-preserving model reduction of complex physical systems. In *Decision and Control, held jointly with the 28th Chinese Control Conference (CDC/CCC), 2009 Proceedings of the 48th IEEE Conference on*, pages 4322–4327. IEEE, 2009.

[19] A. Vandendorpe and P. Van Dooren. Model reduction of interconnected systems. In *Model Order Reduction: Theory, Research Aspects and Applications*, pages 305–321. Springer, 2008.