Average State Observers for Large-Scale Network Systems

Tomonori Sadamoto, *Member, IEEE,* Takayuki Ishizaki, *Member, IEEE,* and Jun-ichi Imura, *Member, IEEE*

***Abstract*—In this paper, we propose a novel observer, called the *average state observer*, for large-scale network systems. This observer estimates the average behavior of the system with an estimation error assurance. To design an average state observer with explicit consideration of the estimation error, we first derive a tractable representation of the estimation error system. On the basis of this representation, we provide a theoretical upper bound of the estimation error for the average state observer. As a result, a systematic procedure to design the average state observers with an estimation error assurance is presented. Finally, we show the efficiency of the average state observer through an example of spatially discretized thermal diffusion networks.**

***Index Terms*—Large-scale systems, Observer design, Model reduction**

I. Introduction

Dynamical systems evolving over large-scale networks appear in engineering and nature, e.g., power networks [1], [2], transportation networks [3], biological networks [4], and spatially discretized meteorological networks [5], [6]. For analysis and control of these networks, various kinds of estimation/prediction techniques have been extensively developed. Examples involve energy management systems (EMS), where computationally efficient estimation/prediction of the behavior of large-scale power networks plays an important role in the improvement of EMS control and operation: see, e.g., [7]. Thus, as in this example, the development of a computationally efficient estimator design method is critical in handling large-scale network systems.

Related to the computationally efficient estimator design, methods based on model reduction techniques for designing low-dimensional observers have been presented. For example in [8], a Luenberger-type low-dimensional observer is designed for a low-dimensional model obtained by the balanced truncation [9], which does not preserve the network structure of the original system. Furthermore, since the relation between the model reduction error and the estimation error is not clear, the construction of low-dimensional observers is rather heuristic. In [10], the authors proposed a low-dimensional observer design method by reducing an original observer having desirable performance such that the performance degradation is small. However, the estimation error for the resultant low-dimensional observer is not explicitly taken into account.

From another viewpoint, distributed estimator design is also gaining attention: see, e.g., [11], [12], [13], [14], [15]. For example, a design method of a distributed Kalman filter consisting of local Kalman filters is presented in [11]. In that work, the information matrix of a centralized Kalman filter, defined as the inverse of covariance matrices, is approximated by smaller information matrices, which yield local Kalman filters for individual subsystems. Early work on distributed estimation was shown in [12], where a set of decentralized observers is constructed for individual subsystems. Indeed, these methods can produce distributed estimators for network systems in a systematic manner. However, they require a priori knowledge of decomposition of the whole system into subsystems, i.e., the set of clusters is assumed to be available in advance. In practical applications, this kind of subsystem decomposition, e.g., coherent generator groups in power networks, is not always given clearly. Therefore, it is crucial to develop an estimator design method that can systematically find a set of subsystem clusters.

A cluster construction method with explicit consideration on the dynamics of network systems has been developed in [16], [17], called clustered model reduction. In this method, using the controllability gramian of the original system, we consider clustering nodes (subsystems) that have similar behavior for input signals. By aggregating the state of clustered nodes into lower-dimensional ones, we obtain an aggregated model that preserves the network structure among the clusters.

On the basis of this clustered model reduction, this paper proposes a novel type of observers that perform the projective state estimation of each cluster for large-scale networks. In fact, large-scale network systems often involve a number of states (nodes) having similar behavior for input signals; see Section II-A for a motivating example. This finding suggests us a possibility to construct a type of observers that can efficiently capture macroscopic system behavior as the average of clustered states having similar behavior. To design such an observer with an estimation error assurance, called *average state observer*, we first derive a tractable representation of the estimation error system, which provides a clear insight into deriving an upper bound of the estimation error by the average state observer. Based on this result, we next propose a systematic procedure for designing the average state observer that satisfies an estimation error bound in terms of the $\mathcal{H}_2$ or $\mathcal{H}_\infty$-norm. The main advantages of the proposed design method are twofold: a relation between the state clustering and the estimation error is theoretically clarified, and a set of

T. Sadamoto, T. Ishizaki and J. Imura are with Department of Mechanical and Environmental Informatics, Graduate School of Information Science and Engineering, Tokyo Institute of Technology; 2-12-1, Meguro, Tokyo, Japan: {sadamoto@cyb., ishizaki@, imura@}mei.titech.ac.jp

T. Sadamoto is Research Fellow of the Japan Society for the Promotion of Science.

All authors are with Japan Science and Technology Agency, CREST 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan

clusters can be systematically constructed while the dynamics of the original network system is taken into account.

The paper [18], which is a preliminary version of this paper, has proposed the framework of low-dimensional observers, including average state observers as a special case. However, [18] does not provide explicit design procedure specialized for average state observers and not discuss the network structure of average state observers to be designed. In contrast, this paper focuses on average state observers and presents a method for designing average state observers with an estimation error assurance. Moreover, in this paper, we consider a thermal diffusion network system as an example, and numerically investigate the efficiency of the average state observers, discuss the network structure of the average state observer from the viewpoint of the dynamical property of thermal diffusion networks, and investigate the trade-off relation between the number of average states and the estimation performance.

This paper is organized as follows. In Section II-A, we first introduce an example to explain why we consider the average state observer design for network systems. In Section II-B, we briefly review the clustered model reduction in [16], [17]. In Section II-C, we propose an average state observer on the basis of this model reduction technique. Next, we formulate an average state observer design problem. In Section III-A, we describe a road map for the systematic design of average state observers and provide an estimation error bound. On the basis of this result, an average state observer design algorithm is provided in Section III-B. In Section IV, we show the efficiency of the average state observer through the numerical example of a thermal diffusion network system. Finally, concluding remarks are provided in Section V.

Notation The following notation is used in this paper:

$\mathbb{R}$	set of real numbers
I_n	n-dimensional identity matrix
e_i^n	the ith column of I_n
$\mathbf{1}_n$	n-dimensional column vector whose each element is 1
$e_{\mathcal{I}}^n$	$e_{\mathcal{I}}^n := [e_{i_1}^n, \ldots, e_{i_m}^n]$ for $i \in \mathcal{I} := \{i_1, \ldots, i_m\}$
$M \prec 0_n$ ($M \succ 0_n$)	negative (positive) definiteness of a symmetric matrix $M \in \mathbb{R}^{n\times n}$
$\mathrm{diag}(M_1, \ldots, M_n)$	block-diagonal matrix having matrices $M_1 \cdots M_n$ on its diagonal blocks
$\mathrm{tr}(M)$	trace of a matrix M
$\|M\|_F$	the Frobenius norm of a matrix M

The $\mathcal{L}_2$-norm of a square integrable function $v(t) \in \mathbb{R}^n$ is defined by

$$\|v(t)\|_{\mathcal{L}_2} := \left(\int_0^\infty v^{\mathsf{T}}(t)v(t)dt\right)^{\frac{1}{2}}.$$

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 respectively defined by

$$\|G(s)\|_{\mathcal{H}_\infty} := \sup_{\omega\in\mathbb{R}} \|G(j\omega)\|,$$
$$\|G(s)\|_{\mathcal{H}_2} := \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{tr}(G(j\omega)G^{\mathsf{T}}(-j\omega))d\omega\right)^{\frac{1}{2}}$$

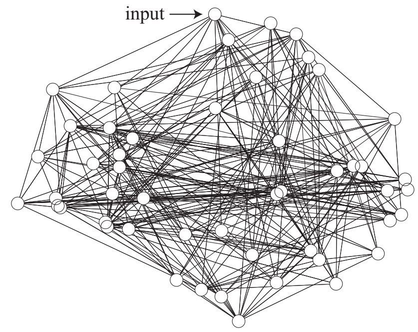


Fig. 1. Network system composed of 50 scalar subsystems.

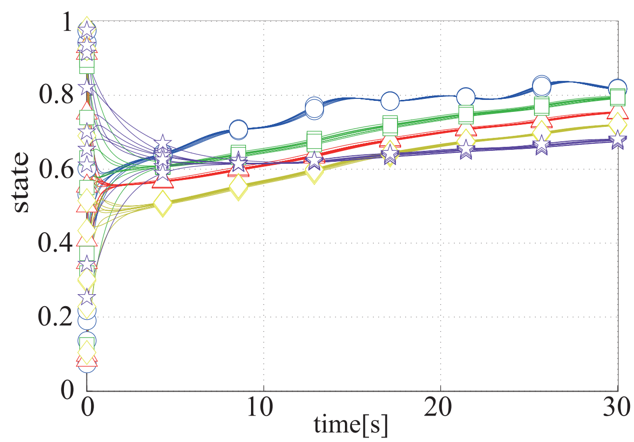


Fig. 2. Transient responses of the network system as shown in Fig. 1

where $\|\cdot\|$ denotes the induced 2-norm.

II. Problem Formulation

A. Motivating Example of Average State Observer Design

In this paper, we deal with a linear dynamical system composed of n subsystems, each of which is supposed to be a one-dimensional system for simplicity. For each $i \in \{1, \ldots, n\}$, the dynamics of the ith subsystem are described by

$$\dot{x}_i = a_{i,i}x_i + \sum_{j\neq i} a_{i,j}x_j + \sum_{k=1}^{m_u} b_{i,k}u_k \tag{1}$$

where $x_i \in \mathbb{R}$ denotes the state and $u_k \in \mathbb{R}$ denotes the kth external input signal for $k \in \{1, \ldots, m_u\}$. In (1), $a_{i,j} = 0$ if the ith subsystem is not instantaneously affected from the jth subsystem.

To explain why we consider the average state observer design for network systems, we introduce the following example. Let us consider a network system in (1) with $n = 50$ and $m_u = 1$. We show the schematic depiction of the network system in Fig. 1, where the circles represent the subsystems. Furthermore, in Fig. 2, we show the transient responses of the all 50 subsystems for a random initial state and input signal. Even though 50 trajectories of the state variables are in this figure, we can see that they are aggregated into five clusters at around $t = 30$, denoted by the lines with circles, squares, triangles, diamonds, and stars. In view of this, we consider a novel low-dimensional observer called *average state observer* for network systems, which estimates only the average of the state variables showing similar behavior to capture the average behavior of the entire network system.

To design the average state observer, in the next subsection, we briefly review the clustered model reduction proposed in [16], [17], where an approximated network model representing the dynamics of the average behavior of the system of interest is provided.

B. Brief Review of Clustered Model Reduction

Let us consider a network system in (1). Using the notation of

$$x = [x_1, \ldots, x_n]^{\mathsf{T}}, \quad u = [u_1, \ldots, u_{m_u}]^{\mathsf{T}}$$

and

$$A := \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix}, \quad B := \begin{bmatrix} b_{1,1} & \cdots & b_{1,m_u} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \cdots & b_{n,m_u} \end{bmatrix},$$

we give the whole network system as in

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

For simplicity, this system is supposed to be stable, i.e., A is a Hurwitz matrix.

Given a network system Σ in (2), we construct a reduced network model approximating the transfer function of Σ from u to x while preserving the network structure of Σ. More specifically, we aggregate the state variables having similar behavior to construct a reduced network model. In what follows, we focus on the behavior of Σ for the input signal u. In view of this, without loss of generality, we assume that $x(0) = 0$.

We first introduce the following notion of *aggregation matrix*:

Definition 1: Given Σ in (2), the family of an index set $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$ for $\mathbb{L} := \{1, \ldots, L\}$ is called a cluster set, each of whose elements is referred to as a cluster if each element $\mathcal{I}_l$ is a disjoint set of $\{1, \ldots, n\}$ such that

$$\bigcup_{l\in\mathbb{L}} \mathcal{I}_l = \{1, \ldots, n\}. \tag{3}$$

Furthermore, define an aggregation matrix compatible with $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$ as in

$$P := \operatorname{diag}\left(\frac{1}{n_1}\mathbf{1}_{n_1}^{\mathsf{T}}, \ldots, \frac{1}{n_L}\mathbf{1}_{n_L}^{\mathsf{T}}\right)\Pi \in \mathbb{R}^{L\times n} \tag{4}$$

where $n_l := |\mathcal{I}_l|$ and $\Pi \in \mathbb{R}^{n\times n}$ is the permutation matrix, i.e.,

$$\Pi := [e^n_{\mathcal{I}_1}, \ldots, e^n_{\mathcal{I}_L}]^{\mathsf{T}} \in \mathbb{R}^{n\times n}. \tag{5}$$

The matrix P in (4) plays a role of aggregation of the states indicated by a cluster set $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$. For the understanding of the structure of P, we provide an example of aggregation matrices as follows. Given Σ in (2) with $n = 5$, consider a cluster set given by

$$\mathcal{I}_1 = \{1\}, \quad \mathcal{I}_2 = \{2, 5\}, \quad \mathcal{I}_3 = \{3, 4\}.$$

The aggregation matrix compatible with this cluster set is given by

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0.5 & 0.5 & 0 \end{bmatrix}.$$

Using the notion of this aggregation matrix, we define the following aggregated network system, where the lth element of its state variable represents the aggregation (average) of a set of states $\{x_i\}_{i\in\mathcal{I}_l}$:

Definition 2: Given Σ in (2), consider a cluster set $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$ and the compatible aggregation matrix P in Definition 1. Define an aggregated network model $\hat{\Sigma}$ by

$$\hat{\Sigma} : \ \dot{\hat{x}} = PAP^{\dagger}\hat{x} + PBu \tag{6}$$

where $P^{\dagger} \in \mathbb{R}^{n\times L}$ is the pseudoinverse of P, i.e.,

$$P^{\dagger} = \Pi^{\mathsf{T}} \operatorname{diag}(\mathbf{1}_{n_1}, \ldots, \mathbf{1}_{n_L}). \tag{7}$$

Note that the aggregated network model $\hat{\Sigma}$ is determined if a cluster set $\{\mathcal{I}_l\}$ is determined. In [16], [17], the authors have present a method to find a cluster set such that the transfer function from u to the lth average state $\hat{x}_l$ approximates that from u to x_i for all $i \in \mathcal{I}_l$ and $l \in \mathbb{L}$.

Remark 1: In Definition 1, the aggregation matrix P in (4) is slightly different from that in [16], [17], where each element of the block-diagonal matrix in (4) is normalized by $\sqrt{n_l}$, but not n_l. In this paper, we use P defined in (4), i.e., that normalized by n_l, because it is useful in understanding the concept of the average state observer introduced in Section II-C.

C. Average State Observer Design Problem

On the basis of the clustered model reduction in Section II-B, in this paper, we propose an observer called the *average state observer* which estimates the average states of the original network system Σ in (2).

Note that the system Σ is assumed to be stable and to have a zero initial state. One may think that the observer design for such a system is trivial. However, the estimation error depends not only on the initial state, but also on the input signal (see Lemma 1 in Section III-A). Furthermore, the average state observer introduced below is not necessarily stable in general even though Σ is stable. Thus, it is not trivial to design an average state observer for Σ satisfying the above assumptions. One extension to the system with a nonzero initial state is described in Remark 4.

Given a stable Σ in (2) with a zero initial state, define the measurement output signal by

$$y = Cx + Du \tag{8}$$

where $y \in \mathbb{R}^{m_y}$, and define the average state observers as follows:

Definition 3: Given Σ in (2) with y in (8), consider a cluster set $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$ and the compatible aggregation matrix P in Definition 1. Define an *average state observer* $\hat{O}$ by

$$\hat{O} : \begin{cases} \dot{\hat{\xi}} = PAP^{\dagger}\hat{\xi} + PBu + H(y - \hat{y}) \\ \hat{y} = CP^{\dagger}\hat{\xi} + Du \end{cases} \tag{9}$$

with $\hat{\xi}(0) = 0$ for simplicity, where $H \in \mathbb{R}^{L\times m_y}$ is an observer gain.

One of the remarkable points we have to note here is that y in (9) is *not* a measurement output signal of the aggregated network model $\hat{\Sigma}$, but of the original network system Σ. Thus, the average state observer $\hat{O}$ is *not* a Luenberger observer for the aggregated network model $\hat{\Sigma}$.

Next, we formulate a problem for designing $\hat{O}$ in (9), i.e., designing H and P (or equivalently, $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$), such that the observer estimates the average behavior of Σ in (2), i.e., $\hat{\xi}_l(t)$ captures all $\{x_i(t)\}_{i\in\mathcal{I}_l}$ for any $l \in \mathbb{L}$. To this end, we define the estimation error of $\hat{O}$ by

$$\delta := x - P^\dagger\hat{\xi} \tag{10}$$

where $P^\dagger$ is defined in (7), and we quantify the magnitude of the estimation error by the $\mathcal{H}_2$-norm of the transfer function from u to δ, denoted by $\Delta(s)$, for simplicity. In this setting, we consider the following average state observer problem as follows:

Problem 1: Given a positive constant $\epsilon \geq 0$ and a stable Σ in (2) with a zero initial state and y in (8), find $\hat{O}$ in (9) satisfying

$$\|\Delta(s)\|_{\mathcal{H}_2} \leq \epsilon. \tag{11}$$

III. Average State Observer Design

A. A Road Map for Average State Observer Design

To solve Problem 1 in the previous section, we provide a tractable representation of the estimation error δ in (10) as follows:

Lemma 1: Consider Σ in (2) with y in (8) and define $\hat{O}$ in (9). Then, δ in (10) obeys

$$\mathcal{E} : \begin{cases} \dot{\mathcal{X}} = \mathcal{A}\mathcal{X} + \mathcal{B}u \\ \delta = \mathcal{P}\mathcal{X} \end{cases} \tag{12}$$

with

$$\mathcal{A} := \begin{bmatrix} PAP^\dagger - HCP^\dagger & (PA - HC)\overline{P}^\dagger\overline{P} \\ 0 & A \end{bmatrix}$$
$$\mathcal{B} := \begin{bmatrix} 0 \\ B \end{bmatrix}, \quad \mathcal{P} := \begin{bmatrix} P^\dagger & \overline{P}^\dagger\overline{P} \end{bmatrix}$$

where $\overline{P} \in \mathbb{R}^{(n-L)\times n}$ and $\overline{P}^\dagger \in \mathbb{R}^{n\times(n-L)}$ satisfy

$$P^\dagger P + \overline{P}^\dagger\overline{P} = I_n, \quad \overline{P}\,\overline{P}^\dagger = I_{n-L} \tag{13}$$

for P in (4).

Proof : Define $\hat{\mathcal{X}} := [\hat{\xi}^\mathsf{T}\ x^\mathsf{T}]^\mathsf{T}$. Then, we have

$$\begin{cases} \dot{\hat{\mathcal{X}}} = \hat{\mathcal{A}}\hat{\mathcal{X}} + \hat{\mathcal{B}}u \\ \delta = \hat{\mathcal{P}}\hat{\mathcal{X}} \end{cases}$$

where

$$\hat{\mathcal{A}} := \begin{bmatrix} PAP^\dagger - HCP^\dagger & HC \\ 0 & A \end{bmatrix}, \quad \hat{\mathcal{B}} := \begin{bmatrix} PB \\ B \end{bmatrix}$$
$$\hat{\mathcal{P}} := [-P^\dagger \quad I_n].$$

Taking a coordinate transformation by

$$T := \begin{bmatrix} -I_L & P \\ 0 & I_n \end{bmatrix} = T^{-1}$$

as $T\hat{\mathcal{A}}T^{-1}, T\hat{\mathcal{B}}$ and $\hat{\mathcal{P}}T^{-1}$, we have the claim. □

In Lemma 1, the error system $\mathcal{E}$ in (12) is the generalization of that for a Luenberger observer, i.e.,

$$\dot{\delta} = (A - HC)\delta.$$

In fact, if $\mathcal{I}_l = \{l\}$, i.e., $P = P^\dagger = I_n$, then $\hat{O}$ in (9) becomes a Luenberger observer and we have

$$\mathcal{A} = \begin{bmatrix} A - HC & 0 \\ 0 & A \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} 0 \\ B \end{bmatrix}, \quad \mathcal{P} = [I_n \quad 0].$$

Thus, δ in this case is independent of any input signals. In contrast, in the average state observer design, the estimation error δ depends on input signals u because the average state observer cannot capture the exact behavior of the system for u.

Note that $\overline{P}^\dagger\overline{P}$ in the error system $\mathcal{E}$ is an orthogonal projection matrix since $\overline{P}^\dagger\overline{P}$ is a complementary of the orthogonal projection matrix $P^\dagger P$. Even though the orthogonal projection matrix $\overline{P}^\dagger\overline{P}$ offers a degree of freedom in choosing its basis, without loss of generality we suppose $\overline{P}^\dagger\overline{P} = \overline{P}^\mathsf{T}\overline{P}$.

On the basis of this error analysis, let us consider the systematic design of $\hat{O}$ in (9). The transfer function of the estimation error system in (12) from u to δ can be written as

$$\Delta(s) = \Xi_{P,H}(s)U_P(s) \tag{14}$$

where

$$\begin{aligned} \Xi_{P,H}(s) &:= P^\dagger(sI_L - A_\Xi)^{-1}(PA - HC)\overline{P}^\mathsf{T} + \overline{P}^\mathsf{T} \\ A_\Xi &:= PAP^\dagger - HCP^\dagger \end{aligned} \tag{15}$$

and

$$U_P(s) := \overline{P}(sI_n - A)^{-1}B. \tag{16}$$

Compared to the clustered model reduction, in the average state observer design, we are required to design not only P, but also H such that $\|\Delta(s)\|_{\mathcal{H}_2}$ is small. However, Δ depends on the design parameters P and H in a nonlinear fashion. Thus, the simultaneous design of P and H making $\|\Delta(s)\|_{\mathcal{H}_2}$ small is difficult.

To overcome this difficulty, we utilize the following facts:

- Parameter H does not affect U_P, but $\Xi_{P,H}$ only.
- U_P depends on P (equivalently $\overline{P}$), not on H.
- It follows that

$$\|\Delta(s)\|_{\mathcal{H}_2} \leq \|\Xi_{P,H}(s)\|_{\mathcal{H}_\infty}\|U_P(s)\|_{\mathcal{H}_2}.$$

On the basis of these facts, we propose a road map for designing P and H as follows. First, we determine P to make $\|U_P(s)\|_{\mathcal{H}_2}$ small. Next, for a fixed P, we determine H to make $\|\Xi_{P,H}(s)\|_{\mathcal{H}_\infty}$ small.

Regarding the first step, we introduce the following lemma from [16], [17], which is useful in determining P (or equivalently, $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$).

Lemma 2: Consider a stable Σ in (2). Let $\Phi \succeq 0_n$ be given satisfying

$$A\Phi + \Phi A^{\mathsf{T}} + BB^{\mathsf{T}} = 0. \tag{17}$$

Define $\Phi_{\frac{1}{2}} \in \mathbb{R}^{n\times n}$ such that $\Phi = \Phi_{\frac{1}{2}}\Phi_{\frac{1}{2}}^{\mathsf{T}}$. Given $\theta \geq 0$, suppose that there exist $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$ in Definition 1 and $\phi_l \in \mathbb{R}^{1\times n}$ such that

$$\left\|(e^n_{\mathcal{I}_l})^{\mathsf{T}}\Phi_{\frac{1}{2}} - \frac{1}{\sqrt{n_l}}\mathbf{1}_{n_l}\phi_l\right\|_F \leq \sqrt{n_l}\theta, \quad l \in \mathbb{L}. \tag{18}$$

Then, $U_P(s)$ in (16) satisfies

$$\|U_P(s)\|_{\mathcal{H}_2} \leq \kappa\theta \tag{19}$$

where

$$\kappa := \left(\sum_{l=1}^{L} n_l(n_l-1)\right)^{\frac{1}{2}}. \tag{20}$$

In this lemma, θ is a design parameter for regulating coarseness for constructing a cluster set $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$. This lemma shows that the norm of U_P can be bounded by using the coarseness parameter θ.

Next, for a given P under the assumptions in Lemma 2, we consider determining H to make $\|\Xi_{P,H}(s)\|_{\mathcal{H}_\infty}$ small. To this end, we give the following theorem:

Theorem 1: Consider Problem 1. Suppose that there exist $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$ and $\phi_l \in \mathbb{R}^{1\times n}$ satisfying (18). Give P in (4). If $(PAP^\dagger, CP^\dagger)$ is detectable, then there exist

$$\gamma > 0, \quad X \succ 0_L, \quad Y \in \mathbb{R}^{L\times m_y}$$

satisfying

$$\begin{bmatrix} \mathrm{sym}(XPAP^\dagger - YCP^\dagger) + (P^\dagger)^{\mathsf{T}}P^\dagger & * \\ \overline{P}A^{\mathsf{T}}P^{\mathsf{T}}X - \overline{P}C^{\mathsf{T}}Y^{\mathsf{T}} & (1-\gamma^2)I_{n-L} \end{bmatrix} \prec 0_n \tag{21}$$

where $\mathrm{sym}(M) := M + M^{\mathsf{T}}$. Furthermore, $\hat{O}$ in (9) with

$$H = X^{-1}Y \tag{22}$$

satisfies

$$\|\Delta(s)\|_{\mathcal{H}_2} < \gamma\kappa\theta \tag{23}$$

with κ in (20) where Δ is defined as the transfer function from u to δ in (10).

Proof : Note that there exists an observer gain H stabilizing $\Xi_{P,H}(s)$ in (15) because $(PAP^\dagger, CP^\dagger)$ is detectable. Thus, there exists γ such that

$$\|\Xi_{P,H}(s)\|_{\mathcal{H}_\infty} < \gamma. \tag{24}$$

Furthermore, it follows from the bounded real lemma [19] and the Schur complement [19] that (24) is equivalent to (21) and (22). Hence, there exist $\gamma > 0$, $X \succ 0_L$, $Y \in \mathbb{R}^{L\times m_y}$ satisfying (21) and (22). Furthermore, combining Lemma 2, (23) follows. □

In this theorem, LMI in (21) is used for determination of H. Furthermore, this theorem provides an upper bound of the estimation error caused by an average state observer designed along with the proposed road map.

Remark 2: The detectability of the pair $(PAP^\dagger, CP^\dagger)$ depends on C, which is not considered in the cluster construction, for which we use the controllability gramian of the original network system. Therefore, the detectability is not always guaranteed in general. One remedy for this is to find P such that $PAP^\dagger$ is Hurwitz, which is sufficient for the detectability. In fact, the stability of $PAP^\dagger$ can be ensured by the clustered model reduction method for diagonally stable systems, i.e. the class of systems admitting a diagonal Lyapunov function, including the bidirectional networks and the positive networks; see [16] and [17] for details.

Remark 3: In this paper, we evaluate the magnitude of $\Delta(s)$ by the $\mathcal{H}_2$-norm and show an upper bound of $\|\Delta(s)\|_{\mathcal{H}_2}$. A similar result based on $\mathcal{H}_\infty$-norm evaluation is also available by constructing a cluster set on the basis of the Hessenberg transformation: see [16] in detail.

Remark 4: We have shown a fundamental result in average state observer design for a network system with a zero initial state. This result can be extended to the case where the system has an unknown nonzero initial state as follows.

Even though the initial state of Σ in (2) is unknown in general, the existing range of the initial state is available in some cases, e.g., thermal diffusion processes starting near an equilibrium state enables us to estimate an existing range for the initial state of the process. In view of this, let us assume that an existing range of the initial state is available, i.e., $x(0)$ is assumed to satisfy

$$x(0)x(0)^{\mathsf{T}} \preceq Q \tag{25}$$

where $Q \succeq 0_n$ is an available bound of the existing range of $x(0)$. In the remainder of this remark, for a given Σ, we consider designing an average state observer $\hat{O}$ in (9) to make the estimation error δ in (10) small for any $x(0)$ satisfying (25).

The estimation error δ in (10) depends on $x(0)$ and u, which is denoted by $\delta(t; x(0), u)$. Since the dynamics of the error system are linear, $\delta(t; x(0), u)$ can be represented as the sum of the error factors caused by $x(0)$ and u, i.e.,

$$\delta(t; x(0), u) = \delta_x(t) + \delta_u(t)$$

where

$$\delta_x(t) := \delta(t; x(0), 0), \quad \delta_u(t) := \delta(t; 0, u).$$

With these settings, we show upper bounds of $\|\delta_x(t)\|_{\mathcal{L}_2}$ and $\|\delta_u(t)\|_{\mathcal{L}_2}$. Let $\Phi \succeq 0_n$ be given satisfying

$$A\Phi + \Phi A^{\mathsf{T}} + BB^{\mathsf{T}} + Q = 0 \tag{26}$$

instead of (17). Suppose the assumptions in Theorem 1 hold. Let $\hat{O}$ in (9) be given with H in (22). Then, δ_x satisfies

$$\|\delta_x(t)\|_{\mathcal{L}_2} < \gamma\kappa\theta + \mathrm{tr}^{\frac{1}{2}}(P^{\mathsf{T}}XPQ) \tag{27}$$

with κ in (20) for any $x(0)$ satisfying (25). Furthermore, δ_u satisfies

$$\|\delta_u(t)\|_{\mathcal{L}_2} < \gamma\kappa\theta \tag{28}$$

for any unit impulse input u.

Remark 5: The proposed method can be extended to the case of multi-dimensional subsystems. More specifically, for network systems where the state of each subsystem represents the same physical quantity, we can design average state observers to estimate the average of the states of subsystems that show similar behavior.

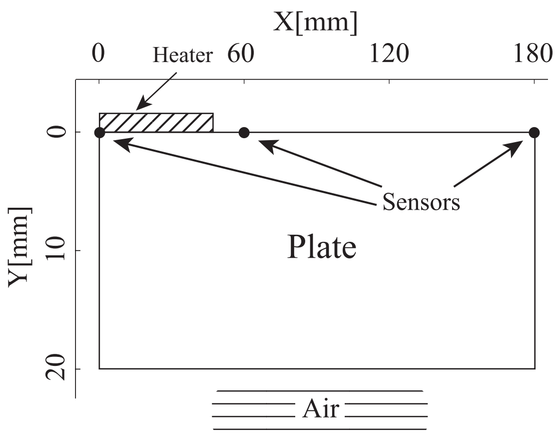


Fig. 3. Thermal diffusion system.

B. Design Algorithm

In this subsection, we present an algorithm for designing an average state observer. For the completeness of the algorithm, we first show the procedure in [16], [17] for constructing a cluster set $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$ satisfying (18) for a given $\theta \geq 0$ as follows.

Suppose that we have l clusters $\{\mathcal{I}_1,\ldots,\mathcal{I}_l\}$ and consider constructing $\mathcal{I}_{l+1}$. To this end, we define

$$\mathcal{J} := \{1,\ldots,n\}\backslash \bigcup_{k=1}^{l} \mathcal{I}_k. \tag{29}$$

First, choose $j \in \mathcal{J}$. Next, find

$$\mathcal{I}_{l+1} = \{i \in \mathcal{J}\backslash\{j\} |\ \left\|\phi_{[i]} - \phi_{[j]}\right\| \leq \theta\} \tag{30}$$

where $\phi_{[i]} \in \mathbb{R}^{1\times n}$ is the ith row vector of $\Phi_{\frac{1}{2}}$. Then, a new cluster $\mathcal{I}_{l+1}$ and $\phi_{l+1} = \sqrt{n_{l+1}}\phi_{[j]}$ satisfy

$$\left\|(e^n_{\mathcal{I}_{l+1}})^{\mathsf{T}}\Phi_{\frac{1}{2}} - \frac{1}{\sqrt{n_{l+1}}}\mathbf{1}_{n_{l+1}}\phi_{l+1}\right\|_F = \sum_{i\in\mathcal{I}_{l+1}} \|\phi_{[i]} - \phi_{[j]}\|.$$

Hence, $\mathcal{I}_{l+1}$ and ϕ_{l+1} satisfy (18).

Next, we summarize an algorithm to solve Problem 1 as follows:

1) Give $\theta \geq 0$.
2) Find $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$ along with the above procedure.
3) Find $X \succ 0_L$ and $Y \in \mathbb{R}^{L\times m_y}$ such that (21) is satisfied while minimizing $\gamma > 0$.
4) If no solutions exist, take a smaller θ, and go back to 2).
5) Construct $\hat{O}$ in (9) with H given by (22).
6) If $\|\Delta(s)\|_{\mathcal{H}_2} > \epsilon$, take a smaller θ, and go back to 2).

Finally, it should be noted that the number of decision variables of LMI in (21) is $\mathcal{O}(L^2)$. This implies that the proposed design algorithm is computationally tractable if L is small.

IV. Numerical Example

A. Spatially Discretized Thermal Diffusion Network

In this section, we show the efficiency of the proposed average state observer. We deal with a network system given by spatial discretization of a thermal diffusion system composed of a metal plate, a heater and the air as shown in Fig. 3.

TABLE I
PARAMETERS OF THERMAL DIFFUSION NETWORK.

Thermal diffusivity[mm^2/s]	λ	29.1
Thermal conductivity[W/(mmK)]	β	2.0×10^4
Heat transfer coef. to air[W/(mm^2K)]	h_a	2.0×10^7
Heat transfer coef. to heater[W/(mm^2K)]	h_h	2.6×10^8

In what follows, we take an XY orthogonal coordinate given by

$$(X,Y) \in \mathcal{D} := [0,180]\times[0,20]$$

as shown in Fig. 3. Let $T(X,Y,t)$ be the temperature of the metal plate at the position (X,Y) and time t. The heat transfer properties of the metal plate are described by a rectangular coordinate diffusion equation [20] as in

$$\frac{\partial T}{\partial t} = \lambda\left(\frac{\partial^2 T}{\partial X^2} + \frac{\partial^2 T}{\partial Y^2}\right), \quad (X,Y)\in \mathrm{int}(\mathcal{D}) \tag{31}$$

where $\mathrm{int}(\mathcal{D})$ is the inside of $\mathcal{D}$ and λ denotes a diffusion coefficient. In addition, the metal plate exchanges heat with the air and the heater at the boundary of $\mathcal{D}$ as follows.

The heat exchange with the air is described with the Neumann type boundary condition

$$\beta\frac{\partial T}{\partial \mathbf{n}} = h_\mathrm{a}(T - T_\mathrm{a}), \quad (X,Y)\in\mathcal{S}_\mathrm{a} \tag{32}$$

where $\mathcal{S}_\mathrm{a}$ is a set of contact points with the air, $\mathbf{n}$ is a unit vector normal to $\mathcal{S}_\mathrm{a}$, β is the coefficient of thermal conductivity, T_a is the temperature of the air and h_a is the coefficient of heat transfer between the air and the metal plate. For simplicity, we suppose $T_\mathrm{a}\equiv 0$ for any t, X and Y.

The heat budget for the heater is described by

$$\beta\frac{\partial T}{\partial Y} = h_\mathrm{h}(T-u), \quad X\in\mathcal{X},\ Y=0 \tag{33}$$

where $\mathcal{X}\subset[0,180]$ is the set of contact points with the heater over $Y=0$ and u is the temperature of the heater. The heater is assumed to have a uniform temperature distribution, i.e., u is independent of X and Y.

Finally, discretizing (31)-(33) with steps δ_X,δ_Y for X and Y axes by means of the finite volume method [21], we have a stable Σ in (2), where $x\in\mathbb{R}^n$ is a vector of spatially discretized temperature T. In addition, a measurement output signal $y\in\mathbb{R}^3$ is taken for the plate temperatures at the positions shown by circles in Fig. 3.

B. Demonstration of Average State Observer Design

In this subsection, we demonstrate the average state observer design for a spatially discretized thermal diffusion network with $\delta_X=\delta_Y=2$[mm], which is a lattice network system with $n=1001$. The parameters of the network system is summarized in Table I. In this demonstration, we evaluate the estimation error of the designed average state observer by the $\mathcal{H}_\infty$-norm to demonstrate the worst case scenario.

We first investigate the relation between the coarseness parameter θ in (18) and a cluster set $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$. To see this, we construct $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$ varying the parameter θ. More specifically,

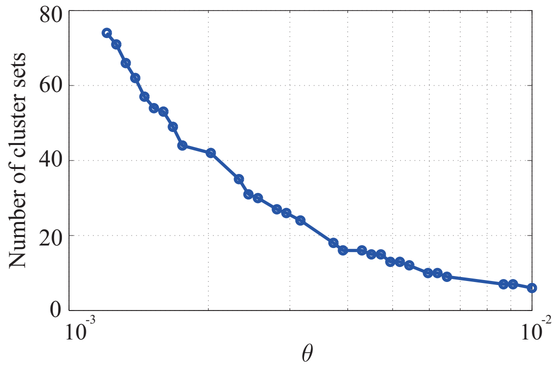


Fig. 4. The number of cluster sets versus θ.

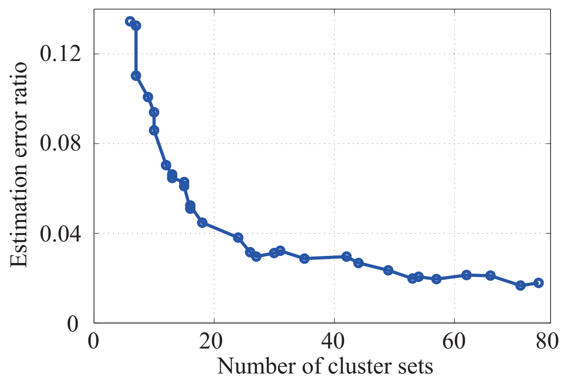


Fig. 5. Estimation error ratio with respect to the number of cluster sets.

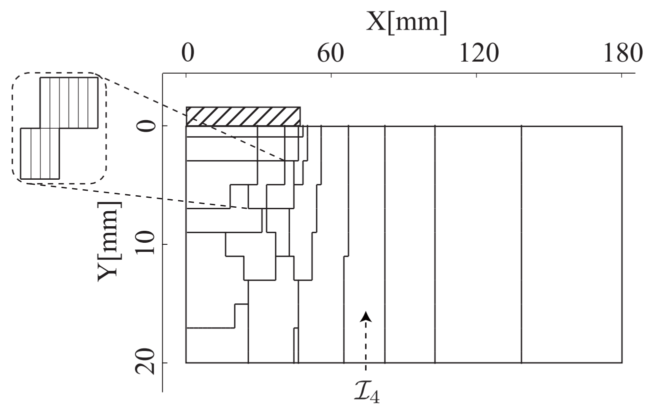


Fig. 6. Set of clusters in the case $L = 24$.

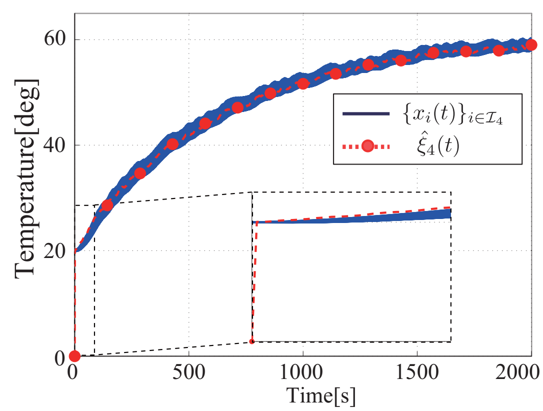


Fig. 7. Trajectories of $\{x_i(t)\}_{i\in\mathcal{I}_4}$ and $\hat{\xi}_4(t)$.

at Step 1) in the algorithm in Section III-B, we vary θ in the range $[10^{-2.9}, 10^{-2}]$. Next, at Step 2), we construct $\{\mathcal{I}_l\}_{l\in\mathbb{L}}$ with respect to each value of θ. In Fig. 4, we plot the number of resultant clusters L versus the value of θ. This figure shows that the number of clusters, which coincides with the dimension of the average state observer, decreases as θ increases.

Next, for each of resultant cluster set, we design average state observer $\hat{O}$ along with the Steps 3)-5). We investigate the relation between the dimension of $\hat{O}$ and the estimation performance of $\hat{O}$, which is quantified by the estimation error ratio $\|\Delta(s)\|_{\mathcal{H}_\infty}/\|X(s)\|_{\mathcal{H}_\infty}$ with $X(s)$ denoting the transfer function from u to x. In Fig. 5, we plot $\|\Delta(s)\|_{\mathcal{H}_\infty}/\|X(s)\|_{\mathcal{H}_\infty}$ versus the dimension of $\hat{O}$. This figure shows that the estimation error ratio decreases as the dimension of $\hat{O}$ increases. Furthermore, Figs. 4 and 5 imply that the estimation performance improves in compensation for the increase of the dimension of the average state observer.

Next, let $\epsilon = 0.09$ in Problem 1. Then, $\|\Delta(s)\|_{\mathcal{H}_2}$ by 24-dimensional $\hat{O}$ becomes 0.08, which is less than ϵ. In Fig. 6, we show the resultant cluster set $\{\mathcal{I}_l\}_{l\in\{1,\ldots,24\}}$. In this figure, there are 24 regions surrounded by lines, where each region represents a set of states belonging to a compatible cluster, e.g., the zoomed up region in this figure represents ten corresponding states (grids) of the original network system. Fig. 6 shows that regions far from the heater, e.g., those around $X \in [120, 180]$, are more roughly clusterized than those close to the heater, e.g., those around $X = 0$. This fact reflects the diffusive property of the system, i.e., the temperature distribution tends to become uniform further as being far from the heater. Thus, the cluster set is constructed with explicit consideration of the dynamical properties of the network system.

Next, we show the efficiency of the 24-dimensional average state observer by comparing the trajectories $x(t)$ in (2) and $\hat{\xi}(t)$ in (9). To see this, we take an input signal as $u(t) = 100 + 100\sin(t^2)$ so that it contains multiple frequency waves, and we take $x(0) = 20 \times \mathbf{1}_{1001}$ and $\hat{\xi}(0) = 0$. In Fig. 7, the blue solid lines and the red dotted line with circles depict $\{x_i(t)\}_{i\in\mathcal{I}_4}$ and $\hat{\xi}_4(t)$ where the cluster $\mathcal{I}_4$ is shown in Fig. 6. We omit the other trajectories since they behave similarly. We can see from Fig. 7 that the estimated signal $\hat{\xi}_4(t)$ stays near the center of the bundles of $\{x_i(t)\}_{i\in\mathcal{I}_4}$. In addition, we show snapshots of the temperature distribution of the original system, i.e., $x(t)$, in the left half of Fig. 8 and show those of the estimated average temperature distributions, i.e., $P\hat{\xi}(t)$, in the right half of Fig. 8. In addition to Figs. 7 and 8, the fact that the resultant estimation error ratio is 0.038 implies that the resultant average state observer captures the average behavior of the original network system.

As shown in Fig. 6, the fine grids of the 1001-dimensional model are clusterized while taking into account the dynamics of the system. To investigate the effectiveness of this cluster construction, we compare the case where a cluster set is given

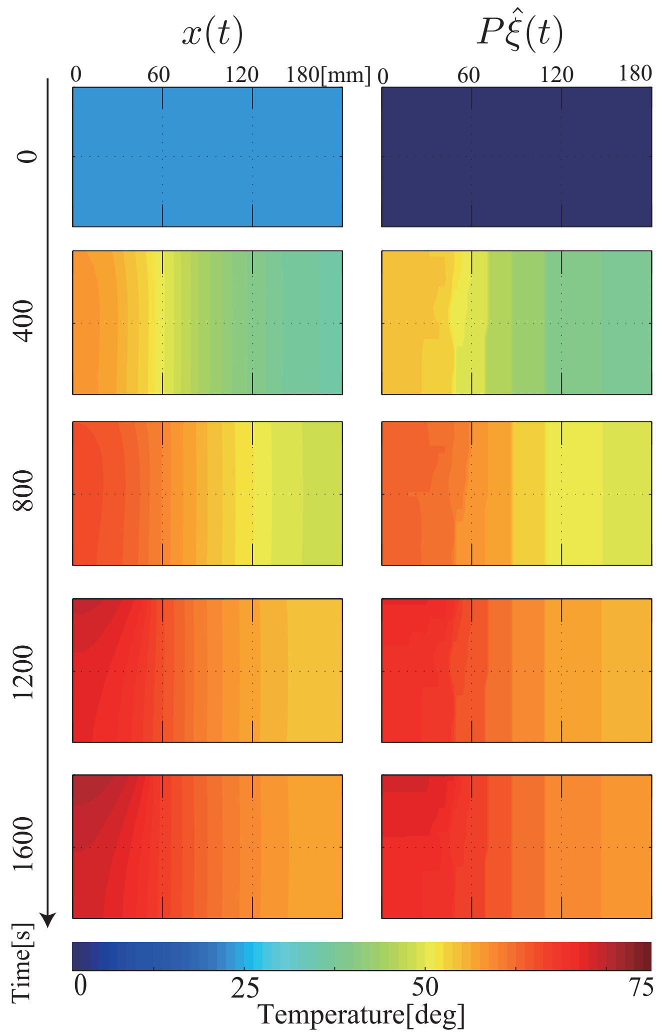


Fig. 8. The entire system behavior $x(t)$ and estimated average behavior $P\hat{\xi}(t)$.

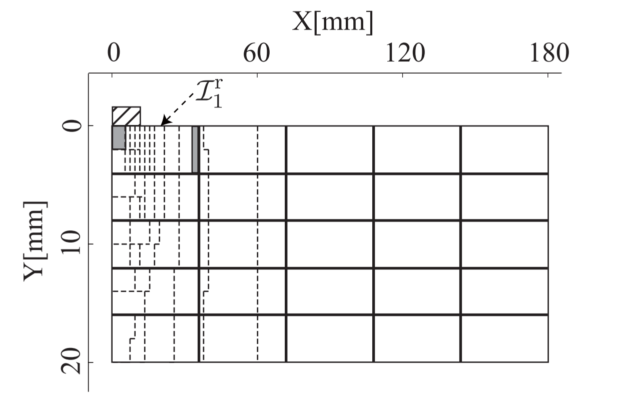


Fig. 9. Set of clusters $\{\mathcal{I}_l^{\mathrm{r}}\}_{l\in\{1,\ldots,25\}}$ compatible with the spatial discretization with the steps $\delta_X = 36$[mm] and $\delta_Y = 4$[mm].

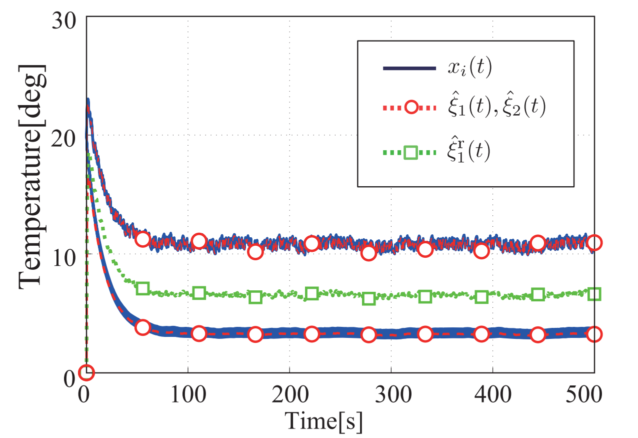


Fig. 10. Trajectories of $x_i(t)$ belonging to the filled region in Fig. 9 and the estimated signals by $\hat{O}$ and $\hat{O}^{\mathrm{r}}$.

a priori without considering the system dynamics.

Let us slightly modify the thermal diffusion system (31)-(33) such that the width of the heater is 10[mm], the coefficient of the heat transfer to air h_{a} is 2.0×10^9, the coefficient of the thermal diffusivity λ is 58.3 and three sensors are added at the points $(X,Y) = (10,0)$, $(20,0)$ and $(30,0)$. Constructing A, B and C by the spatial discretization of this thermal diffusion system, we make another cluster set $\{\mathcal{I}_l^{\mathrm{r}}\}_{l\in\{1,\ldots,25\}}$ compatible with the uniform discretization with the steps $\delta_X = 36$[mm] and $\delta_Y = 4$[mm] as in Fig. 9. For the cluster set, we construct a 25-dimensional observer in (9) whose gain is obtained by solving the LMI in (22). We denote the observer by $\hat{O}^{\mathrm{r}}$ and the state of the observer by $\hat{\xi}^{\mathrm{r}}$.

In Fig. 10, the green dotted line with the squares depicts the first element of $\hat{\xi}^{\mathrm{r}}$, denoted by $\hat{\xi}_1^{\mathrm{r}}$, which represents an estimated average temperature of the region compatible with $\mathcal{I}_1^{\mathrm{r}}$. Furthermore, the blue solid lines in Fig. 10 depict the trajectories of the state corresponding to the two gray-colored regions in Fig. 9. We can see from Fig. 10 that the two bundles of trajectories are distinct from each other. As a result, the $\mathcal{H}_2$-norm of the estimation error is 0.36. Next, we construct a 25-dimensional average state observer $\hat{O}$ in (9) whose cluster set is shown in Fig. 9 as the regions surrounded by the dotted lines. In Fig. 10, the red lines with the circles depict the estimated signals corresponding to the two gray-colored regions in Fig. 9, denoted by $\hat{\xi}_1$ and $\hat{\xi}_2$. We can see from Fig. 10 that each of the estimated signals stays near the center of the corresponding bundle. In this case, the $\mathcal{H}_2$-norm of the estimation error is 0.15. These results demonstrate that the explicit consideration of the system dynamics for cluster construction improves the performance of the average state estimation.

V. Conclusion

In this paper, we have proposed a novel observer called the *average state observer* that can capture the average behavior of large-scale network systems with an estimation error assurance. To design an average state observer with explicit consideration of the estimation error, we have derived a tractable representation of estimation error systems. On the basis of this representation, we have provided a systematic procedure to design the average state observer with an upper bound of the estimation error. The proposed design procedure offers the advantage that we can systematically determine the states showing similar behavior, even though it is nontrivial to find the states that show the average behavior of the system in general. Finally, we have shown the effectiveness of the average state observer through the example of spatially discretized thermal diffusion networks.

Interesting examples of the application of average state estimation include power flow estimation/prediction of large-scale power grids. Since power flow estimation/prediction at low computational cost is a key tool for improving power grid operation, this research is expected to form a fundamental theory of estimation/prediction for large-scale network systems. However, for practical applications to real large-scale network systems, the extension of the average state observers to Kalman filter cases is one of the critical issues. As a first step towards this extension, in [22] we have proposed average state Kalman filters where measurement and system noise are taken into explicit consideration. Application research along this line will be among future topics of study.

ACKNOWLEDGMENT

This work was supported by CREST, JST. This work was also supported by Grant-in-Aid for JSPS Fellows Grant Number 14J11344.

REFERENCES

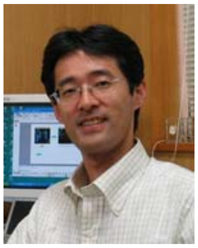
[1] S. M. Amin and B. F. Wollenberg, "Toward a smart grid: power delivery for the 21st century," *Power and Energy Magazine, IEEE*, vol. 3, no. 5, pp. 34–41, 2005.
[2] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, 2001.
[3] M. G. H. Bell and Y. Iida, *Transportation network analysis*. Wiley, 1997.
[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, "Complex networks: Structure and dynamics," *Physics reports*, vol. 424, no. 4, pp. 175–308, 2006.
[5] G. Evensen, *Data assimilation*. Springer, 2007.
[6] M. Verlaan, "Efficient kalman filtering algorithms for hydrodynamic models," *Ph.D. Thesis at Technische Universiteit, Delft, The Netherlands*, 1998.
[7] K.-R. Shih and S.-J. Huang, "Application of a robust algorithm for dynamic state estimation of a power system," *IEEE Transactions on Power Systems*, vol. 17, no. 1, pp. 141–147, 2002.
[8] A. K. Singh and J. Hahn, "State estimation for high-dimensional chemical processes," *Computers & chemical engineering*, vol. 29, no. 11, pp. 2326–2334, 2005.
[9] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2005.
[10] T. Sadamoto, T. Ishizaki, and J. Imura, "Low-dimensional functional observer design for linear systems via observer reduction approach," *in Proc. of the 52th Conference on Decision and Control*, pp. 776–781, 2013.
[11] U. Khan and J. M. F. Moura, "Distributing the kalman filter for large-scale systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4919–4935, 2008.
[12] D. Siljak and M. Vukcevic, "Decentralization, stabilization, and estimation of large-scale linear systems," *IEEE Transactions on Automatic Control*, vol. 21, no. 3, pp. 363–366, 1976.
[13] K. Martensson and A. Rantzer, "A scalable modularized synthesis method for distributed kalman filters," *in Proc. of the 18th IFAC World Congress*, pp. 9133–9138, 2011.
[14] M. K. Sundareshan and P. C. Huang, "On the design of a decentralized observation scheme for large-scale systems," *IEEE Transactions on Automatic Control*, vol. 29, no. 3, pp. 274–276, 1984.
[15] M. K. Sundareshan and R. M. Elbanna, "Design of decentralized observation schemes for large-scale interconnected systems: Some new results," *Automatica*, vol. 26, no. 4, pp. 789–796, 1990.
[16] T. Ishizaki, K. Kashima, J. Imura, and K. Aihara, "Model reduction and clusterization of large-scale bidirectional networks," *IEEE Transactions on Automatic Control*, vol. 59, no. 1, pp. 48–63, 2014.
[17] T. Ishizaki, K. Kashima, A. Girard, J. Imura, L. Chen, and K. Aihara, "Clustered model reduction of positive directed networks," *Automatica*, vol. 59, pp. 238–247, 2015.
[18] T. Sadamoto, T. Ishizaki, and J. Imura, "Projective state observers for large-scale linear systems," *in Proc. of the 13th European Control Conference*, pp. 2969–2974, 2014.
[19] S. P. Boyd, *Linear matrix inequalities in system and control theory*. Siam, 1994.
[20] J. Holman, *Heat Transfer tenth edition*, ser. McGraw-Hill Series in Mechanical Engineering. McGraw Hill Higher Education, 2009.
[21] J. Ferziger and M. Perić, *Computational methods for fluid dynamics*. Springer Berlin etc, 1999.
[22] F. Watanabe, T. Sadamoto, T. Ishizaki, and J. Imura, "Average state kalman filters for large-scale stochastic networked linear systems," *in Proc. of the 14th European Control Conference*, pp. 2823–2828, 2015.

Tomonori Sadamoto was born in Chiba, Japan, in 1987. He received the B.E. and M.E., from Tokyo Institute of Technology, Tokyo, Japan, in 2009 and 2011, respectively. From 2011 through 2013, he was with Technology Research Center of Sumitomo Heavy Industries. From 2014 to 2016, he was a research fellow of the Japan Society for the Promotion of Science. In 2015, he received the Ph.D degree in engineering from Tokyo Institute of Technology. From June in 2015 to March in 2016, he was a Visiting Researcher at School of Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden. Since April 2016, he is a researcher with the Department of Mechanical and Environmental Informatics, Graduate School of Information Science and Engineering, Tokyo Institute of Technology.

Dr. Sadamoto is a member of IEEE, SICE, and ISCIE. He was named as a finalist of the 13st European Control Conference Best Student-Paper Award.

Takayuki Ishizaki was born in Aichi, Japan, in 1985. He received the B.Sc., M.Sc., and Ph.D. degrees in Engineering from Tokyo Institute of Technology, Tokyo, Japan, in 2008, 2009, and 2012, respectively. He served as a Research Fellow of the Japan Society for the Promotion of Science from April 2011 to October 2012. From October to November 2011, he was a Visiting Student at Laboratoire Jean Kuntzmann, Universite Joseph Fourier, Grenoble, France. From June to October 2012, he was a Visiting Researcher at School of Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden. Since November 2012, he has been with the Department of Mechanical and Environmental Informatics, Graduate School of Information Science and Engineering, Tokyo Institute of Technology, where he is currently an Assistant Professor. His research interests include the development of model reduction and its applications.

Dr. Ishizaki is a member of IEEE, SICE, and ISCIE. He was a finalist of the 51st IEEE Conference on Decision and Control Best Student-Paper Award.



Jun-ichi Imura (M'93) was born in Gifu, Japan, in 1964. He received the M.S. degree in applied systems science and the Ph.D. degree in mechanical engineering from Kyoto University, Kyoto, Japan, in 1990 and 1995, respectively.

He served as a Research Associate in the Department of Mechanical Engineering, Kyoto University, from 1992 to 1996, and as an Associate Professor in the Division of Machine Design Engineering, Faculty of Engineering, Hiroshima University, from 1996 to 2001. From May 1998 to April 1999, he was a Visiting Researcher at the Faculty of Mathematical Sciences, University of Twente, Enschede, The Netherlands. Since 2001, he has been with the Department of Mechanical and Environmental Informatics, Graduate School of Information Science and Engineering, Tokyo Institute of Technology, Tokyo, Japan, where he is currently a Professor. His research interests include modeling, analysis, and synthesis of nonlinear systems, hybrid systems, and large-scale network systems with applications to biological systems, power grids, and transportation systems. He is an Associate Editor of Automatica (2009-), the Nonlinear Analysis: Hybrid Systems (2011-), and IEEE Trans. on Automatic Control (2014-).

Dr. Imura is a member of IEEE, SICE, ISCIE, IEICE, and The Robotics Society of Japan.