Average State Kalman Filters for Large-Scale Stochastic Networked Linear Systems

Fumiya Watanabe[1], Tomonori Sadamoto[2,3], Takayuki Ishizaki[2,4], and Jun-ichi Imura[2,4]

***Abstract*— In this paper, we propose a design method of average state Kalman filters for networked linear systems with stochastic noises. The average state Kalman filter is a low-dimensional estimator capturing the average behavior of systems from a macroscopic point of view. In general, it is nontrivial to find a set of states that captures the average behavior of systems. To overcome this difficulty, using the notion of clustering, we devise a systematic design procedure of average state Kalman filters while determining states that capture the average behavior of systems. Furthermore, deriving a tractable representation of the estimation error system, we derive an estimation error bound for the proposed method in a theoretical way. The efficiency of the proposed method is shown by a power system example in smart grid applications.**

I. INTRODUCTION

In recent years, systems of interest to control communities become more complex and larger in scale. For example, in smart grid, we are required to maintain supply-demand balance of power systems including more than one million consumers by controlling a number of power plants [1], [2]. In many cases, such large-scale complex systems are spatially distributed and networked.

Towards establishing a framework for systematic design of controllers for large-scale networked systems, in this paper, we consider designing a state estimator for large-scale networked systems based on input and output sequencial responses. Since measurement outputs in real systems are inevitably contaminated by noises, it is desirable to estimate system states while suppressing the influence of noises on estimated signals. One of well-known such estimators is the Kalman filter [3] that estimates system states with the minimization of the variance of estimation errors. However, the Kalman filter is necessary to have a dimension comparable with a system of interest. Thus, the Kalman filter for large-scale systems does not fully comply with practical application from a viewpoint of computational costs for implementation. In view of this, the development of low-dimensional estimators is crucial to deal with large-scale systems.

In [4], the authors propose a low-dimensional minimum variance estimator that exactly cancels the effect of external input signals with respect to the state estimation error. However, it is difficult to design low-dimensional estimators based on this method in general because the state-space of estimators must include states having even little influence on state estimation.

As one method in which the notion of approximation is introduced, in [5], [6], the authors propose the design method of low-dimensional estimators by constructing a type of Kalman filters for low-dimensional approximate models of systems of interest. However, they do not clarify the relation between the approximation error and the estimation error for the designed low-dimensional estimator. Thus, the construction of approximate models is rather heuristic.

In [7], we have proposed an average state observer capturing behavior of large-scale systems from a macroscopic point of view. We have clarified that the influence of external input signals appears in the estimation error. Furthermore, we have provided a design method to make the influence of inputs on the estimation error small while determining a set of states capturing the average behavior of systems to be estimated on the basis of the clustered model reduction technique [8].

This paper continues upon the research of [7], and extends the case to a type of Kalman filters. In particular, we propose an *average state Kalman filter*, where measurement and system noises are taken into explicit consideration. However, the extenstion of average state observers developed in [7] is not straightforward because the influence of noises appears in the estimation error variance. In addition, as shown in [7], the external input signal has the influence on the estimation error, which results in a non-zero expected value of the estimation error, i.e., unbiased estimation. In view of this, we consider designing it by evaluating the sum of variance and expectancy of estimation errors. Furthermore, we show a theoretical estimation error bound with the provision of a systematic design procedure. The proposed average state Kalman filter is expected to be useful as a first step of practical application, e.g., wave surge prediction in [9] and power estimation in smart grid [10]. Finally, we show the efficiency of our method through a power system example of the IEEE 118 test system provided by [11].

This paper is organized as follows: In Section II-A, we introduce the notion of clustering for networked systems composed of multi-dimensional subsystems. This is the generalized notion in [8], [12] where the clustering technique is proposed for interconnected one-dimensional subsystems and second order subsystems, respectively. Then, we define an average state Kalman filter as an operator to generate signals compatible with states of clustered subsystems. In Section II-

[1]Department of Control and Systems Engineering, Tokyo Institute of Technology; 2-12-1, Meguro, Tokyo, Japan

[2]Department of Mechanical and Environmental Informatics, Graduate School of Information Science and Engineering, Tokyo Institute of Technology; 2-12-1, Meguro, Tokyo, Japan

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

[4]Japan Science and Technology Agency, CREST 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan

{watanabe@cyb., sadamoto@cyb., ishizaki@, imura@}mei.titech.ac.jp

B, using orthogonal projection [13], we formulate average state filters described by linear stochastic discrete-time systems. Furthermore, deriving a tractable representation of the estimation error system, we clarify that the estimation error depends on not only the estimation error variance but also the estimation error bias. Then, we formulate a design problem of the average state Kalman filter. In Section III, on the basis of the error analysis, we show a theoretical estimation error bound with the provision of a systematic design procedure for average state Kalman filters. In Section IV, we show the efficiency of the proposed method through a numerical example of a power system. Finally, concluding remarks are provided in Section V.

Notation The following notation is used:

$\mathbb{R}$	set of real numbers
$\mathbb{Z}$	set of non-negative integers
I_n	unit matrix of size $n \times n$
e_i^n	the ith column of I_n
$e_{\mathcal{I}}^n$	$e_{\mathcal{I}}^n := [e_{i_1}^n, \dots, e_{i_m}^n]$ for $i \in \mathcal{I} := \{i_1, \dots, i_m\}$
$O_{m\times n}$ (O_n)	zero matrix of size $m\times n$ ($n\times n$)
$M \prec O_n$ ($M \succ O_n$)	negative (positive) definiteness of a symmetric matrix $M \in \mathbb{R}^{n\times n}$
$M \preceq O_n$ ($M \succeq O_n$)	negative (positive) semidefiniteness of a symmetric matrix $M \in \mathbb{R}^{n\times n}$
$\mathrm{tr}(M)$	trace of a matrix M
$\mathbf{1}_n$	$\mathbf{1}_n = [1, \dots, 1]^\mathsf{T} \in \mathbb{R}^n$
$\|M\|_F$	the Frobenius norm of a matrix M
$\mathbb{E}[X]$	expectation of a stochastic variable X
$\rho(M)$	$\rho(M) := \max_i\{\lvert\lambda_i\rvert\}$ where λ_i denotes the ith eigenvalue of M

For $\mathcal{N} = \{1, \dots, N\}$, we denote the block-diagonal matrix having matrices $M_1, \dots, M_N$ on its diagonal blocks by $\mathrm{dg}(M_i)_{i\in\mathcal{N}}$. In particular, if not confusing, we omit the subscript of $i \in \mathcal{N}$. Furthermore, the operator $\otimes$ denotes Kronecker product. In addition, we denote Kronecker delta as

$$\delta(t) = \begin{cases} 1 & t = 0 \\ 0 & t \neq 0 \end{cases}, \quad t \in \mathbb{Z}.$$

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

$$\|G(z)\|_{\mathcal{H}_\infty} := \sup_{\omega\in\mathbb{R}} \|G(e^{j\omega})\|,$$
$$\|G(z)\|_{\mathcal{H}_2} := \left(\frac{1}{2\pi}\int_0^{2\pi} \mathrm{tr}(G(e^{j\omega})G^\mathsf{T}(e^{-j\omega}))d\omega\right)^{\frac{1}{2}}$$

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

II. Problem Formulation

A. Introducing Average State Kalman Filter

In this paper, we deal with discrete-time stochastic linear systems composed of N subsystems. Let us consider a set of κ-dimensional subsystems. For each $i \in \mathcal{N} := \{1, \dots, N\}$, the dynamics of the ith subsystem is described by

$$\Sigma_i : \begin{cases} x_i(k+1) = A_i x_i(k) + \sum_{j\neq i}^N A_{ij}x_j(k) + B_i u_i(k) + w_i(k) \\ y_i(k) = C_i x_i(k) + D_i u_i(k) + v_i(k) \end{cases} \tag{1}$$

where $A_i \in \mathbb{R}^{\kappa\times\kappa}$, $B_i \in \mathbb{R}^{\kappa\times m_i}$, $A_{ij} \in \mathbb{R}^{\kappa\times\kappa}$, $C_i \in \mathbb{R}^{p_i\times\kappa}$, $D_i \in \mathbb{R}^{p_i\times m_i}$, and $y_i \in \mathbb{R}^{p_i}$ is a measurement output and $u_i \in \mathbb{R}^{m_i}$ is an input. In addition, the system noise $w_i \in \mathbb{R}^\kappa$ and the measurement noise $v_i \in \mathbb{R}^{p_i}$ are assumed to be zero-mean white Gaussian and mutually uncorrelated with variance $Q_i \succeq O_\kappa$ and $R_i \succeq O_{p_i}$, i.e.

$$\mathbb{E}[w_i(k)] = 0, \quad \mathbb{E}[v_i(k)] = 0$$
$$\mathbb{E}\left[\begin{bmatrix} w_i(k) \\ v_i(k) \end{bmatrix} [w_i^\mathsf{T}(l),\ v_i^\mathsf{T}(l)]^\mathsf{T}\right] = \begin{bmatrix} Q_i & \\ & R_i \end{bmatrix}\delta(k-l) \tag{2}$$

for any $k \in \mathbb{Z}$ and $l \in \mathbb{Z}$. Throughout this paper, we suppose that $x_i \in \mathbb{R}^\kappa$ represents the same physical quantities for all $i \in \mathcal{N}$. For example, as shown in Example 1, we deal with Σ_i as a two-dimensional oscillator having the state variable of $x_i := [\theta_i\ \omega_i]^\mathsf{T}$ where $\theta_i \in \mathbb{R}$ and $\omega_i \in \mathbb{R}$ represent an angle and angular velocity, respectively. In this paper, we consider such subsystems having the same physical quantities among N subsystems.

In addition, we use the notation of

$$n := N\kappa, \quad m := \sum_{i=1}^N m_i, \quad p := \sum_{i=1}^N p_i$$

and

$$A := \begin{bmatrix} A_1 & \cdots & A_{1N} \\ \vdots & \ddots & \vdots \\ A_{N1} & \cdots & A_N \end{bmatrix} \in \mathbb{R}^{n\times n}, \quad B := \mathrm{dg}(B_i) \in \mathbb{R}^{n\times m}$$
$$C := \mathrm{dg}(C_i) \in \mathbb{R}^{p\times n}, \quad D := \mathrm{dg}(D_i) \in \mathbb{R}^{p\times m}.$$

In this notation, we give the dynamics of the whole networked system as

$$\Sigma : \begin{cases} x(k+1) = Ax(k) + Bu(k) + w(k) \\ y(k) = Cx(k) + Du(k) + v(k) \end{cases} \tag{3}$$

where $\circ := [\circ_1^\mathsf{T}, \dots, \circ_N^\mathsf{T}]^\mathsf{T}$ for each $\circ \in \{x, u, y, w, v\}$. For this system, we consider introducing the notion of clustering as follows:

Definition 1: Let Σ in (3) be given. The family of an index set $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ for $\mathbb{L} := \{1, \dots, L\}$ is called a *cluster set*, whose element is referred to as a cluster, if each element $\mathcal{I}_{[l]}$ is a disjoint subset of $\mathcal{N}$ and it satisfies

$$\bigcup_{l\in\mathbb{L}} \mathcal{I}_{[l]} = \mathcal{N}.$$

$$\begin{bmatrix} \theta_i(k+1) \\ \omega_i(k+1) \end{bmatrix} = \begin{bmatrix} 1 & \tau \\ & 1-\frac{D_i\tau}{M_i} \end{bmatrix} \begin{bmatrix} \theta_i(k) \\ \omega_i(k) \end{bmatrix} - \begin{bmatrix} 0 \\ \frac{\tau}{M_i} \end{bmatrix} \sum_{j\neq i} y_{ij}(\theta_i(k)-\theta_j(k)) + \begin{bmatrix} 0 \\ \frac{\tau}{M_i} \end{bmatrix} (u_i(k)+w_i(k)) \tag{6}$$

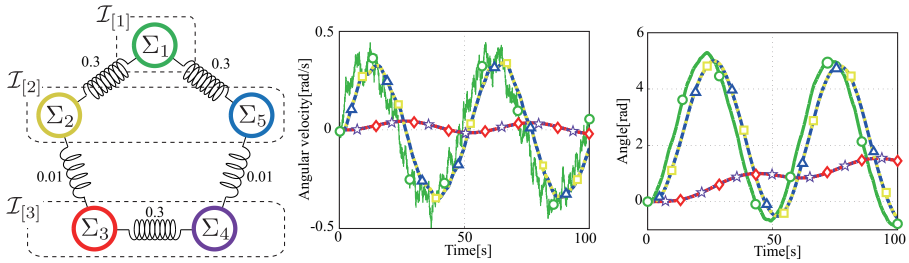

Fig. 1. Illustrative example: Networked oscillators model having five subsystems and their input responses. Angular velocities (in the middle figure) and angles (in the rightmost figure) of individual subsystems are depicted by the lines with markers ○, □, ◇, ★ and △ for i in the ascending order.

Then, an *embedding matrix* compatible with $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ is defined by

$$W := \left(\Pi \ \mathrm{dg}\left(\frac{1}{\sqrt{n_1}}\mathbf{1}_{n_1},\ldots,\frac{1}{\sqrt{n_L}}\mathbf{1}_{n_L}\right)\right) \otimes I_\kappa \in \mathbb{R}^{n\times L\kappa} \tag{4}$$

where n_l is the cardinality of $\mathcal{I}_{[l]}$ and the permutation matrix

$$\Pi := [e^N_{\mathcal{I}_{[1]}},\ldots,e^N_{\mathcal{I}_{[L]}}] \in \mathbb{R}^{N\times N}. \tag{5}$$

In this paper, we consider estimating average behavior of Σ. To illustrate this purpose, let us consider the following example:

Example 1: Let Σ in (3) be composed of five subsystems, i.e., $N = 5$, and each subsystem Σ_i be a discrete-time networked oscillators model in (6) with a sampling interval $\tau = 0.01$, unit mass and a unit damping coefficient, i.e., $M_i = D_i = 1$. The interconnection structure is shown in the leftmost in Fig. 1 where the coefficient between Σ_i and Σ_j denotes $y_{ij} = y_{ji}$ in (6), e.g. $y_{12} = y_{21} = 0.3$. Let $x(0) = 0$, $Q_1 = 0.5$ and $u_1(t) = \sin(50t)$, and $Q_i = 0$ and $u_i(t) = 0$ for $i \in \{2,\ldots,5\}$. In Fig. 1, we plot the resultant trajectories of ω_i and θ_i of individual subsystems by green, yellow, red, purple and blue lines for i in the ascending order. From Fig. 1, we can see that $x_2(t)$ coincides with $x_5(t)$ and $x_3(t)$ coincides with $x_4(t)$. This implies that two bundles of $\{x_i\}_{i\in\{2,5\}}$ and $\{x_i\}_{i\in\{3,4\}}$ can be represented by four-dimensional signal. In other words, by taking into account $x_1 \in \mathbb{R}^2$, estimating a signal $f(k) \in \mathbb{R}^6$ to make the norm of

$$\Delta := x(k) - Wf(k) \tag{7}$$

small in a suitable sense where W is the embedding matrix in (4) with cluster sets

$$\mathcal{I}_{[1]} = \{1\}, \quad \mathcal{I}_{[2]} = \{2,5\}, \quad \mathcal{I}_{[3]} = \{3,4\},$$

we can capture average behavior of Σ instead of estimating all of 10 trajectories of x.

In view of this, we define a filter to estimate average behavior of Σ as follows: Let $\mathbb{U} := \{u : \mathbb{Z} \to \mathbb{R}^m\}$, $\mathbb{Y} := \{y : \mathbb{Z} \to \mathbb{R}^p\}$ and $\mathbb{F} := \{f : \mathbb{Z} \to \mathbb{R}^{L\kappa}\}$. Define the *average state Kalman filter* $\mathcal{F}$ as

$$\mathcal{F} : \mathbb{U}\times\mathbb{Y} \to \mathbb{F} \tag{8}$$

to make the magnitude of estimation error Δ in (7) small in a suitable sense. Owing to the block-diagonal structure of W, we can regard f as an *average state* of x.

In general, we do not know a cluster sets capturing average behavior of Σ in advance. In view of this, we suppose that cluster sets $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ is not given in advance. In what follows, we give a criterion to evaluate the magnitude of the estimation error Δ, and devise a systematic method to design $\mathcal{F}$ as well as determining cluster sets $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$. Note that a method to achieve usual averaging, i.e., normalized by using n_l in (4) but not $\sqrt{n_l}$, is described in Remark 2.

B. Design Problem of Average State Kalman Filters

In this section, we formulate a problem to design average state filters $\mathcal{F}$ in (8). For simplicity, we focus on time-invariant filters and deal with stable Σ in (3), i.e., $\rho(A) < 1$.

Note that the embedding matrix W in (4) satisfies $W^\mathsf{T}W = I_{L\kappa}$ and an output signal f in (8) estimates x in (3) by expanding by W. On the basis of the notion of orthogonal projection [13], as an instance of $\mathcal{F}$ to generate such a signal f, let us consider a linear filter described by

$$\mathcal{F}: \begin{cases} \hat{x}(k+1) = W^\mathsf{T}AW\hat{x}(k) + W^\mathsf{T}Bu(k) + H(y(k)-\hat{y}(k)) \\ \quad \hat{y}(k) = CW\hat{x}(k) + Du(k) \\ \quad f(k) = \hat{x}(k) \end{cases} \tag{9}$$

with $\hat{x}(0) = 0$ where $\hat{n} := L\kappa$ and $H \in \mathbb{R}^{\hat{n}\times p}$ is a filter gain. Without loss of generality, we assume that $\hat{n} \le n$. To analyze the estimation error Δ in (7), we give the following lemma:

Lemma 1: Let Σ in (3) be given. Suppose that a cluster set $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ is given, and an embedding matrix W is defined as in Definition 1. Consider $\mathcal{F}$ in (9). Then

$$\mathcal{E}_W : \begin{cases} \xi(k+1) = \mathcal{A}\xi(k) + \mathcal{B}u(k) + \eta(k) \\ \quad \Delta(k) = \mathcal{W}\xi(k) \end{cases} \tag{10}$$

with $\xi(0) = [x^\mathsf{T}(0)W,\ x^\mathsf{T}(0)]^\mathsf{T}$ where

$$\eta(k) := \begin{bmatrix} W^\mathsf{T}w(k) - Hv(k) \\ w(k) \end{bmatrix}$$

and

$$\mathcal{A} = \begin{bmatrix} W^\mathsf{T}AW - HCW & (W^\mathsf{T}A - HC)(I_n - WW^\mathsf{T}) \\ O_{n\times\hat{n}} & A \end{bmatrix}$$

$$\mathcal{B} = \begin{bmatrix} O_{\hat{n}\times m} \\ B \end{bmatrix}, \quad \mathcal{W} = \begin{bmatrix} W & I_n - WW^\mathsf{T} \end{bmatrix}.$$

Proof: Omit due to page limitation. ■

In Lemma 1, it should be emphasized that Δ in (7) depends on not only $x(0)$ but also $u(k)$ as long as $WW^\mathsf{T} \neq I_n$. Thus, even if a sufficiently long time is elapsed, the expected value of Δ is biased, i.e.,

$$\lim_{k\to\infty} \mathbb{E}\left[\Delta(k)\right] \neq 0$$

unless u is applied to Σ and $\mathcal{F}$ constantly. Although usual n-dimensional Kalman filter in [3] makes magnitude of the variance of Δ small, in our case, we should take into account not only the variance of Δ but also the bias of Δ. Therefore, in this paper, we consider evaluating

$$\lim_{k\to\infty} \mathbb{E}\left[\|\Delta(k)\|\right] \tag{11}$$

for designing $\mathcal{F}$. Note that (11) coincides with the sum of the norm of the bias of Δ and trace of the variance of Δ at the infinite time. To analyze the estimation error Δ in a theoretical way, we evaluate (11) for u of a white Gaussian process with zero-mean and a variance I_m, i.e., for any $k \in \mathbb{Z}$ and $l \in \mathbb{Z}$,

$$\mathbb{E}[u(k)] = 0, \quad \mathbb{E}[u(k)u^\mathsf{T}(l)] = \delta(k-l)I_m. \tag{12}$$

In this setting, we formulate the problem to design average state filters $\mathcal{F}$ as follows:

Problem 1: For a given Σ in (3), consider a cluster set $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ in Definition 1. Consider $\mathcal{F}$ in (9) and Δ in (7). For a given $\epsilon \geq 0$, find $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ and $\mathcal{F}$ such that

$$\lim_{k\to\infty} \mathbb{E}\left[\|\Delta(k)\|\right] \leq \epsilon \tag{13}$$

for a white Gaussian process $\{u(k)\}$ satisfying (12) and all $x(0) \in \mathbb{R}^n$.

III. Design of Average State Kalman Filters

A. A Road Map for Systematic Design

In the previous subsection, we have formulated the problem to design an average state Kalman filter. Problem 1 is formulated as a problem to find W in (4) and a filter gain H in (9) satisfying (13). Note that the left-hand side in (13) is independent from $x(0)$. Taking the z-transformation of Δ in (7), we have

$$\Delta(z) = \Xi_{W,H}(z)X_W(z;u,w) + N_{W,H}(z;w,v) \tag{14}$$

where

$$\Xi_{W,H}(z) := W(zI_{\hat{n}} - A_\Xi)^{-1}B_\Xi + \overline{W} \tag{15}$$

with

$$A_\Xi := W^\mathsf{T}AW - HCW, \quad B_\Xi := (W^\mathsf{T}A - HC)\overline{W} \tag{16}$$

and

$$\begin{aligned} X_W(z;u,w) &:= \overline{W}^\mathsf{T}(zI_n - A)^{-1}[Bu(z) + w(z)] \\ N_{W,H}(z;w,v) &:= W(zI_{\hat{n}} - A_\Xi)^{-1}\left[W^\mathsf{T}w(z) \quad -Hv(z)\right] \end{aligned} \tag{17}$$

with $\overline{W} \in \mathbb{R}^{(n-\hat{n})\times n}$ satisfying

$$WW^\mathsf{T} + \overline{W}\,\overline{W}^\mathsf{T} = I_n. \tag{18}$$

In this notation, it follows that

$$\lim_{k\to\infty} \mathbb{E}[\|\Delta(k)\|] \leq \|\Xi_{W,H}(z)X_W(z)\|_{\mathcal{H}_2} + \|N_{W,H}(z)\|_{\mathcal{H}_2}. \tag{19}$$

Noting that $\Xi_{W,H}(z)X_W(z) = 0$ for $W = I_n$, we can see that (19) coincides with the standard evaluation value of the Kalman filter. Furthermore, it follows that $\|\Xi_{W,H}(z)X_W(z)\|_{\mathcal{H}_2} \neq 0$ unless $WW^\mathsf{T} \neq I_n$ because the average state Kalman filter cannot exactly predict the dynamical behavior of systems for u and w. In this sense, we can see (19) as the generalized evaluation taking into account the estimation error arising from prediction errors.

In what follows, we consider making the first and second term in the right-hand side of (19) sufficiently small. However, it should be noted that the simultaneous design of W and H is difficult because $\Xi_{W,H}$ and $N_{W,H}$ contain parameters W and H in a nonlinear fashion. To overcome this difficulty, we use the following facts:

- X_W depends on W, but not H.
- We have
$$\|\Xi_{W,H}(z)X_W(z)\|_{\mathcal{H}_2} \leq \beta\|X_W(z)\|_{\mathcal{H}_2} \tag{20}$$
where $\beta := \|\Xi_{W,H}(z)\|_{\mathcal{H}_\infty}$.
- The parameters W and H appear in $\Xi_{W,H}$ and $N_{W,H}$.

On the basis of these facts, we first determine W to make $\|X_W(z)\|_{\mathcal{H}_2}$ small. More specifically, let $\Phi \succeq O_n$ be given such that

$$A^\mathsf{T}\Phi A - \Phi + BB^\mathsf{T} + Q = 0. \tag{21}$$

Furthermore, for $j \in \{1,\ldots,\kappa\}$, define

$$\Phi_{\frac{1}{2}}^{(j)} := (e^n_{\mathcal{K}^{(j)}})^\mathsf{T}\Phi_{\frac{1}{2}} \in \mathbb{R}^{N\times n} \tag{22}$$

where $\Phi_{\frac{1}{2}} \succeq O_n$ is a Cholesky factor of Φ, i.e., $\Phi = \Phi_{\frac{1}{2}}\Phi_{\frac{1}{2}}^\mathsf{T}$ and $\mathcal{K}^{(j)} \subseteq \{1,\ldots,n\}$ is the set of indices to represent the jth physical quantity of x_i for all $i \in \mathcal{N}$, i.e., $\mathcal{K}^{(j)}$ is given such that

$$[x_{1,j},\ldots,x_{N,j}]^\mathsf{T} = (e^n_{\mathcal{K}^{(j)}})^\mathsf{T}x \tag{23}$$

where $x_{i,j}$ denotes the jth element of states of the ith subsystem. In Example 1, $\mathcal{K}^{(1)} = \{1,3,5,7,9\}$ and $\mathcal{K}^{(2)} = \{2,4,6,8,10\}$. One approach to construct W in (4) and $\overline{W}$ satisfying (18) while achieving small $\|X_W(z)\|_{\mathcal{H}_2}$ is to take

$$\overline{W} = I_\kappa \otimes (\Pi\mathrm{dg}(\overline{w}_l)_{l\in\mathbb{L}}) \tag{24}$$

with Π in (5) where $\overline{w}_l \in \mathbb{R}^{n_l\times(n_l-1)}$ is given such that $[\overline{w}_l, \mathbf{1}_{n_l}]$ is unitary. Then, (18) holds for any $\overline{w}_l$. Furthermore,

$$\begin{aligned} \|X_W(z)\|^2_{\mathcal{H}_2} &= \|\overline{W}^\mathsf{T}[(\Phi_{\frac{1}{2}}^{(1)})^\mathsf{T},\ldots,(\Phi_{\frac{1}{2}}^{(\kappa)})^\mathsf{T}]^\mathsf{T}\|_F^2 \\ &= \sum_{i=1}^{\kappa}\|\mathrm{dg}(\overline{w}_l^\mathsf{T})_{l\in\mathbb{L}}\Pi^\mathsf{T}\Phi_{\frac{1}{2}}^{(i)}\|_F^2 \end{aligned}$$

holds. Hence, we can construct $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ if there exists a family of $\{\overline{w}_l\}_{l\in\mathbb{L}}$ making $\|X_W(z)\|_{\mathcal{H}_2}$ small. On the basis of this analysis, we have the following lemma:

Lemma 2: Consider Problem 1. Let $\Phi_{\frac{1}{2}}^{(j)}$ in (22) be given for $j \in \{1, \ldots, \kappa\}$. If there exist $\phi_{[l]}^{(j)} \in \mathbb{R}^{1\times n}$ and $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ such that

$$\max_{j\in\{1,\ldots,\kappa\}} \left\| (e_{\mathcal{I}_{[l]}}^N)^\mathsf{T} \Phi_{\frac{1}{2}}^{(j)} - \frac{1}{\sqrt{n_l}} \mathbf{1}_{n_l} \phi_{[l]}^{(j)} \right\|_F^2 \le n_l \sigma^2, \quad l \in \mathbb{L} \tag{25}$$

for a given $\sigma \ge 0$, then

$$\|X_W(z)\|_{\mathcal{H}_2} \le \alpha\sigma \tag{26}$$

where $\alpha := \sqrt{\sum_l^L \kappa n_l (n_l - 1)}$.

Proof: Omit due to page limitation. ■

In this lemma, a design parameter σ represents a coarseness parameter for cluster construction. Supposing that $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ is given satisfying the assumtion in Lemma 2, we consider determining H to make the norm of $N_{W,H}$ in (17) small. To this end, we give the following theorem.

Theorem 1: Consider Problem 1. Suppose that there exist $\phi_{[l]}^{(j)} \in \mathbb{R}^{1\times n}$ and $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ such that (25). Furthermore, if there exist

$$\gamma > 0, \quad X \succ \mathcal{O}_{\hat{n}}, \quad Y \in \mathbb{R}^{\hat{n}\times p}, \quad Z \succ \mathcal{O}_{\hat{n}}$$

such that

$$\begin{bmatrix} X & (XW^\mathsf{T}AW - YCW) & \begin{bmatrix} XW^\mathsf{T}Q_{\frac{1}{2}} & -YR_{\frac{1}{2}} \end{bmatrix} \\ * & X & 0 \\ * & * & I_{n+p} \end{bmatrix} \succ \mathcal{O}_{2\hat{n}+n+p}$$
$$\begin{bmatrix} X & I_{\hat{n}} \\ * & Z \end{bmatrix} \succeq \mathcal{O}_{2\hat{n}}, \quad \mathrm{tr}\,(Z) < \gamma^2 \tag{27}$$

where W is given in (4) and $Q_{\frac{1}{2}}$ and $R_{\frac{1}{2}}$ are Cholesky factors of Q and R, respectively. Then, $\mathcal{F}$ in (9) with

$$H = X^{-1}Y \tag{28}$$

satisfies

$$\lim_{k\to\infty} \mathbb{E}\left[\|\Delta(k)\|\right] < \alpha\beta\sigma + \gamma \tag{29}$$

with σ in (25), α in (26) and β in (20) for a white Gaussian process $\{u(k)\}$ satisfying (12) and all $x(0) \in \mathbb{R}^n$.

Proof: Omit due to page limitation. ■

Theorem 1 provides an explicit bound of the estimation error by the average state Kalman filter. Note that if $\sigma = 0$, i.e., $W = I_n$, then $\mathcal{F}$ in (9) turns out to be the Kalman filter [3].

Remark 1: In Theorem 1, β in the right-hand side in (29) is calculated after designing H. To construct H having a priori error bound of $\|\Xi_{W,H}(z)\|_{\mathcal{H}_\infty}$, it suffices to add an equivalent LMI of $\|\Xi_{W,H}(z)\|_{\mathcal{H}_\infty} \le \beta$ for a given $\beta > 0$ to (27).

B. Designing Algorithm of Average State Kalman Filters

In this subsection, we provide an algorithm to design the average state filter $\mathcal{F}$ in (9) for a given design paramter ϵ in Problem 1 as follows:

Suppose that a design parameter σ in (25) is given and a set of clusters $\{\mathcal{I}_{[1]}, \ldots, \mathcal{I}_{[l]}\}$ is already constructed. Let

$$\mathcal{S} := \{1, \ldots, n\} \backslash \{\mathcal{I}_{[1]}, \ldots, \mathcal{I}_{[l]}\}.$$

Next, we first choose an index $s \in \mathcal{S}$. Subsequently, construct a new cluster $\mathcal{I}_{[l+1]}$ such that

$$\mathcal{I}_{[l+1]} = \left\{ i \in \mathcal{S}\backslash\{s\} \middle| \max_{j\in\{1,\ldots,\kappa\}} \left\| \phi_i^{(j)} - \phi_s^{(j)} \right\| \le \sigma \right\} \tag{30}$$

where $\phi_i^{(j)} \in \mathbb{R}^{1\times n}$ denotes the ith row vector of $\Phi_{\frac{1}{2}}^{(j)}$. We can straightforwardly verify that this newly constructed cluster $\mathcal{I}_{[l+1]}$ satisfies (25).

In this setting, for stable Σ in (3), we summarize the design procedure of an $\hat{n}$-dimensional average state Kalman filter as follows:

a) Give $\epsilon \ge 0$ and $\sigma \ge 0$.
b) Find $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ by the above procedure.
c) Find $X \succ \mathcal{O}_{\hat{n}}$, $Y \in \mathbb{R}^{\hat{n}\times p}$, $Z \succ \mathcal{O}_{\hat{n}}$ and minimal $\gamma > 0$ satisfying (27).
d) If no solutions exist or $\epsilon < \|\Delta(z)\|_{\mathcal{H}_2}$, take smaller σ, and go back to b).
e) Construct an average state Kalman filter $\mathcal{F}$ in (9) with H given by (28).

Finally, it should be noted that since the number of decision variables of LMI given by (27) is $\hat{n}^2 + (p-1)\hat{n} + 1$, this design procedure is computationally tractable if $\hat{n}$ is small.

Remark 2: The average state f given by $\mathcal{F}$ in (9) with an embedding matrix W in (4) implies an average state normalized by $\sqrt{n_l}$. Alternatively, if we define

$$f = S\hat{x}, \quad S := \mathrm{diag}\left(\frac{1}{\sqrt{n_1}}, \ldots, \frac{1}{\sqrt{n_L}}\right) \otimes I_\kappa \in \mathbb{R}^{L\kappa\times L\kappa} \tag{31}$$

then, the norm of estimation error Δ in (7) turns out to be

$$\|\Delta\| = \sqrt{\sum_{l\in\mathbb{L}} \sum_{i\in\mathcal{I}_{[l]}} \|x_i - f_l\|^2} \tag{32}$$

where $f_l \in \mathbb{R}^\kappa$ such that $[f_1^\mathsf{T}, \ldots, f_L^\mathsf{T}]^\mathsf{T} = f$. Thus, signal f generated by the average state filter with (31) coincides with the average state of x in the usual sense.

IV. Numerical Example

In this section, we show the efficiency of the proposed average state filter through a numerical example. We deal with the IEEE 118 power test system provided by [11]. For simplicity, we replace the synchronous condensers in the original test system with generators. In addition, we regard each generator as each subsystem. The number of subsystems is 54, i.e., $N = 54$, which yields $n = 108$. For $i \in \mathcal{N} := \{1, \ldots, 54\}$, the dynamics of the ith generator is given in (6) represented by networked oscillators provided by [14]. In addition, the admittance y_{ij} is calculated by using MATPOWER in [11].

Let the input u be applied to the first to third generators, i.e.,

$$B = \begin{bmatrix} \mathrm{dg}(\frac{\tau}{M_i} e_1^2)_{i\in\{1,2,3\}} \\ 0_{102\times 3} \end{bmatrix} \in \mathbb{R}^{108\times 3}.$$

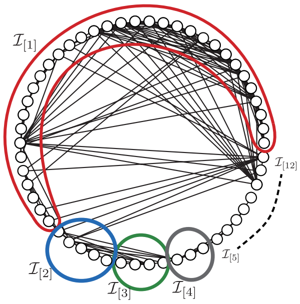

Fig. 2. Network structure of the IEEE 118 test system and resultant cluster sets for $L = 12$.

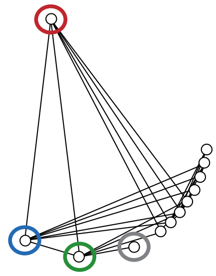

Fig. 3. Clustered network structure with resultant $\{\mathcal{I}_{[l]}\}_{l\in\{1,\ldots,12\}}$.

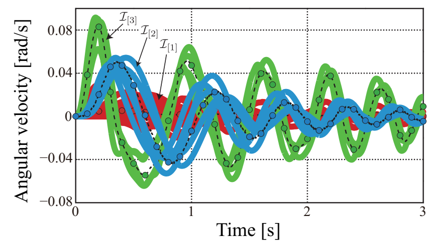

Fig. 4. Trajectories of angular velocities of x and f.

In addition, we take the measurement output y as the states of the first to eighth generators, i.e.,

$$C = \begin{bmatrix} I_{16} & 0_{16\times 92} \end{bmatrix} \in \mathbb{R}^{16\times 108}.$$

Furthermore, we take the variance of w_i and v_i as $Q_i = \mathrm{dg}(0.5, 0)$ for $i \in \{1, 4\}$, and $Q_i = 0$ for $i \in \mathcal{N}\backslash\{1, 4\}$ and $R = 0.001 \times I_{12}$, respectively. Let the sampling interval $\tau = 0.001$, $D_i = 7$ for $i \in \mathcal{N}$, and $M_i = 0.1$ for $i \in \mathcal{J}$, where $\mathcal{J}$ denotes the set of indices of generators that connect to the first to third generators, and $M_i = 1.2$ for $i \in \mathcal{N}\backslash\mathcal{J}$.

We design an average state Kalman filter $\mathcal{F}$ in (9) by the procedure shown in Section III-B for a given value of σ in (25) such that $L = 12$, which yields $\hat{n} = 24$. First, we show the resultant cluster sets $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ in Definition 1 and network structure of the power system in Fig. 2 where individual circles denote generators. In addition, in Fig. 3, we show the resultant network structure composed of the clustered subsystems. Furthermore, taking $x(0) = 0$ and random input signals u containing multiple frequency waves, in Fig. 4, we plot the trajectories of angular velocities of subsystems compatible with $\{\mathcal{I}_{[l]}\}_{l\in\{1,2,3\}}$ and those of average states f compatible with angular velocities. The indication in Fig. 4 is as follows: the trajectories of x are depicted as red, blue and green lines and those of f are depicted as red, blue and green dotted lines with circles. We can see from this figure that each trajectory of f is around the center of colored trajectory sets. Moreover, defining the resultant average behavior as

$$\xi_l(k) := \sum_{i\in\mathcal{I}_{[l]}} \frac{1}{n_l} x_i(k), \quad l \in \mathbb{L},$$

we calculate the resultant estimation error of ξ_l by the average state f_l in (32) normalized by the norm of ξ_l, i.e., $\sum_{l\in\mathbb{L}} \|\xi_l(k) - f_l(k)\| / \sum_{l\in\mathbb{L}} \|\xi_l\| = 0.05$. These results imply that the proposed average state Kalman filter can estimate the average behavior of networked systems.

V. Conclusion

In this paper, we have proposed a design method of average state Kalman filters, which are low-dimensional estimators capturing average behavior of systems, for networked stochastic linear systems. Deriving a tractable representation of the error system. We have shown a theoretical error bound and a systematic design procedure of average state Kalman filters. The efficiency of the proposed method has been shown by using a power system example of the IEEE 118 bus test system. In this paper, we have dealt with time-invariant average state Kalman filters. The extention to time-variant filters is currently under investigation as well as practical application, e.g., wave surge prediction in [9].

VI. Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 14J11344.

References

[1] P. Kundur, *Power system stability and control*. McGraw-Hill Education, 1994.

[2] M. D. Ilic and S. Liu, *Hierarchical power systems control: its value in a changing industry*. Springer Heidelberg, 1996.

[3] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Fluids Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[4] M. Darouach, M. Boutayeb, and M. Zasadzinski, "Reduced order filtering for stochastic discrete-time systems with unknown inputs," in *World Congress*, vol. 15, no. 1, 2002, pp. 508–508.

[5] I. Fukumori and P. Malanotte-Rizzoli, "An approximate kaiman filter for ocean data assimilation: An example with an idealized gulf stream model," *Journal of Geophysical Research: Oceans (1978–2012)*, vol. 100, no. C4, pp. 6777–6793, 1995.

[6] I. Fukumori, "A partitioned kalman filter and smoother," *Monthly Weather Review*, vol. 130, no. 5, pp. 1370–1383, 2002.

[7] T. Sadamoto, T. Ishizaki, and J. Imura, "Projective state observers for large-scale linear systems," in *European Control Conference*. IEEE, 2014, pp. 2969–2974.

[8] T. Ishizaki, K. Kashima, J. Imura, and K. Aihara, "Model reduction and clusterization of large-scale bidirectional networks," *Automatic Control, IEEE Transactions on*, vol. 59, no. 1, pp. 48–63, 2014.

[9] M. Verlaan, "Efficient kalman filtering algorithms for hydrodynamic models," *Ph.D. Thesis at Technische Universiteit, Delft, The Netherlands*, 1998.

[10] 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.

[11] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *Power Systems, IEEE Transactions on*, vol. 26, no. 1, pp. 12–19, 2011.

[12] T. Ishizaki and J. Imura, "Clustered model reduction of interconnected second-order systems, nonlinear theory and its applications," *Institute of Electronics, Information and Communication Engineers, Special Section of Complex Systems Modelling and its Transdisciplinary Applications(in Press)*, vol. E6-N, no. 1, 2014.

[13] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2005.

[14] F. Dörfler and F. Bullo, "Synchronization in complex networks of phase oscillators: A survey," *Automatica*, vol. 50, no. 6, pp. 1539–1564, 2014.