Delay-Dependent Distributed Kalman Fusion Estimation With Dimensionality Reduction in Cyber-Physical Systems

Bo Chen*, Member, IEEE, Daniel W. C. Ho*, Fellow, IEEE, Guoqiang Hu*, Senior Member, IEEE, and Li Yu*, Member, IEEE

Abstract—This article studies the distributed dimensionality reduction fusion estimation problem with communication delays for a class of cyber-physical systems (CPSs). The raw measurements are preprocessed in each sink node to obtain the local optimal estimate (LOE) of a CPS, and the compressed LOE under dimensionality reduction encounters with communication delays during the transmission. Under this case, a mathematical model with compensation strategy is proposed to characterize the dimensionality reduction and communication delays. This model also has the property of reducing the information loss caused by the dimensionality reduction and delays. Based on this model, a recursive distributed Kalman fusion estimator (DKFE) is derived by optimal weighted fusion criterion in the linear minimum variance sense. A stability condition for the DKFE, which can be easily verified by the exiting software, is derived. In addition, this condition can guarantee that the estimation error covariance matrix of the DKFE converges to the unique steady-state matrix for any initial values and, thus, the steady-state DKFE (SDKFE) is given. Note that the computational complexity of the SDKFE is much lower than that of the DKFE. Moreover, a probability selection criterion for determining the dimensionality reduction strategy is also presented to guarantee the stability of the DKFE. Two illustrative examples are given to show the advantage and effectiveness of the proposed methods.

Index Terms—Bandwidth constraints, communication delays, cyber-physical systems (CPSs), distributed fusion estimation, Kalman filtering, stability analysis.

I. INTRODUCTION

INFORMATION fusion has attracted considerable research interest during the past decades, and has found applications in a variety of areas, including the Internet of Things [1]; sensor networks [2], [3]; and cyber–physical systems (CPSs) [4]. Particularly, multisensor fusion estimation utilizes useful information contained in multiple sets of data for the purpose of estimating a quantity or parameter in a process [5]. It is widely used in practical applications because it can potentially improve estimation accuracy and enhance reliability and robustness against faults [5]–[7]. Many fusion estimation approaches have been presented in the literature (see [8]–[13], and the references therein). At the same time, advances in embedded computing, communication, and related hardware technologies have recently brought the paradigm of CPSs to a new research frontier [14]. Moreover, CPSs have found applications in a broad range of areas, such as intelligent transportation systems [15], multirobot systems [16], and smart grid systems [17]. As one of the important issues in CPSs, real-time state estimation based on sensor measurements has recently attracted remarkable research interests because state estimate can provide a CPS with the real-time monitoring and control capability [18], [19]. For example, estimating the real voltage from sensor information must be completed before taking certain actions to regulate the voltage into some desired range in a power grid [17]. It is noted that the accuracy of state estimation has an important impact on computing control commands for safe and efficient operation of a CPS [18]–[21]. Therefore, it is of theoretical significance and practical relevance to investigate the problem of information fusion estimation for the CPSs [22], [23].

There mainly exist two kinds of fusion architectures: 1) centralized fusion structure and 2) distributed fusion structure. However, the distributed fusion structure is generally more robust and fault-tolerant as compared with the centralized fusion structure [6]–[10]. This motivates us to consider the distributed fusion estimation problem in this article for a class of CPS architecture (see Fig. 1), where system state is spatially distributed in the physical space. When the local estimates are transmitted to the fusion center (FC) via communication channels, bandwidth constrains and communication delays are unavoidable in communication networks [24]. Moreover, the above two factors can degrade the fusion...
estimation performance because of the information loss caused by bandwidth and delay constrains [25]–[27]. Thus, how to design distributed fusion methods in the presence of bandwidth and delay constraints is essential for real-time state estimate of CPSs.

A. Related Work

When considering the problem of bandwidth constraints in multisensor systems, as pointed out in [28], there are mainly two approaches to reduce the communication traffic: 1) the quantization method (see [29]–[31] and the references therein) and 2) the dimensionality reduction method (see [32]–[34], and the references therein). Particularly, by analyzing the statistical property of measurement information and resorting to the principal component analysis method, the dimensionality reduction strategy has been designed in [35] to compress the measurement signals, while the dimensionality reduction strategy with the quantization error was developed in [36] to deal with stable multisensor fusion systems. Notice that the dimensionality reduction method in [35] requires to know the global measurement matrix that may be difficult to be satisfied in distributed systems, and solving nonconvex nonlinear optimization problem under this strategy may also add the computational cost and design difficulty. Under the distributed fusion structure, when the physical state $x(t)$ as shown in Fig. 1 is multidimensional (or even high-dimensional) in a CPS, it is unrealistic to completely send the local estimate of the state $x(t)$ to the FC via a bandwidth-constrained communication channel. In this sense, bandwidth constraint in the CPSs is the primary consideration when designing a distributed fusion estimator. Notice that, to reduce the communication traffic, the idea of the dimensionality reduction method is to convert a multidimensional signal into a low-dimensional signal, while the idea of the quantization method is to reduce the number of coding bits for each component of a multidimensional signal before being transmitted. Meanwhile, the quantization usually results in nonlinear dynamics, and it is difficult to find a data compression operator analytically, particularly, for the multidimensional signals. Therefore, the dimensionality reduction method can provide an attractive alternative to solve the distributed fusion estimation problem with bandwidth constraints in the CPSs.

Though the dimensionality reduction fusion estimation algorithms have been proposed in [32]–[36] to reduce the communication traffic, the communication delays, which occur during the transmission, were not taken into account. With the communication delays, the dimensionality reduction fusion estimation must solve two challenging issues: one is how to compensate the information loss caused by the communication delays and bandwidth constraints under a unified mathematical model; The other one is how to fuse the asynchronous local compressed estimates caused by communication delays. Notice that the centralized and distributed fusion estimation algorithms have been proposed in [25] and [37]–[42] based on different communication delay models, however, the main results in [25] and [37]–[42] cannot be extended to the case of the dimensionality reduction estimation with communication delays. The reason is that the data compression and information compensation in dimensionality reduction may change the property of the original measurements (e.g., the statistical correlation in [32] and [34] has been changed under the Kalman fusion structure). Under this case, we have studied the information fusion estimation problem in [22] and [26] for the CPSs with bandwidth constraints and communication delays. It should be pointed out that the steady-state fusion estimator with simple calculation cannot be obtained based on the proposed communication model in [22], while the covariance intersection (CI) fusion strategy in [26] was suboptimal because fusion estimator was determined by minimizing an upper bound of estimation error covariance.

B. Contributions

Motivated by the aforementioned analysis, we study the distributed stochastic dimensionality reduction fusion estimation problem with communication delays for the CPSs. Notice that the information loss is inevitable because of the dimensionality reduction and communication delays, and such a fusion estimation with incomplete information will degrade the estimation performance. Since the delays are caused by communication channels, the key issue is how to design an efficient dimensionality reduction strategy to guarantee the stability of the distributed fusion estimator. Although our previous works in [22], [26], and [32] have studied the related stochastic dimensionality fusion estimation problems, there are still fundamental problems that cannot be solved up to now. In detail:

1) when only considering the stochastic dimensionality reduction strategy, the stable probability selection criterion in [32] was derived from the inequality relaxation of the matrix trace. However, the inequality relaxation will lead to certain conservatism; thus, how to find a new derivation idea to reduce the conservatism is very important for the application of the proposed dimensionality reduction strategy. Note that the stability conditions in [22] were directly derived from a similar derivation in [32] and, thus, the corresponding conservatism also cannot be avoided in [22]. Moreover, the stability...
conditions in [26] also have certain conservatism due to the inequality relaxation of 1-norm and ∞-norm;
2) when considering the stochastic dimensionality reduction strategy under communication delays, the distributed CI fusion estimator in [26] was suboptimal because the corresponding optimization objective was an upper bound of the estimation error covariance matrix. Particularly, the CI fusion results in [26] required to solve nonconvex nonlinear optimization problems online at each time, which may lead to a large number of calculations. Though the distributed fusion estimator in [22] was optimal under the optimal weighted fusion criterion, the model of communication delays cannot be applicable to the case of time-varying delays. More importantly, the computational complexity of the fusion estimator in [22] was also high. Obviously, the common disadvantage of the results in [22] and [26] is the high computational cost, and the optimal weighed fusion criterion can provide the optimal and analytic solutions. Therefore, based on the optimal weighted fusion criterion, how to design steady-state dimensionality reduction fusion estimators with simple calculation is of great significance in the presence of communication delays.

We shall solve the above two problems, and the main contributions of this article can be summarized as follows.

1) By constructing a new common orthogonal space, the cross-covariance matrix is calculated by the recursive form, and then an optimal distributed Kalman fusion estimator (DKFE) is derived in the linear minimum variance sense when there are bandwidth and communication delay constraints in CPSs. Note that each weighting fusion matrix is calculated by the analytic form.

2) A delay-dependent and probability-dependent stability condition is derived such that the fusion estimation error covariance matrix of the DKFE converges to a unique steady-state matrix for any initial values. Under this condition, the steady-state DKFE (SDKFE), which has much lower computational complexity as compared with the DKFE, is given. Moreover, when each communication delay is known, the probability selection criterion for determining the dimensionality reduction strategy is presented to guarantee the stability of the DKFE.

3) Compared with the fusion estimation method in [22], the model of communication delays in this article does not require that each sink node knows the communication delay in advance, and the SDKFE with simple calculation is derived (see Remark 1). Since the CI fusion criterion in [26] is suboptimal, the estimation performance of the designed DKFE must be better than that of the fusion estimator in [26] when each communication delay is constant. Moreover, the computational cost of the SDKFE must be lower than that of the CI fusion estimator in [26] and the conservatism of delay-dependent stability conditions is less than that of the conditions in [26] (see Remarks 2 and 5).

4) When there is no communication delay for the scenario described in Fig. 1, it is shown that the stability condition in this article has less conservatism than the result in [32]. This is because a new derivation idea without any inequality relaxation is proposed to design the stochastic dimensionality reduction strategy. Moreover, when considering communication delays, the corresponding stability analysis is also based on this new derivation idea. Note that it is difficult to obtain the stability condition by using the derivation idea in [32] when the communication delay is modeled in this article (see Remarks 5 and 6).

Notations: The notations used throughout this article are fairly standard. The superscript “T” represents the transpose, and $E[\cdot]$ is the mathematical expectation. $I_m$ represents the identity matrix of size $m \times m$, while $\text{diag}{}[\cdot]$ stands for a block diagonal matrix. $\text{Prob}[A]$ means the occurrence probability of event $A$, while $\text{Tr}(B)$ denotes the trace of the matrix $B$. $||A||_2$ represent the 2-norm of the matrix $A$. $x \perp y$ denotes that $x$ and $y$ are orthogonal vectors, and $\text{col}\{a_1, \ldots, a_L\}$ represents the column vector that is composed of the elements $a_1, \ldots, a_L$. The symbol $\text{lcm}(a, b)$ is the least common multiple of $a$ and $b$, while $\text{rank}(A)$ denotes the rank of matrix $A$. The function $f_0^b(t)$ is defined by $f_0^b(t) \triangleq f(\cdots(f(t)\cdots))$, and $X \succ (\prec)0$ denotes a positive-definite (negative-definite) matrix.

II. PROBLEM FORMULATION

A. Dimensionality Reduction and Communication Delays

Consider the physical process in Fig. 1 described by the following discrete state-space model:

$$x(t+1) = Ax(t) + w(t)$$

where $x(t) \in \mathbb{R}^n$ ($n > 1$) is the state of the process, $w(t)$ is the system noise, and $A$ is a constant matrix with appropriate dimension. As pointed out in [19], model (1) is widely adopted for describing the state dynamics of CPSs, including power systems, smart grid infrastructures, and building automation systems, etc. When the measurements from each sensor are sent to sink nodes, the $i$th sink node’s measurement $y_i(t) \in \mathbb{R}^{n_i}$ is modeled by

$$y_i(t) = C_i x(t) + v_i(t) (i = 1, 2, \ldots, L)$$

where $C_i$ is the measurement matrix with appropriate dimension, and $v_i(t)$ is the measurement noise. Moreover, $w(t)$ and $v_i(t)$ are uncorrelated zero-mean Gaussian white noises satisfying

$$E\left\{[w^T(t) \; v_i^T(t)]^T[w^T(t_1) \; v_i^T(t_1)]\right\} = \delta_{t,t_1} \text{diag}(Q_w, \delta_{i,j} Q_{v_i})$$

where $\delta_{t,t_1}$ is defined by

$$\delta_{t,t_1} = \begin{cases} 1, & \text{if } t = t_1 \\ 0, & \text{if } t \neq t_1. \end{cases}$$

Then, based on the measurements $\{y_1(t), \ldots, y_L(t)\}$, the local optimal estimate (LOE) $\hat{x}_i(t)$ is given by the Kalman filter

$$\hat{x}_i(t) = G_{K_i}(t)A\hat{x}_i(t-1) + K_i(t)y_i(t)$$

$$x(t+1) = Ax(t) + w(t) \quad (1)$$

where $x(t) \in \mathbb{R}^n$ ($n > 1$) is the state of the process, $w(t)$ is the system noise, and $A$ is a constant matrix with appropriate dimension. As pointed out in [19], model (1) is widely adopted for describing the state dynamics of CPSs, including power systems, smart grid infrastructures, and building automation systems, etc. When the measurements from each sensor are sent to sink nodes, the $i$th sink node’s measurement $y_i(t) \in \mathbb{R}^{n_i}$ is modeled by

$$y_i(t) = C_i x(t) + v_i(t) (i = 1, 2, \ldots, L) \quad (2)$$

where $C_i$ is the measurement matrix with appropriate dimension, and $v_i(t)$ is the measurement noise. Moreover, $w(t)$ and $v_i(t)$ are uncorrelated zero-mean Gaussian white noises satisfying

$$E\left\{[w^T(t) \; v_i^T(t)]^T[w^T(t_1) \; v_i^T(t_1)]\right\} = \delta_{t,t_1} \text{diag}(Q_w, \delta_{i,j} Q_{v_i}) \quad (3)$$

where $\delta_{t,t_1}$ is defined by

$$\delta_{t,t_1} = \begin{cases} 1, & \text{if } t = t_1 \\ 0, & \text{if } t \neq t_1. \end{cases} \quad (4)$$

Then, based on the measurements $\{y_1(t), \ldots, y_L(t)\}$, the local optimal estimate (LOE) $\hat{x}_i(t)$ is given by the Kalman filter

$$\hat{x}_i(t) = G_{K_i}(t)A\hat{x}_i(t-1) + K_i(t)y_i(t) \quad (5)$$
where
\[ G_K(t) \triangleq I_n - K(t)C_i. \] (6)

Define \( \bar{x}_i(t) \triangleq x(t) - \hat{x}_i(t) \). Then, the optimal gain matrix \( K_i(t) \) and the local estimation error covariance matrix \( P_{ii}(t) \triangleq E[\bar{x}_i(t)\bar{x}_i^T(t)] \) are calculated by
\[
\begin{align*}
K_i(t) &= P_{ii}^* C_i^T [C_i P_{ii}^* C_i^T + Q_i]^{-1} \\
P_{ii}(t) &= G_K(t) P_{ii}^* (t)
\end{align*}
\] (7)

where \( P_{ii}^* \) denotes the error covariance matrix of one-step prediction. Moreover, it follows from (1), (5), and (7) that the local estimation error cross-covariance matrix \( P_{ij}(t) \triangleq E[\bar{x}_i(t)\bar{x}_j^T(t)](i \neq j) \) is calculated by
\[
P_{ij}(t) = G_K(t)[Q_w + AP_{ij}(t-1)A^T]G_K^T(t). \] (8)

Under the distributed fusion structure, each LOE \( \hat{x}_i(t) \) must be sent to the FC to design an optimal fusion estimator. However, it is unrealistic to send the complete information included in \( \hat{x}_i(t) \in \mathbb{R}^n \) to the FC over communication networks because almost all communication networks can only carry a finite amount of information per unit time. This problem is especially prominent in the fusion estimation for the large-scale CPSs integrated by wireless sensor networks.

To reduce communication traffic, only \( r_i(1 \leq r_i < n) \) components of the \( i \)th LOE \( \hat{x}_i(t) \) are allowed to be transmitted to the FC at each time, and other components are discarded. Compared with the original LOE \( \hat{x}_i(t) \), the dimension of the transmitted signal is reduced. In this sense, the above method can be viewed as one of the dimensionality reduction strategies. According to this dimensionality reduction strategy, the allowed sending components (ASCs) of \( \hat{x}_i(t) \) have \( \Delta_i \) possible cases, where \( \Delta_i = \prod_{j=1}^{r_i} (n - \ell^*_j) / \prod_{j=1}^{r_i} \ell^*_j \). Then, at a particular time, only one vector signal, which is taken from one of the above \( \Delta_i \) cases, is selected and transmitted to the FC, and this selected signal is denoted by \( \hat{x}_{ni}(t) \in \mathbb{R}^n \).

When \( \hat{x}_i(t) \) is sent to the FC by the sink node, the FC will receive the data packet containing \( \hat{x}_i(t) \) at time \( t + d_i \) because of communication delay. Let \( \bar{x}_{ni}(t) \) denote the local estimation information received by the FC at time \( t \). Then, \( \bar{x}_{ni}(t) \) in the FC is given by
\[
\bar{x}_{ni}(t) = \hat{x}_{ni}(t - d_i). \] (9)

It should be pointed out that the communication delays for different sink nodes are not the same (i.e., \( d_i \neq d_j \)), which means that the addressed delays are not constant from the perspective of the whole fusion systems. Moreover, the constant communication delay for each sink node is mainly determined by signal transmission power, and is difficult to avoid due to resource constraints. Particularly, when the FC received the signals from multiple sink nodes, there must exist the problem of resource scheduling. Though time-varying delays may occur for different sink nodes, they must lead to the information disorder that directly adds the design difficulties of dimensionality reduction fusion estimators. In this case, to avoid the information disorder in multisensor fusion, resource scheduling can be controlled by designing the physical mechanism for guaranteeing constant delays (see an example in Remark 1). Therefore, it is of significance in studying the dimensionality reduction fusion estimation problem with constant delays in this article. Up to now, the problem of dimensionality reduction and communication delays has been presented, and the process diagram is shown in Fig. 2.

It is noted that the signal \( \hat{x}_{ni}(t) \) only takes one element from the following finite set:
\[
S_i(t) = \left\{ \hat{x}_{ni}(t) | h_i = 1, 2, \ldots, \Delta_i \right\} \] (10)
where \( \hat{x}_{ni}(t) \in \mathbb{R}^n \) represents one group of ASCs. To characterize the determining process of \( \hat{x}_{ni}(t) \), we introduce the following indicator functions:
\[
\sigma_{h_i}^j(t) = \begin{cases} 1, & \text{if } \hat{x}_{ni}(t) = \hat{x}_{ni}^{h_i}(t) \\ 0, & \text{if } \hat{x}_{ni}(t) \neq \hat{x}_{ni}^{h_i}(t) \end{cases} \] (11)

where \( \sigma_{h_i}^j(t) \) are required to satisfy
\[
\sigma_{h_i}^j(t) \sigma_{h_i}^{j'}(t) = 0 | h_i \neq h_i' | \sum_{h_i=1}^{\Delta_i} \sigma_{h_i}^j(t) = 1 \] (12)
such that \( \hat{x}_{ni}(t) \) only takes one ASC from the set (10) at time \( t \), that is
\[
\hat{x}_{ni}(t) = \sum_{h_i=1}^{\Delta_i} \sigma_{h_i}^j(t) \hat{x}_{ni}^{h_i}(t). \] (13)

Then, it is derived from (9) and (13) that
\[
\bar{x}_{ni}(t) = \sum_{h_i=1}^{\Delta_i} \sigma_{h_i}^j(t - d_i) \hat{x}_{ni}^{h_i}(t - d_i). \] (14)

At time \( t \), if the fusion estimate of \( x(t) \) is directly designed based on \( \bar{x}_{ni}(t) \), the fusion estimation performance must be poor because of the communication delays and the untransmitted component of \( \hat{x}_{ni}(t) \). In this case, the compensating state estimate (CSE) of \( x(t) \), denoted by \( \bar{x}_{\tilde{x}}^c(t) \), can be modeled as follows:
\[
\bar{x}_{\tilde{x}}^c(t) = A^d_{h_i}(t - d_i) \bar{x}_{ni}(t - d_i) \\
+ A^d_i [I_n - H_i(t - d_i)] \bar{x}_{\tilde{x}}^c(t - d_i - 1) \] (15)
where \( H_i(t - d_i) \) is determined by
\[
H_i(t) = \sum_{h_i=1}^{\Delta_i} \sigma_{h_i}^i(t) H_{h_i}^i = \text{diag}\{y_1^i(t), \ldots, y_n^i(t)\}. \tag{16}
\]

Here, \( H_{h_i}^i \) represents a diagonal matrix that contains \( r_i \) diagonal elements “1” and \( n - r_i \) diagonal elements “0.” Then, it follows from (11) and (12) that:
\[
y_i^i(t) \in \{0, 1\}, \sum_{\ell=1}^n y_i^\ell(t) = r_i (i = 1, \ldots, L) \tag{17}
\]

where \( y_i^\ell(t) = 1 \) means that the \( \ell \)-th component of \( \hat{x}_i(t) \) is selected and sent to the FC, while \( y_i^\ell(t) = 0 \) means that the \( \ell \)-th component of \( \hat{x}_i(t) \) is discarded. Particularly, at time \( t \), the compensation strategy in the CSE model (15) is reflected by the following aspects.

1) The untransmitted components of \( \hat{x}_i(t) \) are compensated by the one-step prediction based on \( \hat{x}_i^\ell(t - d_i - 1) \).
2) The delayed information \( \tilde{x}_i(t) \) is compensated by the one-step prediction based on \( \hat{x}_i^\ell(t - d_i) \), where \( \hat{x}_i^\ell(t - d_i) = H_i(t - d_i) \hat{x}_i(t - d_i) + [I - H_i(t - d_i)]A \hat{x}_i^\ell(t - d_i - 1) \).

Remark 1: In [22], at the \( i \)-th sink node, the one-step prediction based on the local estimate \( \hat{x}_i(t) \) was given by \( \hat{x}_i^d(t) = A \hat{x}_i^d(t - 1) \). Due to the bandwidth constraints, only \( r_i \) components of \( \hat{x}_i^d(t) \) were allowed to be sent. Then, the CSE of \( x(t) \), denoted as \( \hat{x}_i^d(t) \), was given by (i.e., the model [22, eq. (18)])
\[
\hat{x}_i^d(t) = H_i(t - d_i) A \hat{x}_i^d(t - d_i) + [I - H_i(t - d_i)] A \hat{x}_i^d(t - 1) \tag{18}
\]

where the definition of \( H_i(t - d_i) \) is the same as that of \( H_i(t - d_i) \), and \( \hat{x}_i^d(t - 1) \) denotes the fusion estimate designed by Chen et al. [22]. For the CSE model (18), the one-step prediction \( \hat{x}_i^d(t) \) must be completed at the sink node, which implies that each sink node must know the communication delay from the sink node to the FC in advance. Under this case, when the communication delay is unknown for the sink node or time-varying, the model (18) will be invalid. Different from the modeling method in [22], the CSE model (15) does not require that each sink node knows the communication delay in advance, and thus the model (15) can be more easily implemented in a practical system. Particularly, when considering the time-varying communication delay \( d_i(t) \), the local estimation information received by the FC, denoted as \( \tilde{x}_i^d(t) \), is given by
\[
\tilde{x}_i^d(t) = \hat{x}_i^d(t - d_i(t)) \tag{19}
\]

where \( \hat{x}_i^d(t) \) denotes the selected ASC at the sink node. Meanwhile, it is reasonable to consider that the time-varying delay \( d_i(t) \) is bounded in practical applications, and satisfies \( d_i(t) \leq d_i^c \). Then, by resorting to the buffers at the FC, each time-varying delay can be prolonged to its upper bound \( d_i^c \) at each time, that is, model (19) is reduced to
\[
\tilde{x}_i^d(t) = \hat{x}_i^d(t - d_i^c). \tag{20}
\]

Since the structure of (20) is the same as that of (9), the case of time-varying delays can still be modeled by (15). Notice that the CSE model (18) in [22] will not be applicable to this case, because the time-varying communication delays are only known to the FC, and each sink node impossibls knows the time-varying delays a priori. On the other hand, the stability condition in [22] could only guarantee the MSE of the fusion estimator converged to a steady-state value. It should be pointed out that the computational complexity of the fusion estimator in [22] is slightly high, yet the corresponding steady-state fusion estimator cannot be derived from the stability condition in [22]. In contrast, the SDKFE with simple calculation can be designed based on the stability condition in Theorem 3.

Remark 2: For the case of time-varying delays, the estimation error cross-covariance matrices cannot be obtained under the dimensionality reduction strategy in this article. Fortunately, the CI fusion criterion does not need the cross-covariance matrices. Therefore, the distributed CI fusion estimation algorithm was developed in [26] to deal with the time-varying delays. Note that the CI fusion criterion is not optimal because the optimization objective is an upper bound of estimation error covariance matrix, and each weighting matrix is obtained by solving nonconvex nonlinear optimization problems at each time. Different from the fusion criterion in [26], the optimal weighted fusion criterion with analytic solutions is used to design the DKFE in this article. Thus, when considering the constant communication delays, the estimation performance of the DKFE is better than that of the fusion estimator in [26]. On the other hand, as pointed out in Remark 1, the designed fusion estimation algorithms in this article can be also applicable to the case of time-varying communication delays. However, it is difficult to show whose estimation performance is better, the DKFE in this article or the fusion estimator in [26], when dealing with time-varying delays. This is because the performance loss in this article is introduced from the delay model (i.e., prolonging the time-varying delay to its upper bound at each time), while the performance loss in [26] is introduced from the CI fusion criterion (i.e., minimizing an upper bound of the fusion estimation error covariance). However, from the perspective of computational complexity, the SDKFE in this article is better than the CI fusion estimator in [26] whenever considering the constant delays or time-varying delays.

B. Problem of Interest

From (13), the selected ASC \( \hat{s}_h(t) (\in R^n) \) at the sink node is determined by the binary variables \( \sigma_{h_i}^i(t) \) (\( h_i = 1, 2, \ldots, \Delta_i \)). Meanwhile, it is known from (15) that the design of optimal \( \sigma_{h_i}^i(t) \) (\( h_i = 1, 2, \ldots, \Delta_i \)) must be completed at the FC, because the communication delay (from the sink node to the FC) and each CSE \( \hat{x}_i(t) \) are only obtained by the FC, but these information are unknown to each sink node. Therefore, an optimal \( \hat{s}_h(t) \) may be difficult to be designed at the sink node. Based on the above consideration, let each binary variable \( \sigma_{h_i}^i(t) \) be generated in a random way at the sink node, and let random variables \( \{\sigma_1^i(t), \sigma_2^i(t), \ldots, \sigma_{\Delta_i}^i(t)\} \) obey the categorical distribution.
satisfying
\[
E\left\{ \sigma_{h_i}^j(t)\sigma_{h_i}^j(t) \right\} = \begin{cases} 
\delta_{h_i,h_j}E\left\{ \sigma_{h_i}^j(t) \right\}, & \text{if } i = j, t = t_1 \\
E\left\{ \sigma_{h_i}^j(t) \right\}E\left\{ \sigma_{h_i}^j(t_1) \right\}, & \text{if } i = j, t \neq t_1 \\
E\left\{ \sigma_{h_i}^j(t) \right\}E\left\{ \sigma_{h_i}^j(t_1) \right\}, & \text{if } i \neq j.
\end{cases}
\]

(21)

Under this case, a group of ASC \( \hat{x}_{h_i}^j(t) (\forall i \in [1, 2, \ldots, L]) \) in set (10) is randomly selected as the \( \hat{x}_{h_i}^j(t) \) at time \( t \). Moreover, the occurrence probabilities of the cases \( \sigma_{h_i}^j(t) = 1 \) and \( \sigma_{h_i}^j(t) = 0 \) are given by \( \text{Prob}[\sigma_{h_i}^j(t) = 1] = \pi_{h_i} \) and \( \text{Prob}[\sigma_{h_i}^j(t) = 0] = 1 - \pi_{h_i} \), where the selection probability \( \pi_{h_i} \geq 0 \) satisfies
\[
\sum_{h_i=1}^{\Delta_i} \pi_{h_i} = 1 \quad (i \in [1, 2, \ldots, L]).
\]

(22)

Then, it is concluded from (16) and (21) that the binary variables \( \gamma_{\ell}^j(t) (\ell = 1, \ldots, n) \) in (17) are independent Bernoulli distributed white noise sequences with \( \text{Prob}[\gamma_{\ell}^j(t) = 1] \triangleq \gamma_{\ell}^j \) and \( \text{Prob}[\gamma_{\ell}^j(t) = 0] \triangleq 1 - \gamma_{\ell}^j \), which yield
\[
H_i \triangleq E[H_i(t - d_i)] = \text{diag}\{\gamma_1^j, \gamma_2^j, \ldots, \gamma_n^j\}. \quad (23)
\]

From (16) and (23), there must exist a constant matrix \( U_{\ell} \in R^{n \times \Delta_i} \) such that
\[
\gamma_{\ell}^j = U_{\ell} \xi_i^j (\ell = 1, \ldots, n; t = 1, \ldots, L)
\]

(24)

where \( \xi_i^j \triangleq \text{col}[\pi_{h_1}^j, \pi_{h_2}^j, \ldots, \pi_{h_{\Delta_i}}^j] \). This means that when each selection probability \( \pi_{h_i}^j \) is given by (21) and (22), \( \gamma_{\ell}^j \) in (23) will be determined by (24). Note that the selection probabilities \( \xi_i^j (i = 1, 2, \ldots, L) \) are to be designed in this article for guaranteeing the stability of the DKFE.

Let \( \hat{x}_{h_i}^j(t) \triangleq x(t) - \hat{x}_{h_i}^j(t) \) denote the estimation error of each CSE. Then, it follows from (1) and (15) that:
\[
\hat{x}_{h_i}^j(t) = A^dH_i(t - d_i)\hat{x}_i(t - d_i) \\
+ A^d[I_n - H_i(t - d_i)]A\hat{x}_i^j(t - d_i - 1) \\
+ A^d[I_n - H_i(t - d_i)]w(t - d_i - 1) + F_w(d_i, t)
\]

(25)

where \( F_w(d_i, t) \) is determined by the following function:
\[
F_w(g, t) \triangleq \sum_{\theta=1}^{g} A^{g-\theta-1}w(t-\theta).
\]

(26)

When \( E[\hat{x}_i^j(-d_n)] = E[x(-d_n)](d_n = 0, 1, \ldots, d) \), it is concluded from (3) and (25) and the fact \( E[x(t)] = E[\hat{x}_i(t)] \) that each CSE \( \hat{x}_{h_i}^j(t) \) is unbiased, that is
\[
E[\hat{x}_{h_i}^j(t)] = E[x(t)](i = 1, 2, \ldots, L).
\]

(27)

According to the CSEs \( \hat{x}_{h_i}^j(t) (i = 1, 2, \ldots, L) \) in the FC, the DKFE for the addressed CPSs is given by
\[
\hat{x}(t) = \sum_{i=1}^{L} \Omega_i(t)\hat{x}_{h_i}^j(t)
\]

(28)

where \( \sum_{i=1}^{L} \Omega_i(t) = I_n \), and combining (27) yields that the DKFE \( \hat{x}(t) \) is unbiased if \( E[\hat{x}_i^j(d_n)] = E[x(d_n)](d_n = 0, 1, \ldots, d) \). Consequently, the problems to be solved in this article are described as follows.

1) When the selection probabilities \( \pi_{h_i}^j (h_i = 1, \ldots, \Delta_i; i = 1, \ldots, L) \) satisfying (22) are given in advance, the aim is to design optimal weighting matrices \( \Omega(t), \ldots, \Omega_L(t) \) such that the MSE of the DKFE \( \hat{x}(t) \) is minimal at each time step, that is
\[
\{\Omega_1(t), \ldots, \Omega_L(t)\}
= \arg \min_{\sum_{i=1}^{L} \Omega_i(t) = I} E\left\{ [x(t) - \hat{x}(t)]^T [x(t) - \hat{x}(t)] \right\}.
\]

(29)

2) Find stability conditions, which are dependent on the communication delay \( d_i \) in (9) and the selection probability \( \pi_{h_i}^j \) in (22), such that the estimation error covariance matrix of the DKFE converges to a unique positive matrix, that is
\[
\lim_{t \to \infty} E\left\{ [x(t) - \hat{x}(t)]^T [x(t) - \hat{x}(t)] \right\} = P
\]

(30)

and \( P \) is independent of the initial values.

Remark 3: When the \( i \)-th sink node knows the selection probability \( \xi_i^j \) in advance, the binary variables \( \sigma_{h_i}^j(t)(h_i = 1, \ldots, \Delta_i) \) obeying the categorical distribution will be randomly generated at each time step, and then the selected ASC \( \hat{x}_i(t) \) can be determined by (13) at the sink node. Under this case, one of the important issues in this article is how to design the satisfactory probability selection criteria, which will be solved in Section IV. On the other hand, when result (30) holds, the limit of each weighting matrix \( \Omega(t) \) must exist, and will be independent of the initial values. This is because the estimation error covariance matrix of the DKFE is dependent on each time-varying matrix \( \Omega(t) \). In such a case, the SDFKFE with simple calculation will be given in this article.

III. Finite-Horizon DKFE for the CPSs

In this section, the recursive DKFE will be derived by using the optimal fusion criterion weighted by matrices in the linear minimum variance sense. Define \( \hat{x}(t) \triangleq x(t) - \hat{x}(t) \) and \( I_a = \text{col}[I_n, \ldots, I_n] \in R^{d \times \Delta_n} \). Then, from the results in [9] and [10], the optimal weighting matrices \( \Omega_1(t), \ldots, \Omega_L(t) \) in (29) and the corresponding fusion estimation error covariance matrix \( P(t) \triangleq E[\hat{x}(t)\hat{x}(t)^T] \) can be calculated by
\[
[\Omega_1(t), \Omega_2(t), \ldots, \Omega_L(t)] = \left(I_a^T \Sigma^{-1}(t)I_a \right)^{-1} I_a^T \Sigma^{-1}(t)
\]

(31)

\[
P(t) = \left(I_a^T \Sigma^{-1}(t)I_a \right)^{-1}
\]

(32)

where the weighting matrices \( \Omega_i(t)(i = 1, 2, \ldots, L) \) determined by (31) satisfy the constraint \( \sum_{i=1}^{L} \Omega_i(t) = I_n \), and
\[
\Sigma(t) = \left[\Sigma_y(t)\right]_{nL \times nL}, \Sigma_y(t) = E\left\{ \hat{x}_i^j(t)^T \hat{x}_i^j(t) \right\}.
\]

(33)
It is concluded from (31) and (33) that if the computation procedure of $\Xi(t)$ is given, then the optimal weighting matrices $\Omega_i(t)$ ($i = 1, \ldots , L$) in (31) can be thus obtained.

In what follows, six lemmas will be given before deriving the recursive form of $\Xi(t)$. For notational convenience, the following indicator function is introduced:

$$
C_0(t_1, t_2) = \begin{cases} 1, & \text{if } t_1 > t_2 \\ 0, & \text{if } t_1 \leq t_2. \end{cases}
$$

(34)

Meanwhile, if $t_1 > t_2$, it will be specified that $\prod_{t_1}^{t_2} F(t) = I_m$ and $\sum_{t_1}^{t_2} G(t) = 0$, where $F(t) \in R^{m \times m}$ and $G(t) \in R^{n \times m}$ represent different matrix functions with respect to the variable $t$.

**Lemma 1** [32]: For stochastic matrices $U$, $B$, and $G$, where

$$
\begin{align*}
U & \triangleq \text{diag}[u_1, \ldots , u_n], \quad B \triangleq \text{diag}[b_1, \ldots , b_n] \\
G & \triangleq \begin{bmatrix}
g_{11} & \cdots & g_{1n} \\
\vdots & \ddots & \vdots \\
g_{n1} & \cdots & g_{mn}
\end{bmatrix}.
\end{align*}
$$

If each random variable $g_{ij}$ in $G$ is independent of any random variables of $u_k$ and $b_k$ ($k = 1, 2, \ldots , n$), then

$$
E(UGB) = E[U \circ B] \otimes E[G]
$$

where “$\otimes$” is defined as $[G^1 \otimes G^2]_{ij} = G^1_{ij}G^2_{ij}$, and the product “$\circ$” for the matrices $U$ and $B$ is defined by

$$
U \circ B = \begin{bmatrix}
u_1 b_1 & \cdots & \nu_1 b_n \\
\vdots & \ddots & \vdots \\
u_n b_1 & \cdots & \nu_n b_n
\end{bmatrix}.
$$

**Lemma 2**: Define

$$
\begin{align*}
\Phi_x(t_1, t_2) & \triangleq \mathbb{E}\{\tilde{x}(t_1)w^T(t_2)\} \\
\Phi^w_x(t_1, t_2) & \triangleq \mathbb{E}\{\tilde{x}(t_1)\tilde{x}^T(t_2)\} \\
\Phi^F_x(t_1, g, t_2) & \triangleq \mathbb{E}\{F_w(g, t_1)w^T(t_2)\}
\end{align*}
$$

(35)

where $G(t)$ and $F_w(g, t_2)$ are determined by (6) and (26), respectively. Then, $\Phi^w_x(t_1, t_2)$, $\Phi^w_x(t_1, t_2)$, $\Phi^F_x(t_1, g, t_2)$, and $\Phi^F_x(t_1, g, t_2)$ are given by

$$
\begin{align*}
\Phi^w_x(t_1, t_2) &= C_0(t_1, t_2) \left\{ \prod_{\varphi_i = 0}^{t_1 - t_2 - 2} \Phi_x(t_1 - \varphi_i) \right\} \\
& \times \Phi_x(t_1, t_2 + 1) \Phi^w_x(t_1, \varphi_i) \\
& \times \Phi^w_x(t_1 + 1, \varphi_i) \\
& \times \Phi^F_x(t_1, g, t_2) = \sum_{\varphi_i = 0}^{g} \Phi^w_x(t_1, t_2 - \varphi_i) \left[ A^{\varphi_i - 1} \right]^{T}
\end{align*}
$$

(36)

$$
\begin{align*}
\Phi^F_x(t_1, g, t_2) &= C_0(t_1, t_2)C_0(t_2, t_1 - g - 1)A^{t_2 - 1}Q_w
\end{align*}
$$

(37)

$$
\begin{align*}
\Phi^F_x(t_1, g, t_2) &= C_0(t_1, t_2)C_0(t_2, t_1 - g - 1)A^{t_2 - 1}Q_w
\end{align*}
$$

(38)

$$
\begin{align*}
\Phi^F_x(t_1, g, t_2) &= C_0(t_1, t_2)C_0(t_2, t_1 - g - 1)A^{t_2 - 1}Q_w
\end{align*}
$$

(39)

where $\delta_{ij}$ and $C_0(t_1, t_2)$ are determined by (4) and (34), respectively. $P_{ij}(t_2)$ in (37) is calculated by (7) or (8).

**Proof**: See [46, Appendix A.1].

**Lemma 3**: Define

$$
\begin{align*}
f_i(t) & \triangleq t - d_i - 1, f_i^0(t) \triangleq t \\
\chi_i(t_1, t_2) & \triangleq \min \left\{ \chi_i(t_1, t_2), \chi_i(t_1, t_2) \right\} \\
\Theta^w_x(t_1, t_2) & \triangleq \mathbb{E}\{\tilde{x}(t_1)w^T(t_2)\} \\
\Theta^F_x(t_1, g, t_2) & \triangleq \mathbb{E}\{F_w(g, t_1)w^T(t_2)\}
\end{align*}
$$

(40)

Then, $\Theta^w_x(t_1, t_2)$, $\Theta^F_x(t_1, g, t_2)$, and $\Theta^F_x(t_1, g, t_2)$ are given by

$$
\Theta^w_x(t_1, t_2) = C_0(t_1 - d_i, t_2)
$$

$$
\times \left\{ \sum_{h=0}^{\chi_i(t_1, t_2) - 1} \left\{ \delta_{h(t_1, t_2)} H_{Ad}^h \tilde{H}_{Ad} Q_w \right\} + \Phi^F_x(d_i, t_1, t_2) \right\}
$$

and

$$
\Theta^F_x(t_1, g, t_2) = \sum_{g=0}^{\chi_i(t_1, t_2) - 1} \left\{ \delta_{g(t_1, t_2)} H_{Ad} H_{Ad} \right\}
$$

$$
\times \left\{ \sum_{h=0}^{\chi_i(t_1, t_2) - 1} \left\{ \delta_{h(t_1, t_2)} H_{Ad}^h \tilde{H}_{Ad} Q_w \right\} + \Phi^F_x(d_i, t_1, t_2) \right\}
$$

(41)

**Lemma 4**: Define

$$
\begin{align*}
\Gamma_{ij}(t) & \triangleq \mathbb{E}\{\tilde{x}(t) \tilde{x}^T(t)\} \\
\Gamma_{ij}(t) & \triangleq \mathbb{E}\{\tilde{x}(t) \tilde{x}^T(t)\} (t_1 > t_2).
\end{align*}
$$

(42)

Then, $\Gamma_{ij}(t)$ is calculated by the following recursive form:

$$
\Gamma_{ij}(t) = \left( \prod_{i=0}^{d_i} \Phi_{K_i}(t - \varphi_i) \right) \Gamma_{ij}(t - d_j - 1)H_{Ad}^T
$$

$$
+ \Phi^F_{x_i}(t, t - d_j)H_{Ad}^T + \Phi^F_{x_i}(t, d_j, t)
$$

$$
+ \Phi^F_{x_i}(t, t - d_j - 1)H_{Ad}^T
$$

(43)

where $H_{Ad}$, $\tilde{H}_{Ad}$, and $H_{Ad}$ are given by (43), while $\Phi^w_x(t_1, t_2)$, $\Phi^F_x(t_1, g, t_2)$, and $\Phi^F_x(t_1, g, t_2)$ are computed by (36)–(38). In this case, $\Gamma_{ij}(t_1, t_2)$ is calculated by

$$
\Gamma_{ij}(t_1, t_2) = \left( \prod_{i=0}^{d_i} \Phi_{K_i}(t - \varphi_i) \right) \Gamma_{ij}(t_2).
$$

(44)

**Proof**: See [46, Appendix A.2].

**Lemma 5**: Define

$$
\Psi_{ij}(t) = \mathbb{E}\{\tilde{x}(t) \tilde{x}^T(t)\} (t_1 > t_2).
$$

(45)

**Proof**: See [46, Appendix A.3].
For $i = j$, $\Psi_i(t)$ is calculated by
\[
\Psi_i(t) = \Phi_{K_i}(t - d_i) \Gamma_{ij}(t - d_i - 1). \tag{48}
\]
For $i \neq j$, let $\eta_{ij} \triangleq \min(\eta_j, \eta_j + d_j + 1) - d_j \geq 0$. Then, $\Psi_i(t)$ is calculated by
\[
\Psi_i(t) = \sum_{k=1}^{\eta_{ij}-1} \left\{ \Phi_{\tilde{X}_i}(t - d_i, f_{ji}^k(t) - d_j) H_{ji}^T + \Phi_{\tilde{Y}_i}(t - d_i, f_{ji}^{k+1}(t)) \bar{H}_{Adij} + \Phi_{\tilde{F}_i}(t - d_i, d_j, f_{ji}^k(t)) \right\} (H_{Adij}^{-1})^T
+ \left( \prod_{\tilde{Y}_i = 0}^{\eta_j(d_j + 1) - d_i} \Phi_{\tilde{K}_i}(t - d_i - \tilde{Y}_i) \right) \times \Gamma_{ij}(f_{ji}^{\eta_{ij}}(t)) (H_{Adij}^{\eta_{ij}-1})^T \tag{49}
\]
where $H_{Adij}$ and $\bar{H}_{Adij}$ are given by (43); $\Gamma_{ij}(t)$ is computed by (45), while $\Phi_{\tilde{X}_i}(t - d_i, f_{ji}^k(t) - d_j)$, $\Phi_{\tilde{Y}_i}(t - d_i, f_{ji}^{k+1}(t))$, and $\Phi_{\tilde{F}_i}(t - d_i, d_j, f_{ji}^k(t))$ are calculated by (36)–(38).

**Proof:** See [46, Appendix A.4].

**Lemma 6:** Define
\[
\begin{align*}
\tau_{ij} &\triangleq \text{lcm}(d_i + 1, d_j + 1), \quad \tau_{dij} \triangleq \tau_{ij}/(d_i + 1) \\
\tilde{\tau}_{ij} &\triangleq \text{lcm}(d_i + 1, d_j + 1), \quad \tilde{\tau}_{dij} \triangleq \tilde{\tau}_{ij}/(d_i + 1) \\
\theta_{ij} &\triangleq \text{lcm}(d_i + 1, d_j + 1), \quad \theta_{dij} \triangleq \theta_{ij}/(d_i + 1)
\end{align*}
\tag{50}
\]
Then, $\tilde{\theta}_{ij}(t)$ can be calculated by (36)–(38) (see Lemma 2), while $\tilde{\theta}_{dij}(t)$ can be calculated by (41) and (42) (see Lemma 3) and (46). In this case, $\tilde{\theta}_{ij}(t)$ is calculated by
\[
\begin{align*}
\tau_{ij}(t) &= H_{Adij}^\tau \hat{\Xi}_i(t - \tau_{ij}) \left( H_{Adij}^\tau \right)^T + \hat{\tilde{\xi}}_{ij}(t) \\
\tilde{\theta}_{ij}(t) &= \left( 1 - \delta_{\tau_{ij}} \right) \left( 1 - \delta_{\tau_{dij}} \right) \sum_{\tilde{\theta}_{ij}(t)} \left( H_{Adij}^{\tau_{ij}} \right)^T \\
+ \left( 1 - \delta_{\tau_{ij}} \right) \sum_{\tilde{\theta}_{ij}(t)} \left( H_{Adij}^{\tau_{ij}} \right)^T
\end{align*}
\tag{52}
\]
where $\delta_{\tau_{ij}}$ is determined in (44), and
\[
\begin{align*}
\sum_{1i} &= H_{di}, H_{Adi}H_{di}, \ldots, H_{Adi}^{d_i-2}H_{di} \\
\sum_{2i} &= H_{Adi}, H_{Adi}H_{Adi}, \ldots, H_{Adi}^{d_i-2}H_{Adi} \\
\sum_{3i} &= I_n, H_{Adi}, \ldots, H_{Adi}^{d_i-2} \\
\sum_{li} &= \sum_{1i}, \sum_{2i}, \sum_{3i}
\end{align*}
\tag{53}
\]
Proving: See [46, Appendix A.5].

According to the results in Lemmas 1–6, the recursive form of $\Xi_i(t)$ in (33) will be given by Theorem 1.

**Theorem 1:** Define
\[
\begin{align*}
\Lambda_i \triangleq E[H_i(t) \odot H_i(t)] \\
V_i \triangleq E[H_i(t) \odot [I_n - H_i(t)]] \\
W_i \triangleq E[[I_n - H_i(t)] \odot [I_n - H_i(t)]]
\end{align*}
\tag{54}
\]
Then, the local estimation error covariance matrix $\Xi_i(t) \triangleq E[\tilde{\Xi}_i(t)\tilde{\Xi}_i(t)^T]$ for each CSE $\tilde{\Xi}_i(t)$ is given by
\[
\Xi_i(t) = A^d [W_i \odot (A \Xi_i(t - d_i - 1) A^T)] (A^d)^T + A^d [\Lambda_i \odot P_{ij}(t - d_i) + W_i \odot Q_w] (A^d)^T
+ A^d [V_i \odot (\Phi_{K_i}(t - d_i) \Gamma_{ij}(t - d_i - 1) A^T + G_{Ki}(t - d_i) Q_{wK_i}(t - d_i) F_{K_i}(t - d_i) + Q_w G_{K_i}(t - d_i))] (A^d)^T
+ \sum_{i=1}^{\theta_{dij}} A^{\theta_{dij}} Q_w [A^T (A^{\theta_{dij}})]^T
\tag{55}
\]
where $P_{ij}(t - d_i)$ and $\Gamma_{ij}(t - d_i - 1)$ are calculated by (7) and (45) (see Lemma 4). On the other hand, the estimation error cross-covariance matrix $\Xi_{ij}(t) \triangleq E[\tilde{\Xi}_i(t)\tilde{\Xi}_j(t)^T]$ is given by
\[
\Xi_{ij}(t) = H_{Adij}^\tau \Xi_{ij}(t - \tau_{ij}) \left( H_{Adij}^\tau \right)^T + \hat{\Xi}_{ij}(t)
+ H_{Adij} \left( \Phi_{\tilde{Y}_i}(t - d_i, t - d_i) H_{Adij}^\tau \right)^T + \Psi_{ij}(t) H_{Adij}^\tau + \Phi_{\tilde{F}_i}(t - d_i, d_j, t)
+ \Phi_{\tilde{F}_i}(t - d_i, d_j, t - d_i - 1) \bar{H}_{Adij}^\tau
+ H_{Adij} \left( \Phi_{\tilde{F}_i}(t - d_i, t - d_i - 1) \bar{H}_{Adij}^\tau \right)^T + \Psi_{ij}(t) H_{Adij}^\tau + \Phi_{\tilde{F}_i}(t - d_i, d_j, t)
+ H_{Adij} \left( \Phi_{\tilde{F}_i}(t - d_i, d_j, t - d_i - 1) \bar{H}_{Adij}^\tau \right)^T + \Psi_{ij}(t) H_{Adij}^\tau + \Phi_{\tilde{F}_i}(t - d_i, d_j, t)
+ \left( \Phi_{\tilde{F}_i}(t - d_i, d_j, t - d_i - 1) \bar{H}_{Adij}^\tau \right)^T + \left( \Phi_{\tilde{F}_i}(t - d_i, d_j, t - d_i - 1) \bar{H}_{Adij}^\tau \right)^T
+ \left( \Phi_{\tilde{F}_i}(t - d_i, d_j, t - d_i - 1) \bar{H}_{Adij}^\tau \right)^T + \left( \Phi_{\tilde{F}_i}(t - d_i, d_j, t - d_i - 1) \bar{H}_{Adij}^\tau \right)^T
\tag{56}
\]
where $\delta_{d_{ij}}$ is determined by (4), and $C_d(d_i, d_j)$ is determined by (34); $H_{Adij}$ and $\bar{H}_{Adij}$ are defined by (43), while $\tau_{ij}$, $\tau_{dij}$, and $\tau_{dij}$ are defined in (50). $\tilde{\theta}_{ij}(t)$ is calculated by (52) (see Lemma 6), and $\Psi_{ij}(t)$ is calculated by (49) (see Lemma 5), $\Theta_{\tilde{Y}_i}(t_1, t_2)$ and $\Theta_{\tilde{F}_i}(t_1, t_2)$ are calculated by (41) and (42) (see Lemma 3), while $\Phi_{\tilde{X}_i}(t_1, t_2)$, $\Phi_{\tilde{Y}_i}(t_1, t_2)$, $\Phi_{\tilde{F}_i}(t_1, t_2)$ are calculated by (36)–(38) (see Lemma 2). Moreover, the relationship
Algorithm 1 For the Given Selection Probabilities $\pi^*_j(h_i = 1, \ldots, \Delta_i; i = 1, \ldots, L)$ Satisfying (22)

1: At each sink node:
2:  for $i : = 1$ to $L$ do
3:  Calculate $G^*_i(t)$ and $K^*_i(t)$ by (6) and (7);
4:  Calculate the LOE $\hat{x}_i(t)$ by (5);
5:  Generate the binary variables $\sigma^*_i(h_i = 1, \ldots, \Delta_i)$ satisfying the categorical distribution, then determine the selected ASC $\hat{x}_s(t)$ by (13);
6:  end for
7: At the FC:
8:  for $i : = 1$ to $L$ do
9:  Calculate $G^*_i(t)$ by (6);
10: Calculate the CSE $\hat{x}_i(t)$ by (15);
11: for $j : = i$ to $L$ do
12: Calculate $P_{ij}(t)$ by (6–8);
13: Calculate $\Gamma_j(t), \Psi_j(t), \Upsilon_j(t)$ by (45), (49), (52);
14: Calculate $\Xi_{ij}(t)$ by (55–56);
15: end for
16: end for
17: Calculate $\Omega_1(t), \Omega_2(t), \ldots, \Omega_L(t)$ by (31);
18: Calculate the DKFE $\hat{x}(t)$ by (28).

between the CSE $\hat{x}_i^c(t)$ and the DKFE $\hat{x}(t)$ is

$$\text{Tr}[P(t)] \leq \text{Tr}[\Xi_{ii}(t)(i \in [1, 2, \ldots, L])].$$

(57)

Proof: See [46, Appendix A.6].

From Theorem 1, $\Xi(t)$ can be calculated by (55) and (56), then the optimal weighting matrices $\Omega_1(t), \ldots, \Omega_L(t)$ are obtained by (31). Moreover, the computation procedures for the DKFE $\hat{x}(t)$ can be summarized by Algorithm 1.

Remark 4: According to (55) and (56), each covariance matrix $\Xi_{ii}(t)$ is independent of the measurement $y_t(t)$ and the LOE $\hat{x}_i(t)$. Thus, $\Xi_{ij}(t)$ can be calculated at the FC when the initial values are given. In this case, only if each selected ASC $\hat{x}_s(t) \in R^n$ is sent to the FC, Algorithm 1 will be implemented in practical applications. On the other hand, when the communication delays and the number of local estimates increase slightly, the computational complexity of Algorithm 1 will be high. In this case, the steady-state DKFE with time-invariant weighting matrices can reduce the amount of computation. Therefore, to obtain the steady-state DKFE, we should find the stability conditions satisfying the following two points: 1) the covariance matrix of the recursive DKFE converges to a positive-definite matrix and 2) the limit of the covariance matrix $P(t)$ is independent of the initial values. Following this idea, the SDKFE is derived in the next section.

IV. STABILITY ANALYSIS FOR THE DKFE

A. Stability Condition of Each Local CSE

The estimation performance of each CSE $\hat{x}_i^c(t)$ will be discussed in this section. First, it is considered that the $i$th subsystem satisfies

$$A, \sqrt{Q_w}$$

is stabilizable and $(A, C_i)$ is detectable. (58)

When the condition (58) holds, it is well known that the estimation error covariance matrix $P_{ii}(t)$ in (7) will converge from any initial conditions $P_{ii}(0) > 0$ to the unique positive semidefinite solution $P_{ii}$. This means that

$$\lim_{t \to \infty} P_{ii}(t) = P_{ii}, \lim_{t \to \infty} [\Phi_{K_i}^*(t) = \Phi_{K_i}]$$

(59)

where the limits $P_{ii}$, $\Phi_{K_i}$, and $K_i$ are independent of the initial values. Moreover, $\Phi_{K_i}$ is a stable matrix. Thus, there must exist an integer $N_{P_i} > 0$ such that, for $t > N_{P_i}$, the estimation error system reduces to

$$\tilde{x}_i(t) = \Phi_{K_i}\tilde{x}_i(t - 1) + G_{K_i}w(t - 1) - K_iy_i(t).$$

(60)

Then, it follows from (60) that

$$\tilde{x}_i(t + 1) = \Phi_{K_i}^{d_i+1}\tilde{x}_i(t - d_i) + \zeta_i^o(t)$$

(61)

where

$$\frac{d_{i+1}}{di} + 1 \quad \begin{array}{c} \frac{d_{i+1}}{di} + 1 \\ \frac{d_{i+1}}{di} + 1 \end{array}$$

(62)

Meanwhile, it is derived from (25) and (60) that

$$\tilde{x}_i(t + 1) = A_{ii}(t)\tilde{x}_i^c(t - d_i) + A_{i2}(t)\tilde{x}_i(t - d_i) + \zeta_i^c(t)$$

(63)

where

$$\left[ \begin{array}{c} A_{ii}(t) = A_{ii}(t - d_i + 1)A \\ A_{i2}(t) = A_{i2}(t - d_i + 1)A \\ \zeta_i^c(t) = A_{ii}(t - d_i + 1)G_{K_i}w(t - d_i) \\ + A_{i2}(t - d_i + 1)w(t - d_i) + F_{(d_i + 1)}(d_i, 1, t + 1) \\ - A_{ii}(t - d_i + 1)w(t - d_i + 1) \end{array} \right].$$

(64)

Define $\tilde{E}_i(t) \triangleq \text{col}([\tilde{x}_i^c(t), \tilde{x}_i(t)])$, $\zeta_i(t) \triangleq \text{col}([\zeta_i^c(t), \zeta_i^o(t)])$. Then, combining (61) and (63) yields that

$$\tilde{E}_i(t + 1) = A_i(t)\tilde{E}_i(t - d_i) + \zeta_i(t)$$

(65)

where

$$A_i(t) = \begin{bmatrix} A_{ii}(t) & A_{i2}(t) \\ 0 & \Phi_{K_i}^{d_i+1} \end{bmatrix}.$$

(66)

It is concluded from the definition of $\tilde{E}_i(t)$ that if the covariance matrix $\tilde{E}_i(t) \triangleq E[\tilde{E}_i(t)[\tilde{E}_i(t)]^T]$ converges to the unique matrix $\tilde{E}_i$, there must exist the unique limit of the estimation error covariance matrix $\Xi_{ii}(t)$ for the $i$th CSE.

Lemma 7: Define

$$f(B) \triangleq E[A_i^T(t)BA_i(t)]$$

(67)

where $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$. Then, $f(B)$ is calculated by:

$$f(B) = \begin{bmatrix} \tilde{B}_{11} & \tilde{B}_{12} \\ \tilde{B}_{21} & \tilde{B}_{22} \end{bmatrix}$$

(68)
where

\[
\begin{align*}
\tilde{B}_{11} &= A^T \left[ W_i \otimes \left( \left( A^{d_i} \right)^T B_{11} A^{d_i} \right) \right] A \\
\tilde{B}_{12} &= A^T \left[ V_i \otimes \left( \left( A^{d_i} \right)^T B_{11} A^{d_i} \right) \right] \Phi_{K_i} \\
&+ A^T \left[ I_n - H_i \right] \left( A^{d_i} \right)^T \Phi_{K_i} \\
\tilde{B}_{21} &= \left( \Phi_{d_i} \right)^{T} \left[ B_{21} A^{d_i} \left( I_n - H_i \right) \right] A \\
&+ \left( \Phi_{d_i} \right)^{T} \left[ B_{21} A^{d_i} \left( I_n - H_i \right) \right] \Phi_{K_i} \\
\tilde{B}_{22} &= \left( \Phi_{d_i} \right)^{T} \left[ B_{22} \left( A^{d_i} \right)^T \right] \left( I_n - H_i \right) \left( A^{d_i} \right)^T \Phi_{K_i} \\
&+ \left( \Phi_{d_i} \right)^{T} \left[ B_{22} \left( A^{d_i} \right)^T \right] \Phi_{K_i} \\
\end{align*}
\]

(69)

where the probability selection matrix \( H_i \) is given by (23), and 
\( W_i, \Lambda_i, \) and \( V_i \) are given by (54). Moreover, for any matrices \( B_1 \) and \( B_2 \), there will be

\[
f(B_1 + B_2) = f(B_1) + f(B_2).
\]

(70)

**Proof:** Equation (68) can be obtained from (16) and (21). Lemma 1, and the definition of \( f(B) \), while (70) is derived from (68). This completes the proof.

Based on Lemma 7, the delay-dependent stability condition of the CSE \( \hat{x}(t) \) will be given in Theorem 2.

**Theorem 2:** For the communication delay \( d_i \) and selection probabilities \( \pi(h_i = 1, \ldots, \Delta_i) \) in (68), if there exist \( D_i > 0 \), \( X_i, Y_i, Z_i, \) and \( S_i > 0 \) such that

\[
\begin{align*}
\begin{bmatrix}
X_i & Y_i \\
Y_i^T & Z_i
\end{bmatrix} &\geq 0 \\
M_i &= \begin{bmatrix}
M_i(1, 1) & -Y_i - d_i Z_i A_i \\
-Y_i^T - d_i A_i^T Z_i & f(D_i) + d f(Z_i) - S_i
\end{bmatrix} < 0
\end{align*}
\]

(71)

(72)

where \( M_i(1, 1) = -d_i I + X_i + Y_i^T + Y_i + d_i Z_i + S_i \) and \( A_i = E(A_i(t)) \), while \( f(D_i) \) and \( f(Z_i) \) are calculated by (68). In Lemma 7, then the covariance matrix \( \Sigma_{u}(t) \) converges to the unique matrix, that is

\[
\lim_{t \to \infty} \Sigma_{u}(t) = \Sigma_{u}
\]

(73)

and the limit \( \Sigma_{u} \) is independent of the initial values.

**Proof:** See [46, Appendix A.7].

**Remark 5:** When \( d_i \neq 0 \), it is calculated from (54) and (55) that \( \text{Tr}[A^{d_i}[V_i \otimes (\Phi_{K_i}(t - d_i - 1)A^T + G_{K_i}(t - d_i - 1)Q_w)(A^{d_i})^T] \neq 0 \text{ and } \text{Tr}[A^{d_i}[V_i \otimes (\Phi_{K_i}(t - d_i - 1)A^T + G_{K_i}(t - d_i - 1)Q_w)(A^{d_i})^T] \neq 0 \), which are different from the results (101) and (102) in [32]. This implies that the stability condition for each CSE is difficult to be obtained by the derivation method based on the property of the operator \( \text{Tr}[] \) in [32]. In contrast, by adopting a novel derivation idea in this article, the stability conditions (71) and (72) are linear matrix inequalities (LMIs), and thus they can be verified by resorting to the MATLAB LMI Toolbox [44]. On the other hand, according to a similar derivation of stability conditions in [26], if the following conditions are satisfied:

\[
\begin{align*}
\zeta_{\infty,i} &\triangleq \| A^{d_i} \|_1 \| (I_n - H_i) \|_2 A \|_1 \\
&\times | A^{d_i} \|_1 \| (I_n - H_i) \|_2 A \|_1  < 1 \\
\zeta_{2,i} &\triangleq \| \Phi_{d_i} \|_2 A^{d_i} (I_n - H_i) A \|_2 < 1
\end{align*}
\]

(74)

where \( \Phi_{d_i} \triangleq \lim_{t \to \infty} \Phi_{K_i}(t) \), while \( \| \cdot \|_1 \) and \( \| \cdot \|_\infty \) represent the 1-norm and \( \infty \)-norm of matrices, respectively. Then, the covariance matrix \( \Sigma_{u}(t) \) will converge. However, the stability conditions in Theorem 2 are derived from the perspective of approaching the spectral radius when using the Laypunov stability theory, while the condition (74) is directly given by using the relaxation technique of matrix norms. Note that, for any matrix “\( o \),” there must be \( \rho(\cdot) \leq \| o \|_i \) for \( i = 1, 2, \ldots, \infty \), where \( \rho(\cdot) \) denotes the spectral radius. In this sense, the stability conditions in Theorem 2 have less conservative than (74) derived by [26]. Moreover, according to the definitions of 1-norm and \( \infty \)-norm, the condition (74) will become more conservative because of the product of \( \| \cdot \|_1 \) and \( \| \cdot \|_\infty \).

The above-mentioned result has been illustrated by Examples 1 and 2.

**Remark 6:** When \( d_i = 0 \), the Lyapunov function candidate can be chosen as \( V_{\xi}(t) = E(x_{\xi}(t)^T(t)D_{\xi}(t)) \). Then, it is concluded from the similar derivation of Theorem 2 that if there exists \( D_i > 0 \) such that

\[
\tilde{f}(D_i) - D_i < 0
\]

(75)

where \( \tilde{f}(D_i) \) is calculated by \( f(D_i) \) [i.e., (68) in Lemma 7] for \( d_i = 0 \), then the corresponding covariance matrix \( \Sigma_{u}(t) \) will converge to the unique steady-state value. Meanwhile, it is concluded from the result in [32] that if

\[
\lambda_{\max}(A^T(I_n - H_i)A) < 1
\]

(76)

where \( H_i \) is determined by (23), then the \( \lim_{t \to \infty} \text{Tr}[\Sigma_{u}(t)] \) will exist for \( d_i = 0 \). It should be pointed out that, under the condition (76), one cannot prove the following results: a) the limit of the covariance matrix \( \Sigma_{u}(t) \) exists and b) the limit of \( \Sigma_{u}(t) \) is independent of the initial conditions. Since the results (a) and (b) are necessary before deriving the steady-state DKFE, the steady-state fusion estimator cannot be given under the condition (76). Moreover, when only considering the convergence of the sequence \{\text{Tr} [\Sigma_{u}(t)]\}, the condition (75) has less conservatism than the condition (76). This is because the relaxation technique of matrix trace inequality is introduced to derive (76), but the condition (75) is derived by the stability theory without any relaxation. This result has been demonstrated by Example 1.

**B. Steady-State DKFE Design**

According to (32) and (33), the stability of the DKFE is also dependent on each estimation error cross-covariance matrix \( \Sigma_{u}(t) \) [see (56)]. Thus, the convergence of the sequence \{\text{Tr}[\Sigma_{u}(t)]\} will first be discussed in this section, and then combining Theorem 2 leads to the delay-dependent stability condition of the DKFE and the SDFKE. The main result will be presented in Theorem 3.

**Theorem 3:** Consider the CPSs (1) and (2) under the condition (58), if the selection probabilities \( \pi(h_i = 1, 2, \ldots, \Delta_i, i = 1, 2, \ldots, L) \) and the communication delays \( d_i(i = 1, 2, \ldots, L) \) satisfy (71), (72), and

\[
\rho(A^{d_i}(I_n - H_i)A) < 1(i = 1, 2, \ldots, L)
\]

(77)
Algorithm 2 For the Given Selection Probabilities $\pi^i_{h_i}(h_i = 1, \ldots, \Delta_i; i = 1, \ldots, L)$ Satisfying (22)

1: Determine the weighting matrices $\Omega_i(i = 1, \ldots, L)$ by (79);
2: At each sink node:
3: for $i : = 1$ to $L$ do
4: Calculate the LOE $\hat{x}_i(t)$ by (5);
5: Generate the binary variables $\sigma^i_{h_i}(h_i = 1, \ldots, \Delta_i)$ satisfying the categorical distribution, then determine the selected ASC $\hat{x}_s(t)$ by (13);
6: end for
7: At the FC:
8: for $i : = 1$ to $L$ do
9: Calculate the CSE $\hat{x}_i^s(t)$ by (15);
10: end for
11: Calculate the SDKFE $\hat{x}_s(t)$ by (80).

where $H_i$ is given by (23), then the fusion estimation error covariance matrix $P(t)$ (32) will converge to the unique matrix $P$, that is

$$\lim_{t \to \infty} P(t) = P$$  \hfill (78)

with $P$ independent of the initial values. Moreover, the steady-state weighing matrices $\Omega_i(i = 1, 2, \ldots, L)$ are calculated by

$$[\Omega_1, \Omega_2, \ldots, \Omega_L] = P^{-1}Q^{-1}$$  \hfill (79)

where $\lim_{t \to \infty} \Xi(t) = \Xi$, and the limit $\Xi$ is independent of the initial values. In this case, the SDKFE $\hat{x}_s(t)$ at the FC side is given by:

$$\hat{x}_s(t) = \sum_{i=1}^{L} \Omega_i \hat{x}_i^s(t)$$  \hfill (80)

where $\hat{x}_i^s(t)$ is calculated by (15).

Proof: See [46, Appendix A.8].

According to Theorem 3, the computation procedures for the SDKFE $\hat{x}_s(t)$ can be summarized by Algorithm 2.

Remark 7: It has been proved in Theorem 3 that when the conditions (71), (72), and (77) hold, the estimation error covariance matrix $P(t)$ can converge to the unique steady-state values for any initial values. Thus, the steady-state weighting matrices (79) can be obtained offline by implementing steps 7–16 of Algorithm 1. It is noted that the computational complexity of the SDKFE obtained by Algorithm 2 is much lower than that of the DKFE obtained by Algorithm 1.

Note that the stability condition in Theorems 2 and 3 is dependent on the communication delays and the selection probabilities of dimensionality reduction. Since each communication delay is determined by the property of communication channel, it is difficult to adjust the parameter $d_i$ to satisfy the stability condition. In this case, from the result (57) in Theorem 1, how to determine the selection probabilities (22) such that the MSE of the DKFE is bounded or convergent will be presented in Theorem 4.

Theorem 4: For the CPSs (1) and (2), when each communication delay $d_i$ is known in prior, two probability criteria to determine the dimensionality reduction strategy are presented as follows.

(C.1) To guarantee that the MSE of the DKFE is bounded, one needs to determine one group of the selection probability $\pi^i = \{\pi^i_1, \ldots, \pi^i_{\Delta_i}\}$ in (24) by (71) and (72).

(C.2) To guarantee the existence of the SDKFE, one needs to determine the $L$ selection probabilities $\pi^i(i = 1, \ldots, L)$ in (24) by (71), (72), and (77).

V. NUMERICAL EXAMPLES

In this section, two illustrative examples are presented to show the advantage and effectiveness of the proposed dimensionality reduction fusion estimation methods.

Example 1: Consider a CPS (1) with the following system parameters [19]:

$$A = \begin{bmatrix} 1.25 & 0 & 0 \\ 1 & 1.1 & 0 \\ \end{bmatrix}, Q_v = \begin{bmatrix} 20 & 0 \\ 0 & 20 \\ \end{bmatrix}$$  \hfill (81)

where the parameters of the first measurement equation in (2) are taken as

$$C_1 = [0 \ 1], Q_{v1} = 2.5.$$  \hfill (82)

Then, it is calculated from (81) and (82) that rank$\{\sqrt{Q_v} A \sqrt{Q_v}\} = 2$ and rank$\{\text{col}\{C_1, C_1 A\}\} = 2$, which means that (58) holds. In this case, one has by (6), (7), and (59) that $G_{K_1} = \begin{bmatrix} 0.4760 & -0.8573 \\ 0.0314 & 0.0376 \end{bmatrix}$. According to the dimensionality reduction strategy, it is considered in this example that only one component of $\hat{x}_i(t)$ is allowed to be transmitted to the FC at each time step. In this case, it is calculated from (54) that

$$\begin{cases}
A_1 = \text{diag}\{\gamma_{11}, 1 - \gamma_{11}\}, W_1 = \text{diag}\{1 - \gamma_{11}, \gamma_{11}\} \\
V_1 = \begin{bmatrix} 0 & \gamma_{11} \\ 1 - \gamma_{11} & 0 \end{bmatrix}, H_1 = \text{diag}\{\gamma_{11}, 1 - \gamma_{11}\}
\end{cases}$$  \hfill (83)

where $0 \leq \gamma_{11} \leq 1$.

To demonstrate the advantage of the designed stability condition, it is assumed that there is no communication delay for this example, and the selection probability $\gamma_{11}$ is taken as $\gamma_{11} = 0.5$. Then, by using the LMI Toolbox in MATLAB to solve the inequality matrix (75), $D_1$ is obtained, while it is calculated from (81) and (83) that

$$\begin{cases}
\lambda_{\max}(A_1^T (I_2 - H_1) A_1) = 1.5887 > 1 \\
\xi_{\infty, 1} = \| (I_2 - H_1) A_1^{\dagger} A_1 \|_{\infty} = 2.3625 > 1.
\end{cases}$$

Therefore, it is concluded from Theorem 2 that the limit of $\text{Tr}(\Xi(t))$ exists, however, the condition (74) derived by [26] and the condition (76) derived by [32] do not hold for this example. Moreover, when choosing $\gamma_{11}$ from 0 to 1, the effectiveness of the conditions (75) and (76) is shown in Table I and Figs. 3 and 4. It can be seen from Figs. 3 and 4 that the sequence of $\text{Tr}(\Xi(t))$ is convergent when $\gamma_{11} \in [0.4, 0.5, 0.6, 0.7, 0.8]$. This result can be directly obtained by the judgement condition (75), however, the judgement condition (74) derived by [26] and the judgement condition (76) derived by [32] are invalid for the above cases. Therefore, the
Table I

| \( \gamma_{11} \) | (75) | (76) derived by [32] | (74) derived by [26] |
|-----------------|------|---------------------|---------------------|
| 0               | False| False               | False               |
| 0.2             | False| False               | False               |
| 0.4             | True | False               | False               |
| 0.6             | True | False               | False               |
| 0.8             | True | False               | False               |
| 0.9             | False| False               | False               |
| 1.0             | False| False               | False               |

where \( \lambda_{\text{max}}(A) = 1.0441 > 1 \) means that this 4-bus smart grid system is unstable, and the covariance of the process noise is taken as

\[
Q_w = \begin{bmatrix}
0.04 & 0.1 & 0.06 & 0.08 \\
0.1 & 0.25 & 0.15 & 0.2 \\
0.06 & 0.15 & 0.09 & 0.12 \\
0.08 & 0.2 & 0.12 & 0.16 \\
\end{bmatrix}.
\]

To monitor the work status of the grid, two sink nodes collect their sensor measurements, and the local estimates computed by the sink nodes are transmitted to the FC (e.g., monitoring center or control center). Then, the measurement matrices in (2) are given by

\[
C_1 = \begin{bmatrix}
1 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 \\
1 & 0 & 1 & 0 \\
\end{bmatrix},
C_2 = \begin{bmatrix}
1 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 \\
\end{bmatrix}
\]

which means that the measurement information on the fourth component of \( x(t) \) cannot be obtained by the first sink node, while the measurement information on the second component of \( x(t) \) cannot be obtained by the second sink node. The covariances of \( \gamma_i(t) (i = 1, 2) \) in (2) are taken as \( Q_{\gamma_1} = \text{diag}(0.9, 0.6, 0.9, 0.4) \) and \( Q_{\gamma_2} = \text{diag}(0.3, 0.4, 0.5, 0.2) \), respectively. Then, it is calculated from (84)–(86) that \( \text{rank}(\text{col}(C_1, C_1A^2, C_1A^3)) = 4 \) \( (i = 1, 2) \) and \( \text{rank}(\sqrt{Q_{\gamma_1}}A, \sqrt{Q_{\gamma_2}}A, \sqrt{Q_{\gamma_1}}A^2, \sqrt{Q_{\gamma_2}}A^2) = 4 \), which means that condition (58) holds. Thus, the limits in (59) exist, and their values are omitted due to page limitation.

For this example, according to the dimensionality reduction strategy, it is considered that only two components of \( \hat{x}(t) \) are allowed to be transmitted to the FC for satisfying the finite bandwidth and, thus, \( r_1 = r_2 = 2 \) and \( \Delta_1 = \Delta_2 = 6 \). In this case, the diagonal matrices \( H_{hi}^1 (i = 1, 2; h_i = 1, 2, 3, 4, 5, 6) \) in (16) are given by

\[
\begin{align*}
H_{h_1}^1 &= \text{diag}(1, 1, 0, 0), \quad H_{h_2}^1 = \text{diag}(0, 0, 1, 0) \\
H_{h_1}^2 &= \text{diag}(1, 0, 0, 1), \quad H_{h_2}^2 = \text{diag}(0, 1, 1, 0) \\
H_{h_1}^3 &= \text{diag}(0, 1, 0, 1), \quad H_{h_2}^3 = \text{diag}(0, 0, 1, 1).
\end{align*}
\]

Then, it follows from (16) and (87) that \( H_i(t) = \text{diag}(\sigma_1^1(t) + \sigma_1^2(t), \sigma_1^1(t) + \sigma_1^2(t), \sigma_1^1(t) + \sigma_1^2(t), \sigma_1^1(t) + \sigma_1^2(t), \sigma_1^1(t) + \sigma_1^2(t)) \), where \( \sigma_i^h(t)(h_i = 1, 2, 3, 4, 5, 6) \) are determined by (11) and (12), and each stochastic process \( \{\sigma_i^h(t)\} \) obeys the categorical distribution. To determine the signal \( \hat{x}_i(t) \) [see (13)], the selection probabilities in (22) are taken as \( \pi_1^1 = 0.3, \pi_2^1 = 0.2, \pi_3^1 = 0.1, \pi_4^1 = 0.1, \pi_5^1 = 0.1, \pi_6^1 = 0.2, \pi_1^2 = 0.2, \pi_2^2 = 0.1, \pi_3^2 = 0.2, \pi_4^2 = 0.1, \pi_5^2 = 0.3, \pi_6^2 = 0.1 \). Thus, the selection probability matrices \( H_1^1 \) and \( H_2^1 \) [see (23)] are given by

\[
\begin{align*}
H_1^1 &= \text{diag}(0.6, 0.5, 0.5, 0.4) \\
H_2^1 &= \text{diag}(0.5, 0.6, 0.3, 0.6).
\end{align*}
\]

When each selected ASC \( \hat{x}_i(t) \) is transmitted to the FC, the communication delays are taken as \( d_1 = 1 \) and \( d_2 = 2 \). In this case, it is calculated from (84) and (88) that \( \rho(A(h_4 - H_1)A) = 0.5759 < 1 \) and \( \rho(A^2(h_4 - H_2)A) = 0.6631 < 1 \), which mean that the condition (77) holds in Theorem 3. Meanwhile, by using the LMI Toolbox in MATLAB, the variables \( D_i, X_i, \)
Y_i, Z_i, and S_i (i = 1, 2) are obtained by solving the matrix inequalities (71) and (72), that is, the conditions (71)–(72) hold for the two local CSEs with different selection probabilities and communication delays. Under this case, it is concluded from Theorem 3 that the fusion estimation covariance matrix P(t) for this example converges to a unique matrix, and the SDKFE exists. Then, implementing Algorithm 1 obtains the steady-state weighting matrices as follows:

\[
\begin{align*}
\Omega_1 &= \begin{bmatrix}
0.6254 & 0.0921 & 0.3294 & 0.044 \\
0.0585 & 0.7874 & 0.2587 & 0.2107 \\
0.1654 & 0.0271 & 0.6857 & 0.2670 \\
0.0065 & 0.0765 & 0.2257 & 0.6729 \\
0.3746 & -0.0921 & -0.3294 & -0.044 \\
-0.0585 & 0.2126 & -0.2587 & -0.2107 \\
-0.1654 & -0.0271 & 0.3143 & -0.2670 \\
-0.0065 & -0.0765 & -0.2257 & 0.3271 \\
\end{bmatrix}, \\
\Omega_2 &= \begin{bmatrix}
\end{align*}
\]

(89)

Thus, the SDKFE \( \hat{x}_i(t) \) for this example is obtained by substituting (89) into (80). Note that one also has by (84) and (88) that

\[
\begin{align*}
\zeta_{\infty, 1} &= ||A||_1 ||(I_n - H_1)^\frac{1}{2} A||_1 \\
&\times ||A||_\infty ||(I_n - H_1)^\frac{1}{2} A||_\infty = 0.8657 < 1 \\
\xi_{\infty, 2} &= ||A^2||_1 ||(I_n - H_1)^\frac{1}{2} A||_1 \\
&\times ||A^2||_\infty ||(I_n - H_1)^\frac{1}{2} A||_\infty = 1.3554 > 1 \\
\end{align*}
\]

(90)

It is obvious that the judgement condition (74) derived by [26] is invalid for the second sink node in this example, which implies that the stability condition in this article has less conservatism than the one in [26].

Let \( P_0(t) \) denote the original DKFE (ODKFE) under the dimensionality reduction when there are no communication delays between the sink nodes and the FC. Then, the estimation performances (assessed by the trace of the estimation error covariance matrix) of the local CSEs, DKFE, and ODKFE are shown in Fig. 5. It is seen from this figure that the estimation performance of the DKFE is better than that of each CSE at each time-step, which is in line with the result (57). However, the estimation performance of the DKFE is worse than that of the ODKFE, which implies that the communication delays can affect the fusion estimation performance. Moreover, it is known from the this figure that the MSEs of the DKFE and CSEs all converge to the steady-state values, which accords with the results (73) and (78).

To demonstrate the effectiveness of the SDKFE, the matrix 2-norms of \( P(t) \) and \( \Omega_i(t) (i = 1, 2) \) are shown in Fig. 6 under different initial values. It is seen from this figure that \( ||P(t)||_2 \) and \( ||\Omega_i(t)||_2 \) (i = 1, 2) can converge to the unique steady-state values under different initial values, which is in line with the result in Theorem 3. On the other hand, let \( E_r(t) = \hat{x}_i(t) - \hat{x}_i(t, 1) \), where \( \hat{x}_i(t) \) represents the ith component of the DKFE \( \hat{x}_i(t) \), and the meaning of \( \hat{x}_i(t, 1) \) is the same as that of \( \hat{x}_i(t, 1) \). Then, implementing Algorithms 1 and 2, the trajectories of \( E_r(t) (i = 1, 2, 3, 4) \) are depicted in Fig. 7, where the measurement sequences \( \{y_i(t) (i = 1, 2) \} \) are the same when computing the DKFE \( \hat{x}_i(t) \) and the SDKFE \( \hat{x}_i(t) \). It is shown from this figure that the errors between the DKFE and SDKFE will converge to 0 as \( t \) increases, which accords with the property of the SDKFE. It should be pointed out that the SDKFE is much easier to implement as compared with the DKFE in practical applications.

VI. CONCLUSION

To guarantee the satisfactory estimation performance in CPSs, the distributed dimensionality reduction fusion estimation problem with communication delays has been studied in this article. Based on the stochastic dimensionality reduction strategy, a mathematical model was proposed to establish the relationship between the dimensionality reduction and communication delays, and then the recursive DKFE was obtained.
by resorting to the optimal weighted fusion criterion. A delay-dependent and probability-dependent condition, which can be easily judged by using the MATLAB LMI Toolbox, was derived for the DKFE such that the fusion estimation error covariance matrix $P(t)$ converges to a unique steady-state matrix. Moreover, when the communication delay $d_i$ is known in advance, the selection probability criterion to determine the dimensionality reduction strategy has also been presented. Finally, two examples were given to demonstrate the advantage and effectiveness of the proposed methods.

REFERENCES

[1] P.-Y. Chen, S.-M. Cheng, and K.-C. Chen, “Information fusion to defend intentional attack in Internet of Things,” IEEE Internet Things J., vol. 1, no. 1, pp. 337–348, Jun. 2014.

[2] H. M. La and W. Sheng, “Distributed sensor fusion for scalar field mapping using mobile sensor networks,” IEEE Trans. Cybern., vol. 43, no. 2, pp. 766–778, Apr. 2013.

[3] Y. Li, D. K. Jha, A. Ray, and T. A. Wettergren, “Information fusion of passive sensors for detection of moving targets in dynamic environments,” IEEE Trans. Cybern., vol. 47, no. 1, pp. 93–104, Jan. 2017.

[4] A. Rosaslovad and S. Veeraraghavan, “Sensor data fusion algorithms for vehicular cyber-physical systems,” IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 9, pp. 1762–1774, Sep. 2012.

[5] X. R. Li, Y. Zhu, J. Wang, and C. Han, “Optimal linear estimation fusion I: Unified fusion rules,” IEEE Trans. Inf. Theory, vol. 49, no. 9, pp. 2192–2208, Sep. 2003.

[6] J. A. Roecker and C. D. McGillen, “Comparison of two-sensor tracking methods based on state vector fusion and measurement fusion,” IEEE Trans. Aerosp. Electron. Syst., vol. 24, no. 4, pp. 447–449, Jul. 1988.

[7] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, Estimation with Applications to Tracking and Navigation. New York, NY, USA: Wiley, 2001.

[8] N. A. Carlson, “Federated square root filter for decentralized parallel processors,” IEEE Trans. Aerosp. Electron. Syst., vol. 26, no. 3, pp. 517–525, May 1990.

[9] S. Sun and Z. Deng, “Multi-sensor optimal information fusion Kalman filter,” Automatica, vol. 40, pp. 1017–1023, Jun. 2004.

[10] Z.-L. Deng, Y. Gao, L. Mao, Y. Li, and G. Hao, “New approach to information fusion to defend intentional attack in Internet of Things,” IEEE Internet Things J., vol. 1, no. 1, pp. 337–348, Aug. 2014.

[11] H. Li, L. Lai, and H. V. Poor, “Multicast routing for decentralized control of cyber physical systems with an application in smart grid,” IEEE J. Sel. Areas Commun., vol. 30, no. 6, pp. 1097–1107, Jul. 2012.

[12] Y.-F. Huang, S. Werner, J. Huang, N. Kashyp, and V. Gupta, “State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid,” IEEE Signal Process. Mag., vol. 29, no. 5, pp. 33–43, Sep. 2012.

[13] X. Cao, P. Cheng, J. Chen, S. S. Ge, Y. Cheng, and Y. Sun, “Cognitive radio based state estimation in cyber-physical systems,” IEEE J. Sel. Areas Commun., vol. 32, no. 3, pp. 489–502, Mar. 2014.

[14] S. Deshmukh, B. Natarajan, and A. Pahwa, “State estimation over a lossy network in spatially distributed cyber-physical systems,” IEEE Trans. Signal Process., vol. 62, no. 15, pp. 3911–3923, Aug. 2014.

[15] Z. Zhang, C. Xia, S. Chen, T. Yang, and Z. Chen, “Reachability analysis of networked finite state machine with communication losses: A switched perspective,” IEEE J. Sel. Areas Commun., vol. 38, no. 5, pp. 845–853, May 2020.

[16] B. Chen, G. Hu, W.-A. Zhang, and L. Yu, “Information fusion estimation for spatially distributed cyber-physical systems with communication delay and bandwidth constraints,” in Proc. Amer. Control Conf., Chicago, IL, USA, 2015, pp. 5152–5157.

[17] B. Chen, D. W. C. Ho, G. Hu, and L. Yu, “Secure fusion estimation for bandwidth constrained cyber-physical systems under replay attacks,” IEEE Trans. Cybern., vol. 48, no. 6, pp. 1862–1876, Jun. 2018.

[18] Z. Wang and Y. Niu, “Distributed estimation and filtering for sensor networks,” Int. J. Syst. Sci. Spec. Issue, vol. 42, pp. 1421–1425, Jul. 2011.

[19] B. Chen, W. Zhang, G. Hu, and L. Yu, “Networked fusion Kalman filtering with multiple uncertainties,” IEEE Trans. Aerosp. Electron. Syst., vol. 51, no. 3, pp. 2332–2349, Jul. 2015.

[20] B. Chen, G. Hu, D. W. C. Ho, and L. Yu, “Distributed covariance intersection fusion estimation for cyber-physical systems with communication constraints,” IEEE Trans. Autom. Control, vol. 61, no. 12, pp. 4020–4026, Dec. 2016.

[21] S. Sun, H. Lin, J. Ma, and X. Li, “Multi-sensor distributed fusion estimation with applications in networked systems: A review paper,” Inf. Fusion, vol. 17, pp. 122–134, Nov. 2012.

[22] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. B. Giannakis, “Distributed compression-estimation using wireless sensor networks,” IEEE Signal Process. Mag., vol. 23, no. 4, pp. 27–41, Jul. 2006.

[23] J. Fang and H. Li, “Hyperplane-based vector quantization for distributed fusion estimation in wireless sensor networks,” IEEE Trans. Inf. Theory, vol. 55, no. 12, pp. 5682–5699, Dec. 2009.

[24] X. Shen, P. K. Varshney, and Y. Zhu, “Robust distributed maximum likelihood estimation with dependent quantized data,” Automatica, vol. 50, no. 1, pp. 169–174, 2014.

[25] D. Li, S. Kar, F. E. Alsaadi, A. M. Dobaie, and S. Cui, “Distributed Kalman filtering with quantized sensing state,” IEEE Trans. Signal Process., vol. 63, no. 19, pp. 5180–5193, Oct. 2015.

[26] B. Chen, W.-A. Zhang, and L. Yu, “Distributed finite-horizon fusion Kalman filtering for bandwidth and energy constrained wireless sensor networks,” IEEE Trans. Signal Process., vol. 62, no. 4, pp. 797–812, Feb. 2014.

[27] H. Ma, Y.-H. Yang, Y. Chen, K. J. R. Liu, and Q. Wang, “Distributed state estimation with dimension reduction preprocessing,” IEEE Trans. Signal Process., vol. 62, no. 12, pp. 3098–3110, Jun. 2014.

[28] B. Chen, W.-A. Zhang, X. Yu, G. Hu, and H. Song, “Distributed fusion estimation with communication bandwidth constraints,” IEEE Trans. Autom. Control, vol. 60, no. 5, pp. 1398–1403, May 2015.

[29] J. Ma and S. Sun, “Centralized fusion estimators for multisensor systems with random sensor delays, multiple packet dropouts and uncertain observations,” IEEE Sensors J., vol. 15, no. 9, pp. 4043–4053, Sep. 2015.

[30] A. Chiuso and L. Schenato, “Information fusion strategies and performance bounds in packet-drop networks,” Automatica, vol. 47, no. 7, pp. 1304–1316, 2011.

[31] Y. Xia, J. Shang, J. Chen, and G.-P. Liu, “Networked data fusion with packet losses and variable delays,” IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 39, no. 5, pp. 1389–1403, Oct. 2009.

[32] R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, “Fusion estimation using measured outputs with random parameter matrices subject to random delays and packet dropouts,” Signal Process., vol. 127, pp. 12–23, Oct. 2016.

[33] Z. Xing and Y. Xia, “Distributed federated Kalman filter fusion over multi-sensor unreliable networked systems,” IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 63, no. 10, pp. 1714–1725, Oct. 2016.
Bo Chen (Member, IEEE) received the B.S. degree in information and computing science from the Jiangxi University of Science and Technology, Ganzhou, China, in 2008, and the Ph.D. degree in control theory and control engineering from the Zhejiang University of Technology, Hangzhou, China, in 2014. He joined the Department of Automation, Zhejiang University of Technology in 2018, where he is currently a Professor. He was a Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, from 2014 to 2015 and from 2017 to 2018. He was also a Postdoctoral Research Fellow with the Department of Mathematics, City University of Hong Kong, Hong Kong, from 2015 to 2017. His current research interests include information fusion, distributed estimation and control, cyber–physical systems security, and networked fusion systems. Prof. Chen was a recipient of the Outstanding Thesis Award of Chinese Association of Automation in 2015.

Daniel W. C. Ho (Fellow, IEEE) received the B.S., M.S., and Ph.D. degrees in mathematics from the University of Salford, Greater Manchester, U.K., in 1980, 1982, and 1986, respectively. From 1985 to 1988, he was a Research Fellow with the Industrial Control Unit, University of Strathclyde, Glasgow, U.K. In 1989, he joined the City University of Hong Kong, Hong Kong, where he is currently a Chair Professor of Applied Mathematics and an Associate Dean of the College of Science. He has over 250 publications in scientific journals. His current research interests include control and estimation theory, complex dynamical distributed networks, multiagent networks, and stochastic systems. Prof. Ho was a recipient of the Chang Jiang Chair Professor Awarded by the Ministry of Education, China, in 2012, and the ISI Highly Cited Researchers Award in Engineering by Clarivate Analytics, from 2014 to 2020. He has been on the Editorial Board of a number of journals, including the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IET CONTROL THEORY AND APPLICATIONS, The Journal of the Franklin Institute, and the Asian Journal of Control.

Guoqiang Hu (Senior Member, IEEE) received the B.Eng. degree in automation from the University of Science and Technology of China, Hefei, China, in 2002, the M.Phil. degree in automation and computer-aided engineering from the Chinese University of Hong Kong, Hong Kong, in 2004, and the Ph.D. degree in mechanical engineering from the University of Florida, Gainesville, FL, USA, in 2007. He joined the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 2011, where he is currently a Full Professor and the Director of the Centre for System Intelligence and Efficiency. He was an Assistant Professor with Kansas State University, Manhattan, KS, USA, from 2008 to 2011. His research interests include distributed control, distributed optimization and game theory, with applications to multirobot systems and smart city systems.

Prof. Hu was a recipient of the Best Paper in Automation Award in the 14th IEEE International Conference on Information and Automation and the Best Paper Award (Guan Zhao-Zhi Award) in the 36th Chinese Control Conference. He serves as an Associate Editor for IEEE TRANSACTIONS ON AUTOMATIC CONTROL and IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY.

Li Yu (Member, IEEE) received the B.S. degree in control theory from Nankai University, Tianjin, China, in 1982, and the M.S. and Ph.D. degrees from Zhejiang University, Hangzhou, China, in 1988 and 1999, respectively. He is currently a Professor with the College of Information Engineering, Zhejiang University of Technology, Hangzhou, China. He has successively presided over 20 research projects. He has published five academic monographs, one textbook, and over 300 journal papers. He has also been authorized over 100 patents for invention and granted five scientific and technological awards. His current research interests include robust control, networked control systems, cyber–physical systems security, and information fusion.