Control-Theoretical and Topological Analysis of Covariance Intersection Based Distributed Kalman Filter

Tsang-Kai Chang and Ankur Mehta

***Abstract*—Covariance intersection (CI) extends Kalman filter (KF) in distributed estimation, since it can fuse Gaussian estimates in the absence of the estimates' correlations. However, even with the preliminary success on the integration of CI and KF, existing discussion limited in global behavior is unable to directly deal with a system with a mixture of unbounded-covariance and bounded-covariance agents. In other words, until this letter there has been no explicit investigation on the analytic relationship between effective observability in each agent and system topology, to the best of our knowledge. To formalize these problems, we establish CI-based KF with explicit CI topology, on top of the conventional KF with observation exchanges. Consequently, the effect of CI on KF can be characterized by the impact of individual CI links on each agents. In particular, we systematically show that CI links can diminish the effective unobservable space, which relaxes the boundedness criterion. In addition, as a conservative fusion scheme, there may exist CI links that provide no improvement on estimation performance but generate additional uncertainty. A method is proposed to identify and then to suppress such redundant CI links for enhanced estimation performance. Finally, the pros and cons of CI on distributed estimation algorithms are comprehensively characterized and substantiated by a numerical example.**

***Index Terms*—Kalman filtering, estimation, distributed control.**

I. Introduction

DUE TO its scalability and robustness, distributed estimation is preferred to its centralized counterpart in distributed networked systems, in particular multirobot systems [1] and sensor networks [2], [3]. The success of the Kalman filter (KF) in centralized systems, thanks to its elegant algorithmic description and explicit covariance characterization, makes it a popular basis for distributed estimation algorithm. The main challenge of extending to distributed KF lies in the unavailability of cross-correlation among distributed agents. More specifically, the fusion of state estimate among agents is relatively difficult without the knowledge of the correlation between those estimates. Lacking correlations, naive fusion may lead to the over-confidence problem, which undermines the reliability of such estimation algorithm.

There are several attempts to extend KF to distributed estimation. In consensus KF [4]–[6], consensus filtering is conducted among agents to replicate a centralized KF. However, the convergence to the consensus requires infinite communication steps between two consecutive local updates, which is far from realizable in any systems. Proposed in [7], Diffusion Kalman filter (D-KF) takes convex combination of estimates in diffusion step such that only local information is required within the algorithm, but the lack of corresponding covariance update makes the covariance terms stored in the algorithm no longer meaningful. Consensus+Innovations KF is proposed in [8] and [9], where consensus KF is improved to allow sensing and communication occurring at the same time scale. Nevertheless, the global covariance is kept and updated in local nodes, which is only possible under time-invariant assumption of system model and communication topology.

Covariance Intersection (CI) is a viable fusion method in the absence of correlations, and the covariance obtained by CI is guaranteed to maintain estimation consistency [10]–[13]. While one cannot retrieve the exact fused covariance with correlations ignored, the covariance provided by CI is no smaller than the real covariance in positive definite sense. In other words, a conservative estimate is obtained from CI regarding the estimation accuracy.

The application of CI on KF, termed CI-KF, is proposed in [14] and [15], respectively. The work in [14] generalizes the diffusion step in D-KF to CI, but the introduction of information consensus to the case without local observability is unnecessary. In fact, the boundedness criterion of CI-KF is milder than local observability, as pointed out in [2]. In [15], a general Kullback-Leibler average of probability distributions is proposed with CI as a special case, and is applied in distributed estimation. The following work [2] further proceeds to one communication step per iteration, and analyzes the covariance boundedness on the whole system. However, since the covariance boundedness is a local property, a system with global unbounded covariance may have

Manuscript received March 6, 2018; revised May 10, 2018; accepted May 30, 2018. Date of publication June 7, 2018; date of current version June 25, 2018. Recommended by Senior Editor G. Yin. *(Corresponding author: Tsang-Kai Chang.)*

The authors are with the Electrical and Computer Engineering Department, University of California at Los Angeles, Los Angeles, CA 90095 USA (e-mail: tsangkaichang@ucla.edu; mehtank@ucla.edu).

Digital Object Identifier 10.1109/LCSYS.2018.2844734

2475-1456 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

agents with bounded covariance. In this case, a global boundedness analysis can not tell whose covariance is bounded, and how to achieve boundedness covariance for the rest agents. On the contrary, a boundedness analysis on each agent gives exact insight on estimation performance, both locally and globally. In [3], the boundedness analysis is established on global observability and strong connectivity, but such criteria are redundant in general. For example, when every agent is locally observable, the strong connectivity is no longer required for the covariance boundedness. This case suggests that the boundedness of CI-KF depends on observability and network topology jointly, which is completely characterized by super-local observability in this letter. In addition, the drawback of redundant CI link is never addressed, to the best of our knowledge.

To comprehensively formalize CI-KF, we first describe the information exchange in the distributed system by observation topology and CI topology. While both topologies resides on communication links, they are distinguished by the information conveyed in such link. We show that by introducing CI steps, the effective observability of the system is altered, and the boundedness criterion of the covariance matrix is relaxed to super-local observable. While CI seems advantageous in distributed estimation, we identify the case where CI step worsens the estimation accuracy, due to the additional uncertainty introduced by CI as a conservative fusion scheme.

The main contributions of this letter include:

- the systematic characterization of CI-KF with explicit separation of observation and CI topologies,
- the covariance boundedness criterion on single agent based on the effective observability of its super-neighbors, and
- the identification of redundant CI links which introduces additional uncertainty.

This letter is organized as follows: We setup the background material in Section II, and the main algorithm CI-KF is summarized in the following section. The boundedness analysis and the redundant CI link are discussed in Sections IV and V, respectively. A numerical example is presented in Section VI and conclusions are delivered in Section VII.

II. Background

A. System Description

We consider a discrete-time linear time-invariant system with time index t as

$$x_{t+1} = Fx_t + Gw_t, \tag{1}$$

where $x_t \in \mathbb{R}^n$ is the state that the distributed agents try to estimate, and $w_t \in \mathbb{R}^m$ is the process noise. We assume that F is nonsingular in this letter. Those agents are labeled by $\Omega = \{1, \ldots, N\}$, and each agent observes the system with the model

$$y_{i,t} = H_i x_t + v_{i,t}, \tag{2}$$

for $i \in \Omega$. In (2), $y_{i,t} \in \mathbb{R}^{p_i}$ is the measurement result and $v_{i,t}$ is the measurement noise. The process and the measurement noises are assumed to be zero-mean, uncorrelated, and white Gaussian with $\mathsf{E}[w_t w_t^\mathsf{T}] = Q \geq 0$, and $\mathsf{E}[v_{i,t} v_{i,t}^\mathsf{T}] = R_i > 0$, respectively. The observability matrix of pair (F, H) is defined as

$$\mathbb{O}(F, H) = \begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-1} \end{bmatrix}, \tag{3}$$

and the unobservable space is defined as the nullspace of the matrix $\mathbb{O}(F, H)$, denoted by $\mathcal{N}(\mathbb{O}(F, H))$.

B. Estimation Consistency and Covariance Intersection

A consistent estimate can be regarded as a conservative estimate regarding the estimation uncertainty. Formally, a consistent estimate is mean-preserving and has no smaller covariance matrix in positive definite sense, with the definition as follows.

Definition 1 (Estimation Consistency): An estimate $\hat{z}$ of z is a Gaussian random vector with mean $\mathsf{E}[\hat{z}]$ and covariance $\Sigma_{\hat{z}}$. The estimation $\hat{z}'$ of z is called *consistent* of $\hat{z}$ if $\mathsf{E}[\hat{z}'] = \mathsf{E}[\hat{z}]$ and $\Sigma_{\hat{z}'} \geq \Sigma_{\hat{z}}$.

Lemma 1 (Covariance Intersection [10]–[13]): Given N consistent estimates $\hat{z}_i$ of $\hat{z}$ with covariances $\Sigma_{\hat{z}_i}$ for $i = 1, \ldots, N$, the estimate $\hat{z}'$ is also consistent of $\hat{z}$ with $\Sigma_{\hat{z}'}^{-1} = \sum_{i=1}^{N} c_i \Sigma_{\hat{z}_i}^{-1}$, and $\mathsf{E}[\hat{z}'] = \Sigma_{\hat{z}'} \sum_{i=1}^{N} c_i \Sigma_{\hat{z}_i}^{-1} \mathsf{E}[\hat{z}_i]$, where the nonnegative coefficients c_i satisfy $\sum_{i=1}^{N} c_i = 1$.

CI is able to combine several consistent estimates which are not necessarily uncorrelated to get a consistent result. The nonnegative coefficients $\{c_i, i = 1, \ldots, N\}$ such that $\sum_{i=1}^{N} c_i = 1$ are called *convex coefficients* in the following.

C. Network Topology

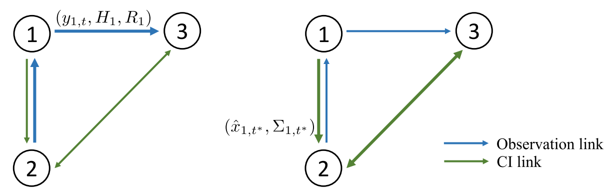
A directed graph $\mathcal{G} = (V, E_\mathcal{G})$ is applied to characterize the information flow via communication. In the graph $\mathcal{G}$, the vertex set V contains all the agents, i.e., $V = \Omega$, and an edge $(j, i) \in E_\mathcal{G}$, $j \neq i$, represents that agent j sends its information to agent i. We may refer an edge as a *link* in this letter. A *path* in $\mathcal{G}$ is given by a sequence of vertices $(v_{i_1}, v_{i_2}, \ldots, v_{i_{m+1}})$ such that $(v_{i_k}, v_{i_{k+1}}) \in E_\mathcal{G}$ for $k = 1, \ldots, m$. A *cycle* in graph $\mathcal{G}$ is defined as a path $(v_{i_1}, v_{i_2}, \ldots, v_{i_{m+1}})$ with $v_{i_1} = v_{i_{m+1}}$. The *neighborhood* of agent i is defined as $N_\mathcal{G}(i) = \{j | (j, i) \in E_\mathcal{G}\}$. That is, the neighborhood of agent i contains all agents that send its information to agent i. The *inclusive neighborhood* of agent i is defined by $N_\mathcal{G}^*(i) = N_\mathcal{G}(i) \cup \{i\}$. Complete treatment of graph theory on multiagent systems can be found in [16].

III. Covariance-Intersection Based Kalman Filter Algorithm

In CI-KF, the common state x_t is estimated by each agent with estimator $\hat{x}_{i,t}$ for $i \in \Omega$ by three steps:

1) observation update,
2) CI update,
3) time update.

In the time interval from t to $t+1$, we use time index t^* for the estimation after observation update, and t^+ for that after CI update. Of these three steps, the first two involve information transmission between agents. Even though the communication



(a) Observation update (b) CI update

Fig. 1. The information exchange in CI-KF described by observation and CI topologies.

channels underlying these two steps may follow similar characteristics, we define observation topology $\mathcal{O}$ and CI topology $\mathcal{C}$ separately, in order to distinguish the types of exchanged information, as in Fig. 1.

In observation update, each agent sends its raw measured result to the recipients according to observation topology. For agent i, the received observation information contains $(y_{j,t}, H_j, R_j)$ for $j \in N_{\mathcal{O}}(i)$. With $N_{\mathcal{O}}(i) = \{i_1, i_2, \dots, i_{n_i}\}$, we can concatenate the raw observation information as

$$\bar{y}_{i,t} = \begin{bmatrix} y_{i,t} \\ y_{i_1,t} \\ \vdots \\ y_{i_{n_i},t} \end{bmatrix} = [y_{j,t}]_{j \in N^*_{\mathcal{O}}(i)}. \tag{4}$$

Equivalently, $\bar{y}_i$ can be obtained from the observation model

$$\bar{y}_{i,t} = \bar{H}_i x_t + \bar{v}_{i,t}, \tag{5}$$

with

$$\bar{H}_i = [H_j]_{j \in N^*_{\mathcal{O}}(i)}. \tag{6}$$

Since all the observation information exchanges occur in single-hop fashion, the noises of the received information are thus independent. That is, the noise $\bar{v}_{i,t}$ in (5) is zero-mean Gaussian with covariance matrix $\bar{R}_i = \mathrm{Diag}(R_i, R_{i_1}, \dots, R_{i_{n_i}})$. Consequently, all agents can directly fuse the received measurement information as in standard KF procedure by

$$\hat{x}_{i,t^*} = \hat{x}_{i,t} + \Sigma_{i,t} \bar{H}_i^{\mathsf{T}} \bar{R}_i^{-1} (\bar{y}_{i,t} - \bar{H}_i \hat{x}_{i,t}), \tag{7}$$

and

$$\Sigma^{-1}_{i,t^*} = \Sigma^{-1}_{i,t} + \bar{H}_i^{\mathsf{T}} \bar{R}_i^{-1} \bar{H}_i. \tag{8}$$

In CI update, agents send their state estimation together with the state covariance based on CI topology $\mathcal{C}$. Once an agent receives all the incoming information, it can fuse all the information by CI without the correlations among all incoming estimates, while maintains estimation consistency. For agent i, the incoming information at CI step is $(\hat{x}_{j,t^*}, \Sigma_{j,t^*})$ for $j \in N_{\mathcal{C}}(i)$. By applying CI formula in Lemma 1, the result of CI update is given by

$$\hat{x}_{i,t^+} = \Sigma_{i,t^+} \sum_{j \in N^*_{\mathcal{C}}(i)} c_j \Sigma^{-1}_{j,t^*} \hat{x}_{j,t^*}, \tag{9}$$

and

$$\Sigma^{-1}_{i,t^+} = \sum_{j \in N^*_{\mathcal{C}}(i)} c_j \Sigma^{-1}_{j,t^*} \tag{10}$$

with convex coefficients $\{c_j, j \in N^*_{\mathcal{C}}(i)\}$. The choice of CI coefficients affects the accuracy of the estimation, and there is plenty of research devoted to the selection of CI coefficients, [2], [3] for instance. However, in this letter, instead of seeking the optimal coefficients in particular cases, we assume constant CI coefficients to focus on the intrinsic uncertainty introduced by CI step, which will be further discussed in Section V.

Because each agent knows the system propagation model, time update can be easily performed distributively without communication. With the zero-mean assumption on noises, time update gives

$$\hat{x}_{i,t+1} = F \hat{x}_{i,t^+}, \tag{11}$$

and

$$\Sigma_{i,t+1} = F \Sigma_{i,t^+} F^{\mathsf{T}} + GQG^{\mathsf{T}}. \tag{12}$$

In summary, while all information exchange relies only on one-hop communication in CI-KF, no sophisticated routing mechanism is required. Additionally, the communication link in both observation and CI topologies is not required bidirectional, which makes CI-KF appealing for distributed system in real implementation.

IV. Boundedness Analysis

Since covariance matrix represents the uncertainty in KF-based estimation algorithms, the boundedness of covariance matrix is essential for estimation algorithm applicability. However, directly investigating the boundedness of $\Sigma_{i,t}$ is tedious and less insightful. Therefore, we instead construct an auxiliary upper bound $\Pi_{i,t}$ such that $\Pi_{i,t} \geq \Sigma_{i,t}$ if $\Pi_{i,0} = \Sigma_{i,0}$. Moreover, $\Pi_{i,t}$ is chosen to follow Riccati recursion, and thus the boundedness of $\Sigma_{i,t}$ can be determined by the convergence of $\Pi_{i,t}$.

By observing that both observation and CI updates are elegant in information form, we attempt to find the upper-bound recursion of $\Pi_{i,t}$ such that these two updates can be combined into a compact information form as

$$\Pi^{-1}_{i,t^+} = c_i \Pi^{-1}_{i,t} + \check{H}_i^{\mathsf{T}} \check{R}_i^{-1} \check{H}_i, \tag{13}$$

with $c_i > 0$. Meanwhile, the upper-bound of time update simply follows

$$\Pi_{i,t+1} = F \Pi_{i,t^+} F^{\mathsf{T}} + GQG^{\mathsf{T}}, \tag{14}$$

as the time update of $\Sigma_{i,t}$ in (12). With explicit information form (13) and covariance form (14) combined together, the recursion of $\Pi_{i,t}$ now simply follows Riccati recursion whose convergence is well-studied in [17]. To clarify the notation, $\bar{H}_i$, defined in (6), is used for the observation matrix when agent i incorporates all the incoming observation information, and $\check{H}_i$, defined in (13), is used in the upper bound $\Pi_{i,t}$ and incorporates both observation and CI updates.

In the following, we demonstrate the way to construct $\Pi_{i,t}$, especially the constants in (13), and investigate the boundedness of $\Sigma_{i,t}$ in consequence.

A. The Construction of $\Pi_{i,t}$

We construct $\Pi_{i,t}$ in three scenarios for completeness. First, the construction of $\Pi_{i,t}$ for those agents without neighbors in CI topology, or $N_{\mathcal{C}}(i) = \varnothing$, is straightforward. If an agent has no neighbors in CI topology, it simply skips CI step, and the rest are only observation and time updates. Explicitly, if $N_{\mathcal{C}}(i) = \varnothing$, by picking $c_i = 1$, $\check{H}_i = \bar{H}_i$ and $\check{R}_i = \bar{R}_i$, then $\Pi_{i,t} \geq \Sigma_{i,t}$ for all t with $\Pi_{i,0} = \Sigma_{i,0}$.

If all the CI neighbors of i have well-defined covariance upper bounds, then $\Pi_{i,t}$ can be constructed recursively. That is, for agent i with CI neighbors $N_{\mathcal{C}}(i)$, $\Pi_{j,t}$ exists for $j \in N_{\mathcal{C}}(i)$. Based on the covariance updates (8), (10) and (12),

$$\begin{aligned}\Sigma_{i,t^+}^{-1} &= c_i\Sigma_{i,t^*}^{-1} + \sum_{j\in N_{\mathcal{C}}(i)} c_j\Big(\Sigma_{j,t}^{-1} + \bar{H}_j^{\mathsf{T}}\bar{R}_j^{-1}\bar{H}_j\Big)\\ &\geq c_i\Sigma_{i,t}^{-1} + c_i\bar{H}_i^{\mathsf{T}}\bar{R}_i^{-1}\bar{H}_i\\ &\quad + \sum_{j\in N_{\mathcal{C}}(i)} c_j\Big(F\Sigma_{j,(t-1)^+}F^{\mathsf{T}} + GQG^{\mathsf{T}}\Big)^{-1},\end{aligned}$$

by omitting $\bar{H}_j^{\mathsf{T}}\bar{R}_j^{-1}\bar{H}_j$. We need the following lemma to convert the order of matrix inversion and multiplication.

Lemma 2 [15]: For a nonsingular matrix $\Omega \geq 0$, there exists a real number $0 < \beta \leq 1$ such that

$$\Big(F\Omega^{-1}F^{\mathsf{T}} + Q\Big)^{-1} \geq \beta F^{-\mathsf{T}}\Omega F^{-1}.$$

With the assumption that F is nonsingular and Lemma 2,

$$\begin{aligned}\Big(F\Sigma_{j,(t-1)^+}F^{\mathsf{T}} + GQG^{\mathsf{T}}\Big)^{-1} &\geq \beta_j F^{-\mathsf{T}}\Sigma_{j,(t-1)^+}^{-1}F^{-1}\\ &\geq \beta_j F^{-\mathsf{T}}\Big(c_j'\Pi_{j,t-1}^{-1} + \check{H}_j^{\mathsf{T}}\check{R}_j^{-1}\check{H}_j\Big)F^{-1}\\ &\geq \beta_j F^{-\mathsf{T}}\check{H}_j^{\mathsf{T}}\check{R}_j^{-1}\check{H}_jF^{-1},\end{aligned}$$

where the second inequality depends on (13) and the third is obtained by omitting $c_j'\Pi_{j,t-1}^{-1}$. Thus, Σ_{i,t^+}^{-1} follows

$$\begin{aligned}\Sigma_{i,t^+}^{-1} &\geq c_i\Sigma_{i,t}^{-1} + c_i\bar{H}_i^{\mathsf{T}}\bar{R}_i^{-1}\bar{H}_i\\ &\quad + \sum_{j\in N_{\mathcal{C}}(i)} c_j\beta_jF^{-\mathsf{T}}\check{H}_j^{\mathsf{T}}\check{R}_j^{-1}\check{H}_jF^{-1}.\end{aligned}$$

Consequently, $\Pi_{i,t}$ in (13) can be constructed with

$$\check{H}_i = \begin{bmatrix}\bar{H}_i\\ [\check{H}_jF^{-1}]_{j\in N_{\mathcal{C}}(i)}\end{bmatrix}, \tag{15}$$

and

$$\check{R}_i^{-1} = \begin{bmatrix}c_i\bar{R}_i^{-1} & 0\\ 0 & \mathsf{Diag}(c_j\beta_j\check{R}_j^{-1})_{j\in N_{\mathcal{C}}(i)}\end{bmatrix}.$$

In (15), the information coming from CI links introduces augmented rows in the observation matrix $\check{H}_i$, which makes the effective unobservable space shrink in agent i. The multiplication of F^{-1} is due to the fact that $\check{H}_j$ comes from the previous time slot, but the multiplication has no effect on the associated unobservable space according to the following lemma.

Lemma 3: For nonsingular F and integer k, $\mathcal{N}(\mathbb{O}(F,H)) = \mathcal{N}(\mathbb{O}(F,HF^k))$.

Proof: To begin with, we let the characteristic polynomial of F as $p_F(\lambda) = \sum_{l=0}^{n}\alpha_l\lambda^l$. Consider $x \in \mathcal{N}(\mathbb{O}(F,H))$, or $HF^lx = 0$, $l = 0,\ldots,n-1$. By Cayley-Hamilton theorem, $\sum_{l=0}^{n}\alpha_lF^l = 0$. With F^n expressed as the linear combination of $I, F, \ldots, F^{n-1}$, we have $HF^nx = 0$, or $x \in \mathcal{N}(\mathbb{O}(F,HF))$, which leads to the fact that $\mathcal{N}(\mathbb{O}(F,H)) \subseteq \mathcal{N}(\mathbb{O}(F,HF))$. The inverse statement can be proven similarly, which jointly concludes that $\mathcal{N}(\mathbb{O}(F,H)) = \mathcal{N}(\mathbb{O}(F,HF))$.

Iteratively, $\mathcal{N}(\mathbb{O}(F,H)) = \mathcal{N}(\mathbb{O}(F,HF^k))$ can be shown for positive integer k. The statement for negative integer k can be similarly established by first showing that $HF^{-1}x = 0$, and the rest follows. ■

For those agents forming a cycle in CI topology $\mathcal{C}$, the aforementioned procedure cannot be used, since the construction so far relies on its neighborhood. To mitigate the problem, we first observe that the information agent i sends out in CI update to the cycle L will be received by agent i again as $\bar{H}_iF^{-k}$, where k is the number of steps to traverse the cycle L. Based on Lemma 3, $\bar{H}_iF^{-k}$ has the exactly the same effect as $\bar{H}_i$ regarding system observability. However, $\bar{H}_i$ is the information already obtained by agent i by itself. Thus, removing the outgoing CI link from agent i to the next agent in the cycle does not affect the effective observability at agent i. Therefore, by removing the CI links starting from i , or $(i,l) \in E_{\mathcal{C}}$ for $l \in \Omega$, the cycle in CI topology no longer exists, and $\Pi_{i,t}$ as well as $\check{H}_i$ can then be constructed with the aforementioned two scenarios.

B. Observability of $(F,\check{H}_i)$

In the following, we characterize the effect of observation links and that of CI links by examining the effective unobservable space at each agent, which is mainly determined by the rows in $\check{H}_i$.

Proposition 1 (Observation Link): If $j \in N_{\mathcal{O}}(i)$, then

$$\mathcal{N}\Big(\mathbb{O}(F,\check{H}_i)\Big) \subseteq \mathcal{N}\Big(\mathbb{O}(F,\bar{H}_i)\Big) \subseteq \mathcal{N}\Big(\mathbb{O}(F,H_j)\Big). \tag{16}$$

Proof: The result is direct from the definition of $\bar{H}_i$ in (6) and that of $\check{H}_i$ in (15). ■

Proposition 2 (CI Link): If $j \in N_{\mathcal{C}}(i)$, then

$$\mathcal{N}\Big(\mathbb{O}(F,\check{H}_i)\Big) \subseteq \mathcal{N}\Big(\mathbb{O}(F,\check{H}_j)\Big). \tag{17}$$

Proof: Proposition 2 is a direct result of the definition in (15) and Lemma 3. ■

Propositions 1 and 2 show that agent i inherits the information from its neighbor by adding additional rows in its effective observation matrix $\check{H}_i$ from both CI and observation links. However, there is a fundamental difference between two link types. If agent j is a neighbor in observation topology $\mathcal{O}$ of agent i, then only H_j is appended in $\check{H}_i$. On the contrary, if agent j is in the CI neighborhood of agent i, the whole $\check{H}_j$ can be affixed in $\check{H}_i$. The difference comes from the fact that the information shared in observation update is only the local observation result, while in CI update, the whole state is used, which contains the information fused up to this agent. Therefore, the next critical question is which information will show up in a single local agent i, in the form of the appended rows in $\check{H}_i$. We collect those agents into the definition of *super neighborhood* $S(i)$ of agent i.

Definition 2 (Super Neighborhood): For $j \neq i$, $j \in S(i)$ if

1) $(j,i) \in E_{\mathcal{O}}$, or

2) there is a path in $\mathcal{C}$ from j to i, or
3) there is an agent k such that $(j,k) \in E_{\mathcal{O}}$ and there is a path in $\mathcal{C}$ from k to i.

Similarly, the inclusive super neighborhood is defined as $S^*(i) = S(i) \cup \{i\}$.

C. Convergence of $\Pi_{i,t}$

By combining (13) and (14), the recursion of $\Pi_{i,t}$ can be expressed as

$$\Pi_{i,t+1} = \check{F}_i\left(\Pi_{i,t}^{-1} + \check{H}_i^{\mathsf{T}}(c_i\check{R}_i)^{-1}\check{H}_i\right)^{-1}\check{F}_i^{\mathsf{T}} + GQG^{\mathsf{T}} \quad (18)$$

with $\check{F}_i = \frac{1}{\sqrt{c_i}}F$. The equation (18) is exactly a Riccati recursion, whose convergence is stated as follows.

Lemma 4 [17]: Given $(\check{F}_i, GQ^{1/2})$ controllable and $(\check{F}_i, \check{H}_i)$ observable and $\Pi_{i,0} \geq 0$, the recursion (18) converges to the unique positive semi-definite solution Π of the discrete algebraic Riccati equation (DARE)

$$\Pi = \check{F}_i\Pi\check{F}_i^{\mathsf{T}} + GQG^{\mathsf{T}} - \Pi\check{H}_i^{\mathsf{T}}\left(c_i\check{R}_i + \check{H}_i^{\mathsf{T}}\Pi\check{H}_i\right)^{-1}\check{H}_i\Pi.$$

To proceed the convergence discussion, we assume that $(F, GQ^{1/2})$ is controllable, and the convergence criterion of interest lies primarily in the pair $(\check{F}_i, \check{H}_i)$.

Definition 3 (Super-Local Observable): An agent i is called super-local observable if $(F, [H_j]_{j\in S^*(i)})$ is observable.

Theorem 1 (Main Convergence Theorem): If F is nonsingular, $(F, GQ^{1/2})$ is controllable, and agent i is super-local observable, then $\Sigma_{i,t}$ is bounded.

Proof: Since agent i is super-local observable, $(\check{F}_i, \check{H}_i)$ is also observable, and then $\Pi_{i,t}$ converges by Lemma 4. Regarding $\Pi_{i,t} \geq \Sigma_{i,t}$, $\Sigma_{i,t}$ is bounded. ■

Theorem 1 indicates the interplay between observability in control theory and connectivity in topologies on the covariance boundedness. In other words, for each agent, the more observable it is, the less dependent it is on the incoming information for a bounded estimation covariance. One can also extend Theorem 1 without difficulty to account for arbitrary set of agents, or even the whole system, on which the analysis in [2] and [3] focuses.

Since the diffusion step of D-KF is a degenerative case of CI without covariance update, the contribution of CI step in CI-KF can be obviously noted by comparing it to D-KF. The convergence criterion for D-KF is locally observable, as investigated in [7]. In fact, if agent i is locally observable, the estimation covariance can be bounded by only executing observation and time updates, which is identical to running D-KF without diffusion update. That is, both convergence criteria with and without diffusion step in D-KF are the same, which makes the effect of diffusion step dubious. For CI-KF, CI step does contribute to the boundedness of the estimation covariance as indicated by Proposition 2. Consequently, super-local observability is milder than local observability, which shows the superiority of CI-KF over D-KF.

V. Removing Redundant CI Links

Even though CI update can carry the information in state estimate and pass it down, the constant multiplication in CI step introduces additional uncertainty to compensate for those unknown correlations. As a result, while some CI links introduce no extra information, CI fusion may impair estimation performance from the constant multiplication. Thus, the redundant CI links should be identified and removed for better estimation result.

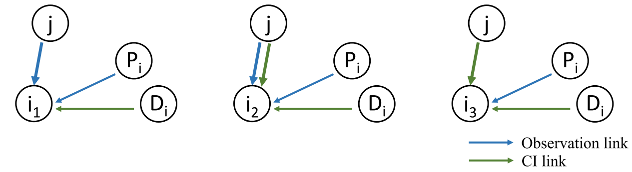


Fig. 2. Observation and CI topologies of agents i and j in Proposition 3. Agents i_2 and i_3 use the identical CI coefficients $\{c_k, k \in D_i^* \cup \{j\}\}$, where they both use c_i on their own estimate in CI step, and i_1 use $\{\frac{c_k}{1-c_j}, k \in D_i^*\}$ for CI coefficients.

Theorem 2 (Redundant CI Link): Given m convex coefficients $\{c_1, \ldots, c_m\}$ and m covariance matrices Σ_i, $i = 1, \ldots, m$, if $(\sum_{i=1}^{m-1}\frac{c_i}{1-c_m}\Sigma_i^{-1})^{-1} \leq \Sigma_m$, then

$$\left(\sum_{i=1}^{m-1}\frac{c_i}{1-c_m}\Sigma_i^{-1}\right)^{-1} \leq \left(\sum_{i=1}^{m}c_i\Sigma_i^{-1}\right)^{-1}. \quad (19)$$

Theorem 2 can be proved with direct algebraic manipulation. By noting that $\{\frac{c_i}{1-c_m}, i = 1, \ldots, m-1\}$ are also convex coefficients, the inequality (19) suggests that the fusion result with only $m-1$ of all m estimates is smaller than the fusion result of all m estimates in positive definite sense, if the criterion is satisfied. Theorem 2 suggests that removing the redundant CI link decreases estimation uncertainty.

One of the common configurations resulting in redundant CI link is the agents without neighbors, as in Fig. 2. Suppose that agent j has no incoming information, or $N_{\mathcal{O}}(j) = N_{\mathcal{C}}(j) = \varnothing$, and agent j sends its information to agent i. Regarding the configuration of observation and CI links, we consider three cases i_1, i_2, and i_3 where

$$\begin{aligned} N_{\mathcal{O}}(i_1) &= P_i \cup \{j\}, & N_{\mathcal{C}}(i_1) &= D_i,\\ N_{\mathcal{O}}(i_2) &= P_i \cup \{j\}, & N_{\mathcal{C}}(i_2) &= D_i \cup \{j\},\\ N_{\mathcal{O}}(i_3) &= P_i, & N_{\mathcal{C}}(i_3) &= D_i \cup \{j\}. \end{aligned}$$

Proposition 3: With the same initial condition, $\Pi_{i_1,0} = \Pi_{i_2,0} = \Pi_{i_3,0}$, $\Pi_{i_2,t} \leq \Pi_{i_3,t}$. Furthermore, if $c_i \geq 1 - c_j$, $\Pi_{i_1,t} \leq \Pi_{i_2,t}$ for all t.

Proposition 3 shows that when agent j has no incoming information, the CI link from j can be removed for better estimation performance, if the information from j is already provided in observation step and is weighted enough in CI step. The intuitive explanation is that CI update provides identical information to the one from observation update, but with additional uncertainty from constant scaling.

The discussion on removing redundant CI links explains the separation of observation and CI topologies at first place. In other words, even though both links rely on the same communication mechanism for information transmission, they have different impacts on the estimation performance. To extend Theorem 2, a more general version of Theorem 2, which is

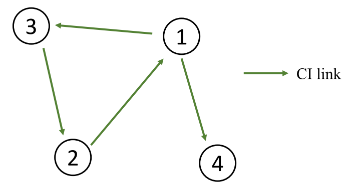


Fig. 3. The topology of four agents in the numerical example.

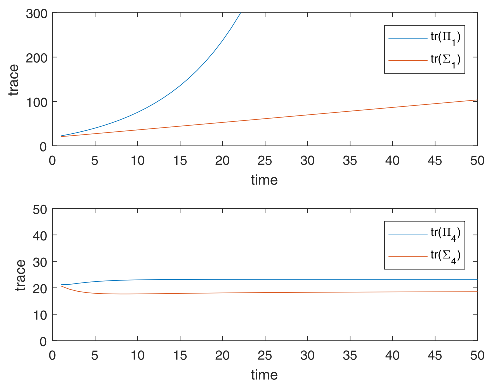


Fig. 4. The traces of covariance and the corresponding upper-bound for agent 1 and 4.

independent of the choice of CI coefficients, is necessary to further identify redundant CI link.

VI. A Numerical Example

In this section, we apply the technique developed in Section IV to find an upper bound for covariance and determine its boundedness. We take $F = I_3$, where I_n is a $n \times n$ identity matrix, and $GQG^{\mathsf{T}} = 1.69I_3$. Four distributed agents try to estimate the common state, with observation matrices

$$H_1 = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix}, \quad H_2 = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix},$$
$$H_3 = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}, \quad H_4 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix},$$

respectively. We let $R_i = 0.1$ for all i. Since observation links can be directly merged as in (5), we focus on CI links with the overall topology in Fig. 3. For agent 4, its inclusive super neighborhood $S^*(4) = \{1, 2, 3, 4\}$, and it is thus super-local observable. By Theorem 1, $\Sigma_{4,t}$ is bounded. But we can tell from the numerical result that $\Sigma_{1,t}$ is unbounded. Previous analysis that focuses on the whole system is not able to handle this mixture scheme directly. For agent 1 in the cycle of CI topology, we can construct the covariance upper bound $\Pi_{1,t}$ according the method in Section IV, and the inequality $\Sigma_{1,t} \leq \Pi_{1,t}$ holds as in the Fig. 4. Even though $\Sigma_{1,t}$ is divergent in Fig. 4, Theorem 1 suggests that either observation link or CI link from agent 4 to agent 1 can make $\Sigma_{1,t}$ bounded, which again shows the benefit of local analysis.

VII. Conclusion

As an essential role to facilitate distributed KF, CI-KF is clearly investigated with its positive and possibly negative effects in this letter. By integrating CI into the observability analysis, we have shown that CI decreases the effective unobservable space of an individual node, and that this decrease depends on the network topologies. Such analysis not only characterizes the covariance boundedness by the effective observability of *super-neighborhood*, it also indicates the additional required information to achieve bounded covariance, as presented in Section VI. We further discussed that the scalar multiplication in the inverse of the covariance matrix introduces additional uncertainty during CI, and such uncertainty may be redundant in certain cases.

CI-KF is particularly suitable for mobile or dynamic agents, e.g., multirobot systems, since every step in the algorithm is robust to time-varying conditions. Our analysis is also general enough for heterogeneous systems composed of agents with various requirements on estimation performance. Further investigation can extend this analysis to establish a tradeoff between the estimation performance and the corresponding resource cost (as in [18]), characterizing additional benefits of CI-KF for systems with power constraints.

References

[1] T.-K. Chang, S. Chen, and A. Mehta, "Multirobot cooperative localization algorithm with explicit communication and its topology analysis," in *Proc. Int. Symp. Robot. Res.*, to be published.

[2] S. Wang and W. Ren, "On the convergence conditions of distributed dynamic state estimation using sensor networks: A unified framework," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 4, pp. 1300–1316, Jul. 2018.

[3] X. He, W. Xue, and H. Fang, "Consistent distributed state estimation with global observability over sensor network," *Automatica*, vol. 92, pp. 162–172, Jun. 2018.

[4] R. Olfati-Saber, "Distributed Kalman filter with embedded consensus filters," in *Proc. 44th IEEE Conf. Decis. Control*, Dec. 2005, pp. 8179–8184.

[5] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proc. 46th IEEE Conf. Decis. Control*, Dec. 2007, pp. 5492–5498.

[6] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, "Distributed Kalman filtering based on consensus strategies," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 4, pp. 622–633, May 2008.

[7] F. S. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Trans. Autom. Control*, vol. 55, no. 9, pp. 2069–2084, Sep. 2010.

[8] S. Das and J. M. F. Moura, "Distributed Kalman filtering with dynamic observations consensus," *IEEE Trans. Signal Process.*, vol. 63, no. 17, pp. 4458–4473, Sep. 2015.

[9] S. Das and J. M. F. Moura, "Consensus+Innovations distributed Kalman filter with optimized gains," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 467–481, Jan. 2017.

[10] S. J. Julier and J. K. Uhlmann, "A non-divergent estimation algorithm in the presence of unknown correlations," in *Proc. Amer. Control Conf.*, vol. 4, Jun. 1997, pp. 2369–2373.

[11] L. Chen, P. O. Arambel, and R. K. Mehra, "Estimation under unknown correlation: Covariance intersection revisited," *IEEE Trans. Signal Process.*, vol. 47, no. 11, pp. 1879–1882, Nov. 2002.

[12] S.-L. Sun, "Multi-sensor optimal information fusion Kalman filters with applications," *Aerosp. Sci. Technol.*, vol. 8, no. 1, pp. 57–62, Jan. 2004.

[13] M. Reinhardt, B. Noack, P. O. Arambel, and U. D. Hanebeck, "Minimum covariance bounds for the fusion under unknown correlations," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1210–1214, Sep. 2015.

[14] J. Hu, L. Xie, and C. Zhang, "Diffusion Kalman filtering based on covariance intersection," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 891–902, Feb. 2012.

[15] G. Battistelli and L. Chisci, "Kullback–Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability," *Automatica*, vol. 50, no. 3, pp. 707–718, Mar. 2014.

[16] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*. Princeton, NJ, USA: Princeton Univ. Press, 2010.

[17] T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2000.

[18] T.-K. Chang and A. Mehta, "Optimal scheduling for resource-constrained multirobot cooperative localization," *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 1552–1559, Jul. 2018.