Split Covariance Intersection with Correlated Components for Distributed Estimation

Colin Cros[1,2], Pierre-Olivier Amblard[1], Christophe Prieur[1], and Jean-François Da Rocha[2]
[1]*Univ. Grenoble Alpes, CNRS, GIPSA-Lab*, Grenoble, France
[2]*Telespazio France*, Toulouse, France
colin.cros@isae.fr, pierre-olivier.amblard@cnrs.fr,
christophe.prieur@gipsa-lab.fr jeanfrancois.darocha@telespazio.com

***Abstract*—This paper introduces a new conservative fusion method to exploit the correlated components within the estimation errors. Fusion is the process of combining multiple estimates of a given state to produce a new estimate with a smaller MSE. To perform the optimal linear fusion, the (centralized) covariance associated with the errors of all estimates is required. If it is partially unknown, the optimal fusion cannot be computed. Instead, a solution is to perform a conservative fusion. A conservative fusion provides a gain and a bound on the resulting MSE matrix which guarantees that the error is not underestimated. A well-known conservative fusion is the Covariance Intersection fusion. It has been modified to exploit the uncorrelated components within the errors. In this paper, it is further extended to exploit the correlated components as well. The resulting fusion rule is integrated into standard distributed algorithms where it allows exploiting the process noise observed by all agents. The improvement is confirmed by simulations.**

***Index Terms*—Conservative fusion, Covariance Intersection, Distributed estimation, Linear fusion**

I. Introduction

Distributed estimation is a recurrent problem in sensor networks. It consists in estimating the state of a dynamical system using a network of agents, i.e., nodes equipped with sensors and with communication capabilities. Each agent is performing independent measurements of the state, and shares its estimate with its neighbors in the network. The fusion of the estimates received by an agent is a complex task, especially when the communication between the agents are limited. To optimally fuse several estimates, i.e., with the smallest resulting Mean Square Error (MSE) matrix, the agent needs to know the covariances of the errors of each estimate and their cross-covariances. For example, the optimal fusion of two estimates is given by the Bar-Shalom-Campo's formulas [3]. However, in cooperating networks, each agent has only local knowledge; it can estimate the covariance of its own error but cannot estimate the cross-covariances with its neighbors' errors, as this would require knowledge of the whole network topology. Without these cross-covariances, the agent cannot calculate the MSE matrix of the fused estimate and must use *conservative* bounds. A conservative fusion method provides a bound on the MSE matrix which guarantees that the estimation error is not underestimated. The first conservative fusion method proposed is known as Covariance Intersection (CI) [8]. CI provides conservative bounds by considering that the estimation errors may be correlated to any degree. Generally, this assumption is very loose and tighter fusion methods have been derived when refined assumptions can be made. If the errors contain independent components, an extension of CI called Split CI (SCI) provides tighter bounds [9]. Furthermore, if the vectors of errors are partitioned into two components with only the cross-covariances between the first components of the vectors unknown, another fusion rule called Partitioned CI (PCI) was proposed in [16] and improved in [1] and also provide better fusion bounds. A third extension is known as Inverse CI (ICI). Initially, it was developed to deal with estimates obtained by combining some common estimate with independent estimates [2], [15]. Under some conditions, ICI can also handle unknown correlated components [14]. All these improvements of CI use the structure of the errors to reduce the set of admissible cross-covariances and tighten the bounds [5]. They have been applied to a wide range of problems: e.g., SLAM [10], cooperative localization [11], or cooperative perception [12].

This work was partially supported by DGA/AID project 2022 65 0082.

In distributed estimation, the fusion of the estimates can be performed using CI. However, two elements can be used to produce tighter bounds. First, the independent measurements induce independent components in the errors. The SCI fusion rule was designed to exploit such components [9]. An alternative is to transmit the measurements to the neighbors, so they can optimally fuse them with their own estimate, this method is known as Diffusion Kalman Filtering (DKF) with CI [4], [7]. The second element to exploit is the process noise. This noise is observed by all the agents and generates a common error to all estimators. As presented in this paper, this common term can also be exploited to produce tighter covariance bounds. The current fusion methods do not take advantage from such common noises as they do not consider correlated components.

This paper introduces an extension of the SCI fusion method, called Extended SCI (ESCI). It is motivated by the distributed estimation problems in which it exploits both the uncorrelated components of the errors (induced by the measurements) and their correlated components (induced by the process noise). It is integrated into standard algorithms and applied to an example inspired by Search-And-Rescue (SAR) missions.

The rest of the paper is organized as follows. Section II recalls the definitions of a conservative fusion and of the SCI fusion method. Then, Section III presents the new ESCI fusion method. Section IV details its integration into standard distributed estimation algorithms. The application is presented in Section V. Finally, Section VI concludes the paper.

Notation. In the sequel, vectors are denoted in lowercase boldface letters e.g., $\boldsymbol{x} \in \mathbb{R}^n$, and matrices in uppercase boldface variables e.g., $\boldsymbol{M} \in \mathbb{R}^{n\times n}$. The notation $\mathrm{E}[\cdot]$ denotes the expected value. For two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, the notation $\boldsymbol{A} \preceq \boldsymbol{B}$ means that the difference $\boldsymbol{B}-\boldsymbol{A}$ is positive semi-definite. The unit simplex of $\mathbb{R}^n$ is denoted as $\mathcal{K}^n \triangleq \{\boldsymbol{x} \in \mathbb{R}^n \mid \forall i,\, x_i \geq 0,\, \boldsymbol{x}^\intercal \boldsymbol{1} = 1\}$.

II. Background

Consider N unbiased estimates $\hat{\boldsymbol{x}}_i$ for $i \in \{1,\dots,N\}$ of a random variable $\boldsymbol{x} \in \mathbb{R}^d$. The estimation errors are denoted as $\tilde{\boldsymbol{x}}_i \triangleq \hat{\boldsymbol{x}}_i - \boldsymbol{x}$ and their covariances as $\tilde{\boldsymbol{P}}_i \triangleq \mathrm{E}\left[\tilde{\boldsymbol{x}}_i\tilde{\boldsymbol{x}}_i^\intercal\right]$. A linear fusion is defined by a matrix of gains $\boldsymbol{K} = \begin{bmatrix}\boldsymbol{K}_1 & \dots & \boldsymbol{K}_N\end{bmatrix} \in \mathbb{R}^{d\times Nd}$, with $\boldsymbol{K}_i \in \mathbb{R}^{d\times d}$ and $\sum_i \boldsymbol{K}_i = \boldsymbol{I}_d$, as:

$$\hat{\boldsymbol{x}}_F(\boldsymbol{K}) \triangleq \sum_{i=1}^N \boldsymbol{K}_i\hat{\boldsymbol{x}}_i = \boldsymbol{K}\hat{\boldsymbol{x}}_\mathrm{c}, \tag{1}$$

where $\hat{\boldsymbol{x}}_\mathrm{c} \triangleq \begin{pmatrix}\hat{\boldsymbol{x}}_1^\intercal & \dots & \hat{\boldsymbol{x}}_N^\intercal\end{pmatrix}^\intercal \in \mathbb{R}^{Nd}$. The covariance of the error of the fused estimate depends on the gain $\boldsymbol{K}$ and on the covariance of the error of $\hat{\boldsymbol{x}}_\mathrm{c}$, $\tilde{\boldsymbol{P}}_\mathrm{c}$:

$$\tilde{\boldsymbol{P}}_F(\boldsymbol{K},\tilde{\boldsymbol{P}}_\mathrm{c}) = \boldsymbol{K}\tilde{\boldsymbol{P}}_\mathrm{c}\boldsymbol{K}^\intercal. \tag{2}$$

If $\tilde{\boldsymbol{P}}_\mathrm{c}$ is not entirely known but is only assumed to belong to some subset of admissible covariance matrices $\mathcal{A}$, then $\tilde{\boldsymbol{P}}_F(\boldsymbol{K},\tilde{\boldsymbol{P}}_\mathrm{c})$ cannot be computed. In this case, an alternative is to provide a *conservative* bound. A couple $(\boldsymbol{K},\boldsymbol{B}_F)$ is said to generate a conservative fusion for a $\mathcal{A}$ if:

$$\forall \boldsymbol{P}_\mathrm{c} \in \mathcal{A}, \qquad \tilde{\boldsymbol{P}}_F(\boldsymbol{K},\boldsymbol{P}_\mathrm{c}) \preceq \boldsymbol{B}_F. \tag{3}$$

In other words, fusing the estimates with the gain $\boldsymbol{K}$ ensures that the covariance of the error is bounded by $\boldsymbol{B}_F$.

CI considers that the covariances of the errors $\tilde{\boldsymbol{P}}_i$ are known but not their cross-covariances $\tilde{\boldsymbol{P}}_{i,j} \triangleq \mathrm{E}\left[\tilde{\boldsymbol{x}}_i\tilde{\boldsymbol{x}}_j^\intercal\right]$. For any $\boldsymbol{\omega} \in \mathcal{K}^N$, CI provides a conservative estimate defined as:

$$\hat{\boldsymbol{x}}_F = \boldsymbol{B}_F\sum_{i=1}^N \omega_i\tilde{\boldsymbol{P}}_i^{-1}\hat{\boldsymbol{x}}_i, \qquad \boldsymbol{B}_F^{-1} = \sum_{i=1}^N \omega_i\tilde{\boldsymbol{P}}_i^{-1}. \tag{4}$$

CI considers that the errors may be completely correlated. In distributed estimation, the estimators integrate independent measurements $\boldsymbol{z}_i$ and have the following structure:

$$\hat{\boldsymbol{x}}_i = (\boldsymbol{I}-\boldsymbol{K}\boldsymbol{H}_i)\hat{\boldsymbol{x}}_i^- + \boldsymbol{K}\boldsymbol{z}_i.$$

Therefore, the errors $\tilde{\boldsymbol{x}}_i$ cannot be perfectly correlated and SCI produces tighter bounds. It considers that the estimation errors are split into a correlated and an uncorrelated component as:

$$\tilde{\boldsymbol{x}}_i = \tilde{\boldsymbol{x}}_i^{(1)} + \tilde{\boldsymbol{x}}_i^{(2)}, \tag{5}$$

where the components $\tilde{\boldsymbol{x}}_i^{(1)}$ are correlated to an unknown degree while the components $\tilde{\boldsymbol{x}}_i^{(2)}$ are uncorrelated between each other and with the $\tilde{\boldsymbol{x}}_i^{(1)}$. The covariances of $\tilde{\boldsymbol{x}}_i^{(1)}$ and $\tilde{\boldsymbol{x}}_i^{(2)}$ are denoted as $\tilde{\boldsymbol{P}}_i^{(1)}$ and $\tilde{\boldsymbol{P}}_i^{(2)}$ (and are assumed known). For any $\boldsymbol{\omega} \in \mathcal{K}^N$, SCI provides a conservative estimate defined as:

$$\hat{\boldsymbol{x}}_F = \boldsymbol{B}_F\sum_{i=1}^N \omega_i\left(\tilde{\boldsymbol{P}}_i^{(1)} + \omega_i\tilde{\boldsymbol{P}}_i^{(2)}\right)^{-1}\hat{\boldsymbol{x}}_i, \tag{6a}$$

$$\boldsymbol{B}_F^{-1} = \sum_{i=1}^N \omega_i\left(\tilde{\boldsymbol{P}}_i^{(1)} + \omega_i\tilde{\boldsymbol{P}}_i^{(2)}\right)^{-1}. \tag{6b}$$

The parameter $\boldsymbol{\omega}$ must be chosen: optimized or empirically tuned with e.g., the methods in [13] or [6].

III. Extended SCI

A. Motivation: Limits of the SCI fusion

CI considers that the errors can be correlated to any degree. In distributed estimation problems, the estimates incorporate independent measurements, and therefore their errors contain independent components. SCI has been proposed to exploit theses independent terms.

Moreover, the state to estimate is often disturbed by an additive process noise $\boldsymbol{w}$. This noise is added to all the estimation errors during the prediction step. Thus, all the errors share a common component. This common component also reduces the space of admissible centralized covariance matrices $\mathcal{A}$: e.g., the errors cannot be perfectly negatively correlated. However, SCI only handles uncorrelated components, so cannot exploit this common componenet. The new ESCI fusion rule is defined to overcome this limitation.

B. Definition of the ESCI fusion

Consider that the estimation errors are split into two components as in (5). The first components $\tilde{\boldsymbol{x}}_i^{(1)}$ are still correlated to an unknown degree. The second components $\tilde{\boldsymbol{x}}_i^{(2)}$ are not assumed uncorrelated, but are assumed to have known second moments. Introduce the centralized errors,

$$\tilde{\boldsymbol{x}}_\mathrm{c}^{(l)} \triangleq \begin{pmatrix}\tilde{\boldsymbol{x}}_1^{(l)\intercal} & \dots & \tilde{\boldsymbol{x}}_N^{(l)\intercal}\end{pmatrix}^\intercal, \qquad l \in \{1,2\},$$

whose covariances and cross-covariances are denoted as $\tilde{\boldsymbol{P}}_\mathrm{c}^{(l)} \triangleq \mathrm{E}\left[\tilde{\boldsymbol{x}}_\mathrm{c}^{(l)}\tilde{\boldsymbol{x}}_\mathrm{c}^{(l)\intercal}\right]$ and $\tilde{\boldsymbol{P}}_\mathrm{c}^{(1,2)} \triangleq \mathrm{E}\left[\tilde{\boldsymbol{x}}_\mathrm{c}^{(1)}\tilde{\boldsymbol{x}}_\mathrm{c}^{(2)\intercal}\right]$. The matrices $\tilde{\boldsymbol{P}}_\mathrm{c}^{(2)}$ and $\tilde{\boldsymbol{P}}_\mathrm{c}^{(1,2)}$ are known, but only the diagonal blocks of $\tilde{\boldsymbol{P}}_\mathrm{c}^{(1)}$ (corresponding to the covariances $\tilde{\boldsymbol{P}}_i^{(1)}$) are known. If $\tilde{\boldsymbol{P}}_\mathrm{c}^{(1,2)} \neq \boldsymbol{0}$, the errors (5) are virtually re-splittable to set $\tilde{\boldsymbol{P}}_\mathrm{c}^{(1,2)} = \boldsymbol{0}$ by letting:

$$\tilde{\boldsymbol{x}}_\mathrm{c}^{(1)} \leftarrow \tilde{\boldsymbol{x}}_\mathrm{c}^{(1)} - \tilde{\boldsymbol{P}}_\mathrm{c}^{(1,2)}(\tilde{\boldsymbol{P}}_\mathrm{c}^{(2)})^{-1}\tilde{\boldsymbol{x}}_\mathrm{c}^{(2)}, \tag{7a}$$

$$\tilde{\boldsymbol{x}}_\mathrm{c}^{(2)} \leftarrow \tilde{\boldsymbol{x}}_\mathrm{c}^{(2)} + \tilde{\boldsymbol{P}}_\mathrm{c}^{(1,2)}(\tilde{\boldsymbol{P}}_\mathrm{c}^{(2)})^{-1}\tilde{\boldsymbol{x}}_\mathrm{c}^{(2)}. \tag{7b}$$

The errors $\tilde{\boldsymbol{x}}_\mathrm{c}^{(1)}$ and $\tilde{\boldsymbol{x}}_\mathrm{c}^{(2)}$ satisfy the same properties: only the off-diagonal blocks of $\tilde{\boldsymbol{P}}_\mathrm{c}^{(1)}$ are unknown. We therefore assume without loss of generality that $\tilde{\boldsymbol{P}}_\mathrm{c}^{(1,2)} = \boldsymbol{0}$. In this splitting, $\tilde{\boldsymbol{x}}_i^{(2)}$ contains all *known* components. In distributed estimation, it will contain the independent measurement noise plus the common process noise as illustrated in the next section.

For any $\boldsymbol{\omega} \in \mathcal{K}^N$, the ESCI fusion is defined as:

$$\hat{\boldsymbol{x}}_F = \boldsymbol{B}_F \boldsymbol{B}_c^{-1} \boldsymbol{H} \hat{\boldsymbol{x}}_c, \qquad \boldsymbol{B}_F^{-1} = \boldsymbol{H}^\intercal \boldsymbol{B}_c^{-1} \boldsymbol{H}, \tag{8a}$$

with:

$$\boldsymbol{H} = \mathbf{1}_N \otimes \boldsymbol{I}_d, \tag{8b}$$

$$\boldsymbol{B}_c^{(1)} = \operatorname{diag}\left(\frac{1}{\omega_1}\tilde{\boldsymbol{P}}_1^{(1)}, \dots, \frac{1}{\omega_N}\tilde{\boldsymbol{P}}_N^{(1)}\right), \tag{8c}$$

$$\boldsymbol{B}_c = \boldsymbol{B}_c^{(1)} + \tilde{\boldsymbol{P}}_c^{(2)}. \tag{8d}$$

The ESCI is a generalization of the SCI in the sense that if $\boldsymbol{P}_c^{(2)}$ is block diagonal, then (8) and (6) define the same fusion.

Theorem 1. *For any $\boldsymbol{\omega} \in \mathcal{K}^N$, the ESCI fusion defined in* (8) *is conservative for the set:*

$$\mathcal{A}_{\text{ESCI}} \triangleq \left\{ \boldsymbol{P}_c^{(1)} + \tilde{\boldsymbol{P}}_c^{(2)} \mid \boldsymbol{P}_c^{(1)} \succeq \mathbf{0} \text{ and } \forall i \in \{1, \dots, N\}, \boldsymbol{P}_i^{(1)} = \tilde{\boldsymbol{P}}_i^{(1)} \right\}. \tag{9}$$

Proof. The proof is the same as for the SCI fusion rule [9]. For any $\boldsymbol{\omega} \in \mathcal{K}^N$, the matrix $\boldsymbol{B}_c^{(1)}$ is a conservative bound on the centralized covariance of $\tilde{\boldsymbol{x}}_c^{(1)}$ [9]. Therefore, $\boldsymbol{B}_c$ is a conservative bound for the centralized covariance of $\tilde{\boldsymbol{x}}_c$. The fusion is then obtained by applying the gain: $\boldsymbol{K} = \left(\boldsymbol{H}^\intercal \boldsymbol{B}_c^{-1} \boldsymbol{H}\right)^{-1} \boldsymbol{H}^\intercal \boldsymbol{B}_c^{-1}$. $\square$

C. *Special case of a common noise*

When the correlated components of the errors come from a common noise $\boldsymbol{w}$, the ESCI expressions can be simplified. Consider that the estimation errors are split as :

$$\tilde{\boldsymbol{x}}_i = \tilde{\boldsymbol{x}}_i^{(1)} + \tilde{\boldsymbol{x}}_i^{(\text{ind})} + \boldsymbol{M}_i \boldsymbol{w}, \tag{10}$$

where the components $\tilde{\boldsymbol{x}}_i^{(1)}$ are correlated to an unknown degree, the components $\tilde{\boldsymbol{x}}_i^{(\text{ind})}$ are uncorrelated between each other, with the $\tilde{\boldsymbol{x}}_i^{(1)}$ and with $\boldsymbol{w}$, the matrices $\boldsymbol{M}_i$ are known, and $\boldsymbol{w}$ is a common independent noise. The covariances of each component are known and denoted as $\boldsymbol{P}_i^{(\text{ind})}$ for $\tilde{\boldsymbol{x}}_i^{(\text{ind})}$ and $\boldsymbol{Q}$ for $\boldsymbol{w}$. In this case the fusion (8) becomes:

$$\hat{\boldsymbol{x}}_F = \boldsymbol{B}_F \sum_{i=1}^{N} \omega_i (\boldsymbol{I}_d - \boldsymbol{S}_1 \boldsymbol{S}_0^{-1} \boldsymbol{M}_i^\intercal) \tilde{\boldsymbol{P}}_i'^{-1} \hat{\boldsymbol{x}}_i, \tag{11a}$$

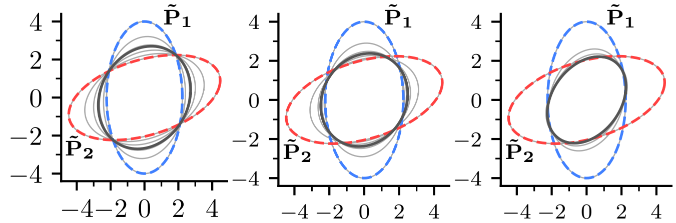
$$\boldsymbol{B}_F^{-1} = \sum_{i=1}^{N} \omega_i \tilde{\boldsymbol{P}}_i'^{-1} - \boldsymbol{S}_1 \boldsymbol{S}_0^{-1} \boldsymbol{S}_1^\intercal, \tag{11b}$$

with $\tilde{\boldsymbol{P}}_i' \triangleq \tilde{\boldsymbol{P}}_i^{(1)} + \omega_i \tilde{\boldsymbol{P}}_i^{(\text{ind})}$ and :

$$\boldsymbol{S}_0 = \sum_{i=1}^{N} \omega_i \boldsymbol{M}_i^\intercal \tilde{\boldsymbol{P}}_i'^{-1} \boldsymbol{M}_i + \boldsymbol{Q}^{-1}, \tag{11c}$$

$$\boldsymbol{S}_1 = \sum_{i=1}^{N} \omega_i \tilde{\boldsymbol{P}}_i'^{-1} \boldsymbol{M}_i. \tag{11d}$$

The advantage of (11) over (8) is that (11) requires to invert $N+1$ matrices of size d while (8) requires the inversion of one matrix of size Nd. As the cost of an inversion of a matrix of size n is a $O(n^3)$, (11) is more efficient.



(a) CI fusion. (b) SCI fusion. (c) ESCI fusion.

Fig. 1: Comparison of the bounds provided by CI, SCI and ESCI. The dotted ellipses represent the covariances $\tilde{\boldsymbol{P}}_1$ and $\tilde{\boldsymbol{P}}_2$, the grey ellipses are the bound obtained with $\boldsymbol{\omega} = \begin{pmatrix} 2k/10 & 1-2k/10 \end{pmatrix}^\intercal$ and $k \in \{0, \dots, 5\}$, and the dark ellipse is the bound that minimizes the trace. The numerical values used are: $\tilde{\boldsymbol{P}}_1^{(1)} = [[1,-2],[-2,5]]$, $\tilde{\boldsymbol{P}}_2^{(1)} = [[9,-1],[-1,1]]$, $\tilde{\boldsymbol{P}}_1^{(\text{ind})} = [[2,0],[0,9]]$, $\tilde{\boldsymbol{P}}_2^{(\text{ind})} = [[9,3],[3,2]]$, $\boldsymbol{Q} = [[2,2],[2,2]]$, and $\boldsymbol{M}_1 = \boldsymbol{M}_2 = \boldsymbol{I}$.

If all the matrices $\boldsymbol{M}_i = \boldsymbol{I}_d$, then (11) simplifies further to:

$$\hat{\boldsymbol{x}}_F = \boldsymbol{B}_0 \sum_{i=1}^{N} \omega_i \tilde{\boldsymbol{P}}_i'^{-1} \hat{\boldsymbol{x}}_i, \qquad \boldsymbol{B}_F = \boldsymbol{B}_0 + \boldsymbol{Q}, \tag{12}$$

where $\boldsymbol{B}_0^{-1} = \sum_{i=1}^{N} \omega_i \tilde{\boldsymbol{P}}_i'^{-1}$. This case is equivalent to first fuse the uncorrupted estimates $\tilde{\boldsymbol{x}}_i^{(1)} + \tilde{\boldsymbol{x}}_i^{(\text{ind})}$ using SCI and then add the noise.

To illustrate the interest of the ESCI fusion, consider the fusion of two estimates whose errors are split according to (10). Theses estimates can be split using CI (without considering the splitting), using SCI (by grouping the correlated component $\boldsymbol{M}_i \boldsymbol{w}$ with the first component), or using ESCI. Fig. 1 compares the bounds obtained with the three fusions. It can be observed that the ESCI bounds are tighter, as expected.

IV. Distributed estimation algorithms

This section presents the integration of the ESCI fusion method into distributed estimation algorithms.

A. *System model*

Consider a system parameterized by a discrete-time state-space model. The state at time $k \in \mathbb{N}$ is denoted as $\boldsymbol{x}(k) \in \mathbb{R}^d$. It is assumed to follow the following linear dynamics:

$$\boldsymbol{x}(0) \sim \mathcal{N}(\boldsymbol{x}_0, \tilde{\boldsymbol{P}}_0), \tag{13a}$$

$$\forall k \in \mathbb{N}, \qquad \boldsymbol{x}(k+1) = \boldsymbol{F}\boldsymbol{x}(k) + \boldsymbol{w}(k+1), \tag{13b}$$

where $\boldsymbol{x}_0$ is the initial state, $\tilde{\boldsymbol{P}}_0$ the covariance matrix of the initial uncertainty, $\boldsymbol{F} \in \mathbb{R}^{d \times d}$ the evolution matrix, and $\boldsymbol{w}(k)$ the process noise at time k. The system is estimated by a network of N agents. The agents are equipped with sensors to observe the state, and have (limited) communication capabilities with their neighbors in the network. As the estimation algorithms are symmetrical between the agents, a focus is made on one particular agent indexed by $i \in \{1, \dots, N\}$.

The set of neighbors of Agent i is denoted as $\mathcal{N}_i$ and its neighborhood as $\mathcal{M}_i \triangleq \mathcal{N}_i \cup \{i\}$. At each time step $k \in \mathbb{N}$, Agent i performs the measurement $\boldsymbol{z}_i(k) \in \mathbb{R}^{m_i}$:

$$\boldsymbol{z}_i(k) = \boldsymbol{H}_i \boldsymbol{x}(k) + \boldsymbol{v}_i(k), \tag{14}$$

where $\boldsymbol{H}_i \in \mathbb{R}^{m_i \times d}$ is the observation matrix and $\boldsymbol{v}_i(k)$ is the observation noise at time k.

The process noise and the observation noises of the agents are assumed: (i) zero-mean, (ii) white, and (iii) uncorrelated between each other and with the error of the initial state. The covariance matrix of the process noise is denoted as $\boldsymbol{Q}$ and that of the measurement noise of Agent i as $\boldsymbol{R}_i$. In this model, all the matrices have been assumed time-independent for the sake of clarity. This work can however be adapted with time-varying matrices.

The estimate of Agent i at time k based on the measurements up to time l is denoted as $\hat{\boldsymbol{x}}_i(k|l)$. The estimation error is $\tilde{\boldsymbol{x}}_i(k|l) \triangleq \hat{\boldsymbol{x}}_i(k|l) - \boldsymbol{x}(k)$. Agent i also estimates a conservative bound, denoted as $\boldsymbol{P}_i(k|l)$, of the covariance of its error. Here, conservative means that:

$$\boldsymbol{P}_i(k|l) - \mathrm{E}\left[\tilde{\boldsymbol{x}}_i(k|l)\tilde{\boldsymbol{x}}_i(k|l)^\intercal\right] \succeq \boldsymbol{0}. \tag{15}$$

At each time step, Agent i exchanges information with its neighbors $j \in \mathcal{N}_i$. Three levels of communication are considered, labeled L1, L2 and L3. They restrict the amount of information that the agents can send (Level L1 being the more restrictive). The information transmitted in each level is given in Table I and is described in the next paragraph.

B. Algorithm description

The distributed estimation algorithm described below was proposed in [9] for Levels L1 and L2. For Level L3, there are slight modifications, the resulting algorithm is the DKF algorithm proposed in [7]. These two algorithms are represented by the diagrams in Fig. 2. The estimation algorithm has four steps detailed below from the perspective of Agent i.

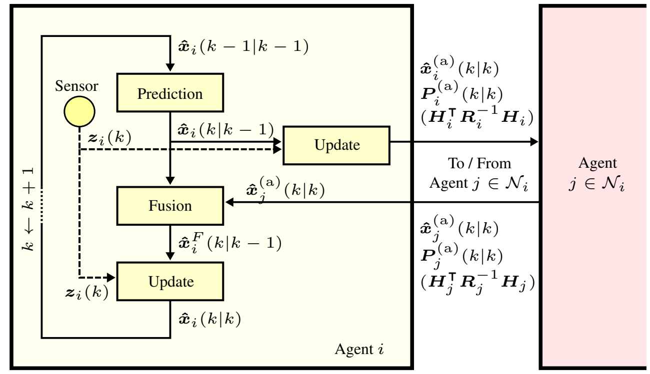
1) **Prediction.** The state is predicted using the evolution model.

$$\hat{\boldsymbol{x}}_i(k|k-1) = \boldsymbol{F}\hat{\boldsymbol{x}}_i(k-1|k-1), \tag{16a}$$
$$\boldsymbol{P}_i(k|k-1) = \boldsymbol{F}\boldsymbol{P}(k-1|k-1)\boldsymbol{F}^\intercal + \boldsymbol{Q}. \tag{16b}$$

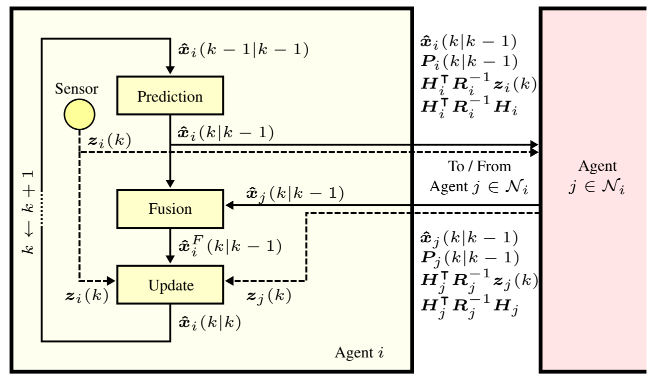
2) **Update.** The prediction is updated using the measurement $\boldsymbol{z}_i(k)$ (as in a Kalman filter).

$$\hat{\boldsymbol{x}}_i^{(\mathrm{a})}(k|k) = \boldsymbol{P}_i^{(\mathrm{a})}(k|k)\boldsymbol{P}_i(k|k-1)^{-1}\hat{\boldsymbol{x}}_i(k|k-1) + \boldsymbol{P}_i^{(\mathrm{a})}(k|k)\boldsymbol{H}_i^\intercal \boldsymbol{R}_i^{-1}\boldsymbol{z}_i(k), \tag{17a}$$
$$\boldsymbol{P}_i^{(\mathrm{a})}(k|k)^{-1} = \boldsymbol{P}_i(k|k-1)^{-1} + \boldsymbol{H}_i^\intercal \boldsymbol{R}_i^{-1}\boldsymbol{H}_i. \tag{17b}$$

The superscript (a) (for autonomous) indicates that this estimate is obtained solely from the measurements of Agent i. Agent i sends this estimate to its neighbors, and receives their estimates. The complete list of parameters transmitted for each communication level is given in Table I.



(a) Without measurement transmissions (Levels L1 and L2).



(b) With measurement transmissions (Level L3).

Fig. 2: Diagrams of the distributed algorithms.

Level	Parameters transmitted
C1	$\hat{\boldsymbol{x}}_i^{(a)}(k\|k)$, $\boldsymbol{P}_i^{(a)}(k\|k)$
C2	$\hat{\boldsymbol{x}}_i^{(a)}(k\|k)$, $\boldsymbol{P}_i^{(a)}(k\|k)$, $\boldsymbol{H}_i^\intercal \boldsymbol{R}_i^{-1}\boldsymbol{H}_i$
C3	$\hat{\boldsymbol{x}}_i(k\|k-1)$, $\boldsymbol{P}_i(k\|k-1)$, $\boldsymbol{H}_i^\intercal \boldsymbol{R}_i^{-1}\boldsymbol{z}_i(k)$, $\boldsymbol{H}_i^\intercal \boldsymbol{R}_i^{-1}\boldsymbol{H}_i$

TABLE I: Parameters transmitted by communication levels.

3) **Fusion.** The estimates received $\hat{\boldsymbol{x}}_j$, $j \in \mathcal{N}_i$, are fused with the prediction $\hat{\boldsymbol{x}}_i(k|k-1)$. This fusion step provides gains $\boldsymbol{K}_j$ for $j \in \mathcal{M}_i$ and a conservative bound $\boldsymbol{B}_F$.

$$\hat{\boldsymbol{x}}_i^F(k|k-1) = \boldsymbol{K}_i\hat{\boldsymbol{x}}_i(k|k-1) + \sum_{j\in\mathcal{N}_i}\boldsymbol{K}_j\hat{\boldsymbol{x}}_j \tag{18a}$$
$$\boldsymbol{P}_i^F(k|k-1) = \boldsymbol{B}_F. \tag{18b}$$

The details of the fusion step are given in Section IV-C.

4) **Update.** The fused estimate is updated using the measurement $\boldsymbol{z}_i(k)$ as in step 2.

$$\hat{\boldsymbol{x}}_i(k|k) = \boldsymbol{P}_i(k|k)\boldsymbol{P}_i^F(k|k-1)^{-1}\hat{\boldsymbol{x}}_i(k|k-1) + \boldsymbol{P}_i(k|k)\boldsymbol{H}_i^\intercal \boldsymbol{R}_i^{-1}\boldsymbol{z}_i(k), \tag{19a}$$
$$\boldsymbol{P}_i(k|k)^{-1} = \boldsymbol{P}_i^F(k|k-1)^{-1} + \boldsymbol{H}_i^\intercal \boldsymbol{R}_i^{-1}\boldsymbol{H}_i. \tag{19b}$$

In Level L3, the measurements of the neighbors are optimally fused during the update of Step 4. Therefore, there is no need to perform Step 2, and (19) becomes:

$$\hat{\boldsymbol{x}}_i(k|k) = \boldsymbol{P}_i(k|k)\boldsymbol{P}_i^F(k|k-1)^{-1}\hat{\boldsymbol{x}}_i(k|k-1) + \boldsymbol{P}_i(k|k)\sum_{j\in\mathcal{M}_i}\boldsymbol{H}_j^\intercal\boldsymbol{R}_j^{-1}\boldsymbol{z}_j(k), \tag{20a}$$

$$\boldsymbol{P}_i(k|k)^{-1} = \boldsymbol{P}_i^F(k|k-1)^{-1} + \sum_{j\in\mathcal{M}_i}\boldsymbol{H}_j^\intercal\boldsymbol{R}_j^{-1}\boldsymbol{H}_j. \tag{20b}$$

C. Fusions

This paragraph details the fusion step of the distributed estimation algorithm. It is during this step that ESCI is used instead of the classical fusions to exploit the process noise.

1) Level L1: With Level L1, the fusion step can only be performed using CI (4). Neither the independent components induced by the measurements, nor the correlated components induced by the process noise can be used.

2) Level L2: With Level L2, the authors of [9] propose to fuse the estimates using SCI in order to exploit the independent component. Indeed, the error on the transmitted estimate is:

$$\tilde{\boldsymbol{x}}_j^{(\mathrm{a})}(k|k) = \boldsymbol{P}_j^{(\mathrm{a})}(k|k)\boldsymbol{P}_j(k|k-1)^{-1}\tilde{\boldsymbol{x}}_j(k|k-1) + \boldsymbol{P}_j^{(\mathrm{a})}(k|k)\boldsymbol{H}_j^\intercal\boldsymbol{R}_j^{-1}\boldsymbol{v}_j(k). \tag{21}$$

It is split as (5) with:

$$\tilde{\boldsymbol{x}}_j^{(1)} = \boldsymbol{P}_j^{(\mathrm{a})}(k|k)\boldsymbol{P}_j(k|k-1)^{-1}\tilde{\boldsymbol{x}}_j(k|k-1), \tag{22a}$$

$$\tilde{\boldsymbol{x}}_j^{(2)} = \boldsymbol{P}_j^{(\mathrm{a})}(k|k)\boldsymbol{H}_j^\intercal\boldsymbol{R}_j^{-1}\boldsymbol{v}_j(k). \tag{22b}$$

The covariances of both components are computable from the parameters transmitted. This decomposition takes advantage from the independent measurements but ignores the fact that the terms $\tilde{\boldsymbol{x}}_j^{(1)}$ all contain the process noise $\boldsymbol{w}(k)$:

$$\tilde{\boldsymbol{x}}_j(k|k-1) = \boldsymbol{F}\tilde{\boldsymbol{x}}_j(k-1|k-1) - \boldsymbol{w}(k). \tag{23}$$

The error (23) can be rexpressed as (10) with:

$$\tilde{\boldsymbol{x}}_j^{(1)} = \boldsymbol{P}_j^{(\mathrm{a})}(k|k)\boldsymbol{P}_j(k|k-1)^{-1}\boldsymbol{F}\tilde{\boldsymbol{x}}_j(k-1|k-1), \tag{24a}$$

$$\tilde{\boldsymbol{x}}_j^{(\mathrm{ind})} = \boldsymbol{P}_j^{(\mathrm{a})}(k|k)\boldsymbol{H}_j^\intercal\boldsymbol{R}_j^{-1}\boldsymbol{v}_j(k), \tag{24b}$$

$$\boldsymbol{M}_j = -\boldsymbol{P}_j^{(\mathrm{a})}(k|k)\boldsymbol{P}_j(k|k-1)^{-1}. \tag{24c}$$

All the terms required for applying the ESCI fusion method (11) are also computable from the parameters transmitted.

3) Level L3: With Level L3, the authors of [7] propose to fuse the estimates using CI. However, as expressed in (23), the predicted estimates are all corrupted by the process noise. It is therefore more interesting to fuse them using ESCI (12). This fusion rule is equivalent to fuse the estimates $\hat{\boldsymbol{x}}_j(k-1|k-1)$ using CI before performing the prediction step.

V. Simulations

To illustrate the interest of the ESCI fusion method, the algorithms are applied in an example inspired by SAR missions. A distress signal is sent by an emitter of unknown position and is received by a network of satellites. The objective is to estimate the position $\boldsymbol{p} \in \mathbb{R}^3$ of the emitter. To do so, the satellites measure the times of reception of the distress signal

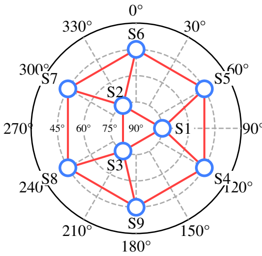


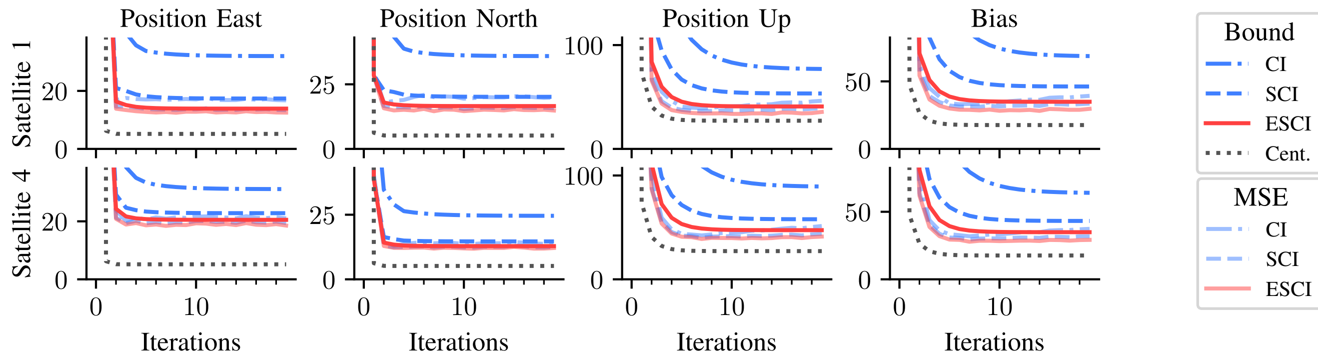
Fig. 3: Skyplot (azimuth / elevation) of the positions of the satellites. (A satellite in the center has an elevation of 90°, i.e., is at zenith.) The lines represent the edges of the network.

and deduce the *pseudo-ranges* from the emitter. The pseudo-range is a biased version of the distance obtained by measuring the time-of-flight of the signal. The bias comes from the fact that the time of emission of the signal is unknown: the unknown clock-offet τ generates a bias $\beta \triangleq c\tau$ on the range measurement. In this context, the state to estimate is the location of the emitter and the bias: $\boldsymbol{x} = \begin{pmatrix}\boldsymbol{p}^\intercal & \beta\end{pmatrix}^\intercal \in \mathbb{R}^4$. This state is modeled with a slow dynamic, the evolution matrix is $\boldsymbol{F} = \boldsymbol{I}_4$ and the process noise has covariance $\boldsymbol{Q} = \sigma_w^2\boldsymbol{I}_4$ with $\sigma_w = 5$ m. The observation matrix of Satellite i is $\boldsymbol{H}_i = \begin{bmatrix}\boldsymbol{u}_i & 1\end{bmatrix} \in \mathbb{R}^{1\times 4}$ where $\boldsymbol{u}_i$ is the unit vector pointing from the satellite to the emitter. The variance of the measurement noise is $R_i = \sigma_m^2$, with $\sigma_m = 10$ m. The network of satellites is represented with a skyplot in Fig. 3.

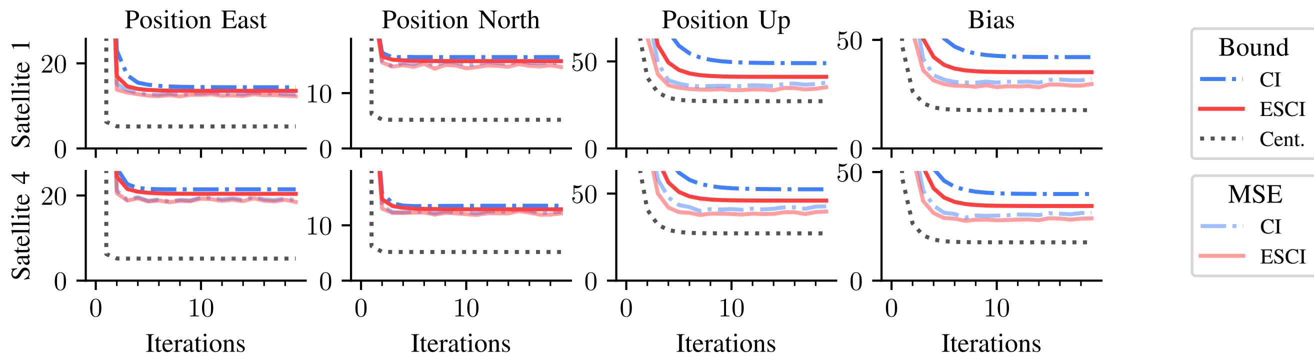
The algorithms without and with measurement transmissions have been applied with the standard fusion methods and with the new ESCI fusion rule. The estimation using a centralized Kalman filter was also computed in order to visualize the smallest reachable error. This filter is idealistic since it requires to track the cross-covariances between all the errors. For each fusion method the parameter $\boldsymbol{\omega}$ has been optimized to minimize the trace of the bound. Fig. 4 presents the evolution of the estimated variance bounds and the MSEs computed over $10,000$ runs of 20 iterations. We observe that for all fusion methods the MSEs are, as expected, lower than the bounds which confirms the conservativeness. For both algorithms, the ESCI fusion rule provides tighter bounds. For the algorithm without measurement transmissions, we observe reductions of the variance bounds with respect to the SCI fusion rule of about 19% for Sat. 1 and 11% for Sat. 4 in the horizontal plane (North / East) and of about 23% for Sat. 1 and 18% for Sat. 4 in the vertical plane (Up). For the algorithm with measurement transmissions, the reductions of the variance bounds with respect to the CI fusion rule are of about 5% for both satellites in the horizontal plane (North / East) and of about 16% for Sat. 1 and 12% for Sat. 4 in the vertical plane (Up). This example confirms that exploiting the common process noise helps to improve the accuracy.

VI. Conclusion

This paper has introduced a new conservative fusion rule inspired by the SCI fusion rule. This new method is designed



(a) Without measurement transmissions (Levels L1 and L2).



(b) With measurement transmissions (Level L3).

Fig. 4: Estimated variance bounds (matte curves) and MSEs (semi-transparent curves) for Satellite 1 and Satellite 4. The bound labeled 'Cent.' were obtained with a centralized Kalman Filter to represent the optimal reachable performances.

to exploit all the known components in the estimation errors, uncorrelated or not. When applied to distributed estimation problems, this fusion rule exploits the commonly observed process noise to produce tighter bounds. This fusion rule can be used with standard distributed algorithms, such as the DKF, with no additional communication requirements or complexity with respect to the current fusion methods. The simulations have confirmed the interest of this fusion rule with significant bound reductions, particularly important for applications such as SAR.

References

[1] J. Ajgl and O. Straka. Rectification of partitioned covariance intersection. In *2019 American Control Conference (ACC)*, pages 5786–5791. IEEE, 2019.

[2] J. Ajgl and O. Straka. Inverse covariance intersection fusion of multiple estimates. In *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*, pages 1–8. IEEE, 2020.

[3] Y. Bar-Shalom and L. Campo. The effect of the common process noise on the two-sensor fused-track covariance. *IEEE Transactions on aerospace and electronic systems*, (6):803–805, 1986.

[4] F. S. Cattivelli and A. H. Sayed. Diffusion strategies for distributed Kalman filtering and smoothing. *IEEE Transactions on automatic control*, 55(9):2069–2084, 2010.

[5] R. Forsling, A. Hansson, F. Gustafsson, Z. Sjanic, J. Löfberg, and G. Hendeby. Conservative linear unbiased estimation under partially known covariances. *IEEE Transactions on Signal Processing*, 70:3123–3135, 2022.

[6] D. Franken and A. Hupper. Improved fast covariance intersection for distributed data fusion. In *2005 7th International Conference on Information Fusion*, volume 1, pages 7–pp. IEEE, 2005.

[7] J. Hu, L. Xie, and C. Zhang. Diffusion Kalman filtering based on covariance intersection. *IEEE Transactions on Signal Processing*, 60(2):891–902, 2011.

[8] S. J. Julier and J. K. Uhlmann. A non-divergent estimation algorithm in the presence of unknown correlations. In *Proceedings of the 1997 American Control Conference*, volume 4, pages 2369–2373. IEEE, 1997.

[9] S. J. Julier and J. K. Uhlmann. General decentralized data fusion with covariance intersection (CI). *Handbook of Multisensor Data Fusion*, 2001.

[10] S. J. Julier and J. K. Uhlmann. Using covariance intersection for SLAM. *Robotics and Autonomous Systems*, 55(1):3–20, 2007.

[11] H. Li and F. Nashashibi. Cooperative multi-vehicle localization using split covariance intersection filter. *IEEE Intelligent transportation systems magazine*, 5(2):33–44, 2013.

[12] A. Lima, P. Bonnifait, V. Cherfaoui, and J. Al Hage. Data fusion with split covariance intersection for cooperative perception. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pages 1112–1118. IEEE, 2021.

[13] W. Niehsen. Information fusion based on fast covariance intersection filtering. In *Proceedings of the Fifth International Conference on Information Fusion. FUSION 2002.(IEEE Cat. No. 02EX5997)*, volume 2, pages 901–904. IEEE, 2002.

[14] B. Noack, J. Sijs, and U. D. Hanebeck. Inverse covariance intersection: New insights and properties. In *2017 20th International Conference on Information Fusion (Fusion)*, pages 1–8. IEEE, 2017.

[15] B. Noack, J. Sijs, M. Reinhardt, and U. D. Hanebeck. Decentralized data fusion with inverse covariance intersection. *Automatica*, 79:35–41, 2017.

[16] A. Petersen and M.-A. Beyer. Partitioned covariance intersection. In *Proceedings of International Symposium Information on Ships*. Citeseer, 2011.