Decentralized Data Fusion with Inverse Covariance Intersection

Benjamin Noack [a], Joris Sijs [b], Marc Reinhardt [a], Uwe D. Hanebeck [a]

[a] *Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology (KIT), Germany.*

[b] *TNO Technical Sciences, Den Haag, The Netherlands.*

Abstract

In distributed and decentralized state estimation systems, fusion methods are employed to systematically combine multiple estimates of the state into a single, more accurate estimate. An often encountered problem in the fusion process relates to unknown common information that is shared by the estimates to be fused and is responsible for correlations. If the correlation structure is unknown to the fusion method, conservative strategies are typically pursued. As such, the parameterization introduced by the ellipsoidal intersection method has been a novel approach to describe unknown correlations, though suitable values for these parameters with proven consistency have not been identified yet. In this article, an extension of ellipsoidal intersection is proposed that guarantees consistent fusion results in the presence of unknown common information. The bound used by the novel approach corresponds to computing an outer ellipsoidal bound on the intersection of inverse covariance ellipsoids. As a major advantage of this inverse covariance intersection method, fusion results prove to be more accurate than those provided by the well-known covariance intersection method.

Key words: State Estimation, Data Fusion, Sensor Fusion, Decentralized Kalman Filtering, Covariance Intersection.

1 Introduction

In typical network-based sensor systems (Hall et al., 2013), it is often not a single instance that computes an estimate but a multiplicity of nodes, each of which is equipped with its own state estimation system. In order to improve the estimation quality, the locally computed estimates are fused into a single estimate. The problem of fusing estimates has received particularly strong attention in the context of distributed target-tracking applications (Bar-Shalom and Li, 1995), where the treatment of cross-correlations between estimates continues to pose a challenging problem. An optimal fusion method typically requires a full-rate communication to the data sink or an augmented state representation as studied in Chong et al. (2014). Alternatives to fusion methods are consensus and diffusion schemes—see, for instance, Battistelli et al. (2015), and Cattivelli and Sayed (2010), respectively. In these cases, the nodes reach a consent on a global estimate by employing specific averaging techniques, which typically also require frequent communication and are not designed to minimize the fused error covariance matrix.

In particular for fully decentralized sensor networks, fusion rules for arbitrary estimates have been proposed that provide suboptimal but consistent fusion results irrespective of the underlying correlation structure. The requirement of consistency guarantees that the actual error covariance matrix is not underestimated but conservatively bounded. A well-known conservative fusion rule is *covariance intersection* (CI), which has been introduced by Julier and Uhlmann (1997) and has given rise to many further developments and applications such as in Deng et al. (2012) or in Hu et al. (2012). Reinhardt et al. (2015) have shown that CI tightly bounds the entirety of possible error covariance matrices. More precisely, if the correlations between two estimates to be fused are entirely unknown, CI encompasses the optimal fusion method in terms of a minimum mean-squared error and also other optimality criteria. However, CI often provides too conservative fusion results as typical estimation tasks and communication networks, in general, prevent extremal correlation terms to occur.

A recent advancement is *ellipsoidal intersection* (EI) (Sijs and Lazar, 2012) that employs a common error term to model unknown correlations and reports a far less conservative result as compared to CI. The EI algo-

Email addresses: benjamin.noack@ieee.org (Benjamin Noack), joris.sijs@tno.nl (Joris Sijs), marc.reinhardt@ieee.org (Marc Reinhardt), uwe.hanebeck@ieee.org (Uwe D. Hanebeck).

Published in Automatica
DOI: 10.1016/j.automatica.2017.01.019
 21 December 2016

rithm is based on a special decomposition of the fusion problem, which is revisited in this article. Although consistency has not been shown yet for EI, it proves to be effective in numerous applications. Therefore, this article is concerned with the question under which conditions EI guarantees to be a consistent fusion method. For this purpose, a detailed study on the structure of the fusion problem and its solutions is presented. It turns out that EI guarantees consistent estimates only under rather restrictive conditions. By identifying the missing pieces, a consistent extension is derived that opens up EI to a wide variety of fusion tasks. The novel approach achieves a conservative fusion result by computing a bound on the intersection of inverse covariance ellipsoids, which gives reason to name it *inverse covariance intersection* (ICI).

2 Conventions and Preliminaries

Underlined variables $\underline{x} \in \mathbb{R}^n$ denote real-valued vectors, and lowercase boldface letters $\underline{\mathbf{x}}$ are used for random quantities. Matrices are written in uppercase boldface letters $\mathbf{C} \in \mathbb{R}^{n \times n}$, and $\mathbf{C}^{-1}$ and $\mathbf{C}^{\mathrm{T}}$ are the inverse and transpose, respectively. $(\mathbf{C})_{ij}$ is the entry in the i-th row and j-th column of matrix $\mathbf{C}$. The inequality $\mathbf{C} \geq \mathbf{C}'$ means that the difference $\mathbf{C} - \mathbf{C}'$ is positive semi-definite. By the tuple $(\hat{\underline{\mathbf{x}}}, \mathbf{C})$, we denote an estimate with mean $\hat{\underline{\mathbf{x}}}$ and error covariance matrix $\mathbf{C} = \mathrm{E}\left[\tilde{\underline{\mathbf{x}}}\tilde{\underline{\mathbf{x}}}^{\mathrm{T}}\right]$, where $\tilde{\underline{\mathbf{x}}} = \hat{\underline{\mathbf{x}}} - \underline{\mathbf{x}}$ is the error with respect to the state $\underline{\mathbf{x}}$ to be estimated. The matrix $\mathbf{I}$ denotes the identity matrix of appropriate dimension. An ellipsoid with center $\hat{\underline{c}}$ and shape matrix $\mathbf{X}$ is denoted by $\mathcal{E}(\hat{\underline{c}}, \mathbf{X}) = \left\{ \underline{x} \in \mathbb{R}^n \mid (\hat{\underline{c}} - \underline{x})^{\mathrm{T}} \mathbf{X}^{-1} (\hat{\underline{c}} - \underline{x}) \leq 1 \right\}$.

3 Analysis of the Fusion Problem

The challenge of fusing multiple estimates is typically addressed from the perspective of a two-sensor fusion problem. As such, we consider two unbiased estimates $(\hat{\underline{\mathbf{x}}}_{\mathsf{A}}, \mathbf{C}_{\mathsf{A}})$ and $(\hat{\underline{\mathbf{x}}}_{\mathsf{B}}, \mathbf{C}_{\mathsf{B}})$ provided by the sensor nodes A and B. In order to compute the fused estimate $(\hat{\underline{\mathbf{x}}}_{\mathrm{fus}}, \mathbf{C}_{\mathrm{fus}})$, the linear combination

$$\hat{\underline{\mathbf{x}}}_{\mathrm{fus}} = \mathbf{K}_{\mathrm{fus}} \hat{\underline{\mathbf{x}}}_{\mathsf{A}} + \mathbf{L}_{\mathrm{fus}} \hat{\underline{\mathbf{x}}}_{\mathsf{B}} \tag{1}$$

with the fusion gains $\mathbf{K}_{\mathrm{fus}}$ and $\mathbf{L}_{\mathrm{fus}}$ is considered. The corresponding error covariance matrix yields

$$\begin{aligned} \mathbf{C}_{\mathrm{fus}} = \mathbf{K}_{\mathrm{fus}} \mathbf{C}_{\mathsf{A}} \mathbf{K}_{\mathrm{fus}}^{\mathrm{T}} + \mathbf{K}_{\mathrm{fus}} \mathbf{C}_{\mathsf{AB}} \mathbf{L}_{\mathrm{fus}}^{\mathrm{T}} \\ + \mathbf{L}_{\mathrm{fus}} \mathbf{C}_{\mathsf{BA}} \mathbf{K}_{\mathrm{fus}}^{\mathrm{T}} + \mathbf{L}_{\mathrm{fus}} \mathbf{C}_{\mathsf{B}} \mathbf{L}_{\mathrm{fus}}^{\mathrm{T}} \, , \end{aligned} \tag{2}$$

which also depends on the cross-covariance term $\mathbf{C}_{\mathsf{AB}} = \mathbf{C}_{\mathsf{BA}}^{\mathrm{T}} = \mathrm{E}\left[\tilde{\underline{\mathbf{x}}}_{\mathsf{A}} \tilde{\underline{\mathbf{x}}}_{\mathsf{B}}^{\mathrm{T}}\right]$ The gains are typically designed to minimize the trace of the error matrix (2), which corresponds to minimizing the mean squared error. However, in decentralized estimation networks, the cross-covariance matrix $\mathbf{C}_{\mathsf{AB}}$ is often unknown to the fusion method, and only an approximation of $\mathbf{C}_{\mathrm{fus}}$ is attainable. In order to not underestimate the actual error matrix, the fusion result $(\hat{\underline{\mathbf{x}}}_{\mathrm{fus}}, \mathbf{C}_{\mathrm{fus}})$ is required to be *consistent*.

Definition 1 (Consistency)**.** An estimate $(\hat{\underline{\mathbf{x}}}, \mathbf{C})$ is *consistent* if the actual error covariance matrix is bounded by the reported covariance matrix $\mathbf{C}$, i.e., $\mathbf{C} \geq \mathrm{E}\left[\tilde{\underline{\mathbf{x}}}\, \tilde{\underline{\mathbf{x}}}^{\mathrm{T}}\right]$ with $\tilde{\underline{\mathbf{x}}} = (\hat{\underline{\mathbf{x}}} - \underline{\mathbf{x}})$.

3.1 Common Information

For the derivation of EI, Sijs and Lazar (2012) have proposed to utilize a special parameterization and decomposition of the estimates to be fused. The means and covariance matrices of the estimates are decomposed into

$$\hat{\underline{\mathbf{x}}}_{\mathsf{A}} = \mathbf{C}_{\mathsf{A}} \left((\mathbf{C}_{\mathsf{A}}^{\mathrm{I}})^{-1} \hat{\underline{\mathbf{x}}}_{\mathsf{A}}^{\mathrm{I}} + \mathbf{\Gamma}^{-1} \hat{\underline{\gamma}} \right) , \tag{3a}$$

$$\hat{\underline{\mathbf{x}}}_{\mathsf{B}} = \mathbf{C}_{\mathsf{B}} \left((\mathbf{C}_{\mathsf{B}}^{\mathrm{I}})^{-1} \hat{\underline{\mathbf{x}}}_{\mathsf{B}}^{\mathrm{I}} + \mathbf{\Gamma}^{-1} \hat{\underline{\gamma}} \right) \tag{3b}$$

and

$$\mathbf{C}_{\mathsf{A}}^{-1} = (\mathbf{C}_{\mathsf{A}}^{\mathrm{I}})^{-1} + \mathbf{\Gamma}^{-1} \, , \tag{4a}$$

$$\mathbf{C}_{\mathsf{B}}^{-1} = (\mathbf{C}_{\mathsf{B}}^{\mathrm{I}})^{-1} + \mathbf{\Gamma}^{-1} \, , \tag{4b}$$

respectively. This parameterization is based on a common estimate $(\hat{\underline{\gamma}}, \mathbf{\Gamma})$ that is shared by the sensor nodes A and B. Exclusive information is represented by the partial estimate $(\hat{\underline{\mathbf{x}}}_{\mathsf{A}}^{\mathrm{I}}, \mathbf{C}_{\mathsf{A}}^{\mathrm{I}})$ at sensor node A and by $(\hat{\underline{\mathbf{x}}}_{\mathsf{B}}^{\mathrm{I}}, \mathbf{C}_{\mathsf{B}}^{\mathrm{I}})$ at sensor node B. In particular, (3a) and (4a) correspond to the fusion of $(\hat{\underline{\mathbf{x}}}_{\mathsf{A}}^{\mathrm{I}}, \mathbf{C}_{\mathsf{A}}^{\mathrm{I}})$ and $(\hat{\underline{\gamma}}, \mathbf{\Gamma})$ with the gains $\mathbf{K}_{\mathrm{fus}} = \mathbf{C}_{\mathsf{A}} (\mathbf{C}_{\mathsf{A}}^{\mathrm{I}})^{-1}$ and $\mathbf{L}_{\mathrm{fus}} = \mathbf{C}_{\mathsf{A}} \mathbf{\Gamma}^{-1}$ and zero cross-covariance matrix. The same applies to (3b) and (4b) with $\mathbf{K}_{\mathrm{fus}} = \mathbf{C}_{\mathsf{B}} (\mathbf{C}_{\mathsf{B}}^{\mathrm{I}})^{-1}$ and $\mathbf{L}_{\mathrm{fus}} = \mathbf{C}_{\mathsf{B}} \mathbf{\Gamma}^{-1}$. In the above decompositions, the errors $\tilde{\underline{\mathbf{x}}}_{\mathsf{A}}^{\mathrm{I}}$, $\tilde{\underline{\mathbf{x}}}_{\mathsf{B}}^{\mathrm{I}}$, and $\tilde{\underline{\gamma}}$ related to the partial estimates $\hat{\underline{\mathbf{x}}}_{\mathsf{A}}^{\mathrm{I}}$, $\hat{\underline{\mathbf{x}}}_{\mathsf{B}}^{\mathrm{I}}$, and $\hat{\underline{\gamma}}$ each have zero mean and are assumed to be mutually uncorrelated, i.e.,

$$\mathrm{E}[\tilde{\underline{\mathbf{x}}}_{\mathsf{A}}^{\mathrm{I}} (\tilde{\underline{\mathbf{x}}}_{\mathsf{B}}^{\mathrm{I}})^{\mathrm{T}}] = \mathrm{E}[\tilde{\underline{\mathbf{x}}}_{\mathsf{A}}^{\mathrm{I}}\, \tilde{\underline{\gamma}}^{\mathrm{T}}] = \mathrm{E}[\tilde{\underline{\mathbf{x}}}_{\mathsf{B}}^{\mathrm{I}}\, \tilde{\underline{\gamma}}^{\mathrm{T}}] = \mathbf{0} \, . \tag{5}$$

By considering (3), it can be seen that the estimation errors $\tilde{\underline{\mathbf{x}}}_{\mathsf{A}}$ and $\tilde{\underline{\mathbf{x}}}_{\mathsf{B}}$ obey the same decompositions

$$\begin{aligned} \tilde{\underline{\mathbf{x}}}_{\mathsf{A}} = \hat{\underline{\mathbf{x}}}_{\mathsf{A}} - \underline{\mathbf{x}} &= \mathbf{C}_{\mathsf{A}} \left((\mathbf{C}_{\mathsf{A}}^{\mathrm{I}})^{-1} \hat{\underline{\mathbf{x}}}_{\mathsf{A}}^{\mathrm{I}} + \mathbf{\Gamma}^{-1} \hat{\underline{\gamma}} \right) - \underline{\mathbf{x}} \\ &= \mathbf{C}_{\mathsf{A}} \left((\mathbf{C}_{\mathsf{A}}^{\mathrm{I}})^{-1} (\hat{\underline{\mathbf{x}}}_{\mathsf{A}}^{\mathrm{I}} - \underline{\mathbf{x}}) + \mathbf{\Gamma}^{-1} (\hat{\underline{\gamma}} - \underline{\mathbf{x}}) \right) \\ &= \mathbf{C}_{\mathsf{A}} \left((\mathbf{C}_{\mathsf{A}}^{\mathrm{I}})^{-1} \tilde{\underline{\mathbf{x}}}_{\mathsf{A}}^{\mathrm{I}} + \mathbf{\Gamma}^{-1} \tilde{\underline{\gamma}} \right) , \end{aligned} \tag{6a}$$

$$\tilde{\underline{\mathbf{x}}}_{\mathsf{B}} = \mathbf{C}_{\mathsf{B}} \left((\mathbf{C}_{\mathsf{B}}^{\mathrm{I}})^{-1} \tilde{\underline{\mathbf{x}}}_{\mathsf{B}}^{\mathrm{I}} + \mathbf{\Gamma}^{-1} \tilde{\underline{\gamma}} \right) . \tag{6b}$$

Due to (5), it is only the common partial estimate $(\hat{\underline{\gamma}}, \mathbf{\Gamma})$, which is responsible for correlations between the estimation errors (6a) and (6b), while the partial errors $\tilde{\underline{\mathbf{x}}}_{\mathsf{A}}^{\mathrm{I}}$ and $\tilde{\underline{\mathbf{x}}}_{\mathsf{B}}^{\mathrm{I}}$ are uncorrelated. The cross-covariance matrix yields

$$\mathbf{C}_{\mathsf{AB}} = \mathbf{C}_{\mathsf{A}} \mathbf{\Gamma}^{-1} \mathrm{E}\left[\tilde{\underline{\gamma}}\, \tilde{\underline{\gamma}}^{\mathrm{T}}\right] \mathbf{\Gamma}^{-1} \mathbf{C}_{\mathsf{B}}^{\mathrm{T}} = \mathbf{C}_{\mathsf{A}} \mathbf{\Gamma}^{-1} \mathbf{C}_{\mathsf{B}} \, . \tag{7}$$

This result has a strong implication for the design of fusion methods. The set of possible error covariance matrices (2) is consequently parameterized by $\mathbf{\Gamma}$, and only those $\mathbf{\Gamma}$ come into consideration that are *admissible*.

Definition 2 (Admissibility)**.** Given two estimates $(\hat{\underline{\mathbf{x}}}_\mathsf{A}, \mathbf{C}_\mathsf{A})$ and $(\hat{\underline{\mathbf{x}}}_\mathsf{B}, \mathbf{C}_\mathsf{B})$, the positive definite matrix $\mathbf{\Gamma}$ is *admissible* if there are decompositions (3) and (4).

3.2 Optimal Fusion with Common Information

In the presence of a common component $(\hat{\underline{\gamma}}, \mathbf{\Gamma})$, the optimal and consistent combination of the local estimates from nodes A and B corresponds to the optimal fusion of the three independent estimates $\hat{\underline{\mathbf{x}}}^\mathrm{I}_\mathsf{A}$, $\hat{\underline{\mathbf{x}}}^\mathrm{I}_\mathsf{B}$, $\hat{\underline{\gamma}}$ according to (1). The result can be simplified to

$$\begin{aligned}\hat{\underline{\mathbf{x}}}_\mathbf{\Gamma} &= \mathbf{K}_\mathrm{fus}\,\hat{\underline{\mathbf{x}}}^\mathrm{I}_\mathsf{A} + \mathbf{L}_\mathrm{fus}\,\hat{\underline{\mathbf{x}}}^\mathrm{I}_\mathsf{B} + \mathbf{M}_\mathrm{fus}\,\hat{\underline{\gamma}} \\ &= \mathbf{C}_\mathbf{\Gamma}\left((\mathbf{C}^\mathrm{I}_\mathsf{A})^{-1}\hat{\underline{\mathbf{x}}}^\mathrm{I}_\mathsf{A} + (\mathbf{C}^\mathrm{I}_\mathsf{B})^{-1}\hat{\underline{\mathbf{x}}}^\mathrm{I}_\mathsf{B} + \mathbf{\Gamma}^{-1}\hat{\underline{\gamma}}\right) \\ &= \mathbf{C}_\mathbf{\Gamma}\left(\mathbf{C}^{-1}_\mathsf{A}\hat{\underline{\mathbf{x}}}_\mathsf{A} + \mathbf{C}^{-1}_\mathsf{B}\hat{\underline{\mathbf{x}}}_\mathsf{B} - \mathbf{\Gamma}^{-1}\hat{\underline{\gamma}}\right)\end{aligned} \quad (8a)$$

with the error covariance matrix

$$\begin{aligned}\mathbf{C}^{-1}_\mathbf{\Gamma} &= (\mathbf{C}^\mathrm{I}_\mathsf{A})^{-1} + (\mathbf{C}^\mathrm{I}_\mathsf{B})^{-1} + \mathbf{\Gamma}^{-1} \\ &= \mathbf{C}^{-1}_\mathsf{A} + \mathbf{C}^{-1}_\mathsf{B} - \mathbf{\Gamma}^{-1}\ ,\end{aligned} \quad (8b)$$

were the optimal fusion gains $\mathbf{K}_\mathrm{fus} = \mathbf{C}_\mathbf{\Gamma}(\mathbf{C}^\mathrm{I}_\mathsf{A})^{-1}$, $\mathbf{L}_\mathrm{fus} = \mathbf{C}_\mathbf{\Gamma}(\mathbf{C}^\mathrm{I}_\mathsf{B})^{-1}$, and $\mathbf{M}_\mathrm{fus} = \mathbf{C}_\mathbf{\Gamma}\mathbf{\Gamma}^{-1}$ have been employed.

In the special case $\mathbf{C}_\mathsf{A} = \mathbf{C}^\mathrm{I}_\mathsf{A}$, $\mathbf{C}_\mathsf{B} = \mathbf{C}^\mathrm{I}_\mathsf{B}$, and an absent $(\hat{\underline{\gamma}}, \mathbf{\Gamma})$, the estimates have independent errors, i.e., $\mathbf{C}_\mathsf{AB} = \mathbf{0}$. The fused estimate $\hat{\underline{\mathbf{x}}}_\mathrm{in}$ and its error covariance matrix $\mathbf{C}_\mathrm{in} = \mathrm{E}[\tilde{\underline{\mathbf{x}}}^\mathrm{T}_\mathrm{in}\,\tilde{\underline{\mathbf{x}}}_\mathrm{in}]$ are then given by

$$\hat{\underline{\mathbf{x}}}_\mathrm{in} = \mathbf{K}_\mathrm{in}\hat{\underline{\mathbf{x}}}_\mathsf{A} + \mathbf{L}_\mathrm{in}\hat{\underline{\mathbf{x}}}_\mathsf{B}\ , \quad (9a)$$

and

$$\mathbf{C}_\mathrm{in} = \mathbf{K}_\mathrm{in}\mathbf{C}_\mathsf{A}\mathbf{K}^\mathrm{T}_\mathrm{in} + \mathbf{L}_\mathrm{in}\mathbf{C}_\mathsf{B}\mathbf{L}^\mathrm{T}_\mathrm{in} = (\mathbf{C}^{-1}_\mathsf{A} + \mathbf{C}^{-1}_\mathsf{B})^{-1}. \quad (9b)$$

The gains $\mathbf{K}_\mathrm{in} = \mathbf{C}_\mathrm{in}\mathbf{C}^{-1}_\mathsf{A}$ and $\mathbf{L}_\mathrm{in} = \mathbf{C}_\mathrm{in}\mathbf{C}^{-1}_\mathsf{B}$ are designed to minimize the mean squared error, which corresponds to minimizing the trace of (9b).

A *naive* fusion rule that ignores common information $(\hat{\underline{\gamma}}, \mathbf{\Gamma})$ and erroneously assumes $\mathbf{C}_\mathsf{AB} = \mathbf{0}$ in (7) computes the result in (9). However, by inserting (3) and (4) into (9), we can easily accept that $(\hat{\underline{\gamma}}, \mathbf{\Gamma})$ will be incorporated twice, which is known as the problem of *double-counting*. For instance, (9b) becomes $\mathbf{C}_\mathrm{in} = \left((\mathbf{C}^\mathrm{I}_\mathsf{A})^{-1} + (\mathbf{C}^\mathrm{I}_\mathsf{B})^{-1} + 2\mathbf{\Gamma}^{-1}\right)$. Such a naive fusion result is then inconsistent, i.e. $\mathbf{C}_\mathrm{in} < \mathrm{E}[\tilde{\underline{\mathbf{x}}}_\mathrm{in}\tilde{\underline{\mathbf{x}}}^\mathrm{T}_\mathrm{in}]$.

The lower part in each equation of (8) unveils that both estimates can be fused by (9) as if they are independent, and the common part can then be subtracted to prevent double-counting. This technique is, for instance, employed by Grime and Durrant-Whyte (1994) for the channel filter, which separately keeps track of common information between sensor nodes.

3.3 Consistent and Tight Fusion

The optimal fusion result $(\hat{\underline{\mathbf{x}}}_\mathbf{\Gamma}, \mathbf{C}_\mathbf{\Gamma})$ in (8) can only be computed when the common information $(\hat{\underline{\gamma}}, \mathbf{\Gamma})$ is known and can be exploited to minimize the mean squared error $\mathrm{E}[\tilde{\underline{\mathbf{x}}}^\mathrm{T}_\mathbf{\Gamma}\tilde{\underline{\mathbf{x}}}_\mathbf{\Gamma}] = \mathrm{trace}(\mathbf{C}_\mathbf{\Gamma})$. In particular, decentralized estimation schemes may render it difficult or even impossible to keep track of common information. Therefore, the primary objective of our study is to derive a fusion method that yields consistent results without knowing $(\hat{\underline{\gamma}}, \mathbf{\Gamma})$. We additionally strive for fusion results that are not related to an unnecessarily large error covariance matrix and enclose the entirety of possible optimal fusion results as *tight* as possible.

Definition 3 (Tightness)**.** Let $\mathbf{\Lambda} \geq 0$ be an upper bound for every possible $\mathbf{C}_\mathbf{\Gamma}$, i.e., $\mathbf{C}_\mathbf{\Gamma} \leq \mathbf{\Lambda}$ for all admissible $\mathbf{\Gamma}$. A fusion result $(\hat{\underline{\mathbf{x}}}_\mathrm{fus}, \mathbf{C}_\mathrm{fus})$ is *tight* if the implication $\mathbf{C}_\mathbf{\Gamma} \leq \mathbf{\Lambda} \leq \mathbf{C}_\mathrm{fus} \Longrightarrow \mathbf{\Lambda} = \mathbf{C}_\mathrm{fus}$ holds for all admissible $\mathbf{\Gamma}$.

With this definition, we arrive at the problem statement:

Problem 4. Given consistent estimates $(\hat{\underline{\mathbf{x}}}_\mathsf{A}, \mathbf{C}_\mathsf{A})$ and $(\hat{\underline{\mathbf{x}}}_\mathsf{B}, \mathbf{C}_\mathsf{B})$ of $\underline{\mathbf{x}}$ according to (3) and (4), we seek to compute a *consistent* and *tight* fusion result $(\hat{\underline{\mathbf{x}}}_\mathrm{fus}, \mathbf{C}_\mathrm{fus})$ irrespective of the actual values of $\hat{\underline{\gamma}}$ and $\mathbf{\Gamma}$.

4 Review of Conservative Fusion Rules

For the purpose of solving Problem 4, conservative fusion methods have to be employed that provide consistent and tight results without having access to $(\hat{\underline{\gamma}}, \mathbf{\Gamma})$. In the following, we provide a review of covariance intersection and ellipsoidal intersection, which are both important concepts in decentralized data fusion. By identifying the missing pieces, a basis for the novel fusion method in Sec. 5 is established.

The most prominent option to achieve a consistent fusion result is the *covariance intersection* (CI) algorithm proposed by Julier and Uhlmann (1997). It uses the gains $\mathbf{K}_\mathrm{CI} = \omega\mathbf{C}_\mathrm{CI}\mathbf{C}^{-1}_\mathsf{A}$ and $\mathbf{L}_\mathrm{CI} = (1-\omega)\mathbf{C}_\mathrm{CI}\mathbf{C}^{-1}_\mathsf{B}$ to obtain

$$\hat{\underline{\mathbf{x}}}_\mathrm{CI} = \mathbf{K}_\mathrm{CI}\,\hat{\underline{\mathbf{x}}}_\mathsf{A} + \mathbf{L}_\mathrm{CI}\,\hat{\underline{\mathbf{x}}}_\mathsf{B} \quad (10a)$$

and reports the conservative covariance matrix

$$\mathbf{C}_\mathrm{CI} = \left(\omega\mathbf{C}^{-1}_\mathsf{A} + (1-\omega)\mathbf{C}^{-1}_\mathsf{B}\right)^{-1}\ , \quad (10b)$$

with $\mathbf{C}_\mathrm{CI} \geq \mathrm{E}[\tilde{\underline{\mathbf{x}}}_\mathrm{CI}\tilde{\underline{\mathbf{x}}}^\mathrm{T}_\mathrm{CI}]$. ω has to be chosen in the interval $[0, 1]$. CI was designed as a universal fusion rule and,

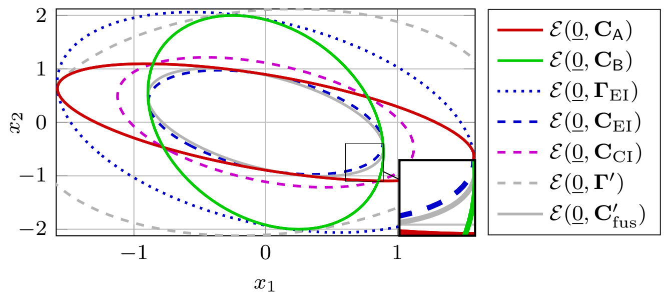


Fig. 1. A maximum inner ellipsoid $\mathcal{E}(\underline{0}, \mathbf{C}_\mathrm{EI})$ is obtained by modeling $\mathbf{\Gamma}_\mathrm{EI}$ as a tight upper bound $\mathcal{E}(\underline{0}, \mathbf{\Gamma}_\mathrm{EI})$ on $\mathcal{E}(\underline{0}, \mathbf{C}_\mathsf{A})$ and $\mathcal{E}(\underline{0}, \mathbf{C}_\mathsf{B})$. The common estimate related to $\mathcal{E}(\underline{0}, \mathbf{\Gamma}')$ and the corresponding fusion result $\mathcal{E}(\underline{0}, \mathbf{C}'_\mathrm{fus})$ represent the counterexample in Sec. 4.1.

as such, it is not restricted to a specific cross-covariance structure like (7). CI even constitutes a tight fusion rule as shown by Reinhardt et al. (2015) if arbitrary cross-covariance structures are possible and need to be considered. However, CI often—and in particular for the considered setup in Problem 4—proves to be too conservative. In this case, CI is not tight as illustrated in Sec. 4.1, which motivates the study of alternative concepts.

As an alternative to CI, *ellipsoidal intersection* (EI) has been derived in Sijs and Lazar (2012) from the observation that (4) implies that every admissible matrix $\mathbf{\Gamma}$ obeys the inequalities

$$\mathbf{C}_\mathsf{A} \leq \mathbf{\Gamma} \quad \text{and} \quad \mathbf{C}_\mathsf{B} \leq \mathbf{\Gamma} \, . \tag{11}$$

In order to account for unknown common information, EI computes a common estimate $(\hat{\underline{\gamma}}_\mathrm{EI}, \mathbf{\Gamma}_\mathrm{EI})$ that has a maximum possible $\mathbf{\Gamma}_\mathrm{EI}^{-1}$ and can be subtracted in (8). For the corresponding covariance ellipsoids, this implies that $\mathbf{\Gamma}_\mathrm{EI}$ is designed to be the shape matrix of the smallest ellipsoid $\mathcal{E}(\underline{0}, \mathbf{\Gamma}_\mathrm{EI})$ that encloses the covariance ellipsoids $\mathcal{E}(\underline{0}, \mathbf{C}_\mathsf{A})$ and $\mathcal{E}(\underline{0}, \mathbf{C}_\mathsf{B})$ related to $\mathbf{C}_\mathsf{A}$ and $\mathbf{C}_\mathsf{B}$. As elucidated in Sijs and Lazar (2012), the covering ellipsoid can be computed by means of a joint transformation

$$\mathbf{D}_\mathsf{A} = \mathbf{T}\mathbf{C}_\mathsf{A}\mathbf{T}^\mathrm{T} \quad \text{and} \quad \mathbf{D}_\mathsf{B} = \mathbf{T}\mathbf{C}_\mathsf{B}\mathbf{T}^\mathrm{T} \tag{12}$$

such that $\mathbf{D}_\mathsf{A}$ and $\mathbf{D}_\mathsf{B}$ are diagonal. The transformation $\mathbf{T}$ can be computed with the aid of an eigenvalue decomposition as in Sijs and Lazar (2012, see eq. (12) therein). The component-wise maximum $(\bar{\mathbf{D}})_{ii} = \max\{(\mathbf{D}_\mathsf{A})_{ii}, (\mathbf{D}_\mathsf{B})_{ii}\}$ then provides us with $\mathbf{\Gamma}_\mathrm{EI} := \mathbf{T}^{-1}\bar{\mathbf{D}}\mathbf{T}^{-\mathrm{T}}$. In compliance with (8b), this matrix is removed from the standard fusion result according to

$$\mathbf{C}_\mathrm{EI}^{-1} = \mathbf{C}_\mathsf{A}^{-1} + \mathbf{C}_\mathsf{B}^{-1} - \mathbf{\Gamma}_\mathrm{EI}^{-1} \, . \tag{13}$$

Fig. 1 illustrates the derivation of $\mathbf{\Gamma}_\mathrm{EI}$ as a covering ellipsoid, and the result (13) is far smaller than (10b). The computation of the common vector $\hat{\underline{\gamma}}_\mathrm{EI}$ and the fused estimate $\hat{\underline{\mathbf{x}}}_\mathrm{EI}$ is described in Sijs and Lazar (2012).

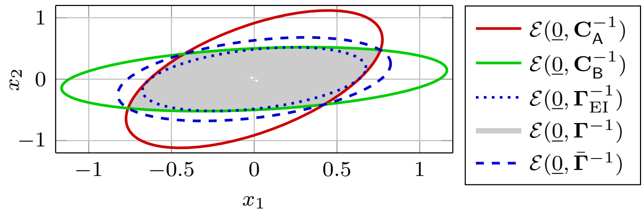


Fig. 2. Ellipsoids correspond to the inverse covariance matrices of Fig. 1. Different ellipsoids $\mathcal{E}(\underline{0}, \mathbf{\Gamma}^{-1})$ for admissible $\mathbf{\Gamma}^{-1}$ are shown. The ellipsoid $\mathcal{E}(\underline{0}, \bar{\mathbf{\Gamma}}^{-1})$ with $\bar{\mathbf{\Gamma}} = \omega\mathbf{C}_\mathsf{A} + (1-\omega)\mathbf{C}_\mathsf{B}, \omega = 0.5$ bounds the intersection.

4.1 *Comparison of* CI *and* EI

In order to compare CI and EI, it is worth pointing out that the problem statement in Sec. 3 itself constitutes a reason why the EI fusion method differs from CI. The derivations of EI exploit the decompositions (3) and (4), while CI is not tailored to a specific correlation structure.

As it has been observed by Julier (2008), CI prevents double-counting of the common estimate, and $(\hat{\underline{\gamma}}, \mathbf{\Gamma})$ is only incorporated once. By applying CI to the decompositions (3) and (4), the covariance matrix (10b) becomes

$$\mathbf{C}_\mathrm{CI} = \left(\omega(\mathbf{C}_\mathsf{A}^\mathrm{I})^{-1} + (1-\omega)(\mathbf{C}_\mathsf{B}^\mathrm{I})^{-1} + \mathbf{\Gamma}^{-1}\right)^{-1} \, .$$

By comparing (8b) with $\mathbf{C}_\mathrm{CI}$, we observe from

$$\mathbf{C}_\mathbf{\Gamma}^{-1} - \mathbf{C}_\mathrm{CI}^{-1} = (1-\omega)(\mathbf{C}_\mathsf{A}^\mathrm{I})^{-1} + \omega(\mathbf{C}_\mathsf{B}^\mathrm{I})^{-1} \geq \mathbf{0}$$

that the difference is, in general, strictly positive definite. For the estimates depicted in Fig. 1, any convex combination of $(\mathbf{C}_\mathsf{A}^\mathrm{I})^{-1}$ and $(\mathbf{C}_\mathsf{B}^\mathrm{I})^{-1}$ is positive definite, irrespective of $\mathbf{\Gamma}$ and $\omega \in [0,1]$. Hence, **CI is not tight**.

Fig. 1 already reveals an example where EI is inconsistent. The depicted ellipsoid $\mathcal{E}(\underline{0}, \mathbf{\Gamma}')$ also covers both covariance ellipsoids; hence, the matrix $\mathbf{\Gamma}'$ satisfies the inequalities (11) and is admissible. If the estimates share the common information $(\underline{0}, \mathbf{\Gamma}')$ and are fused by means of EI, the actual error covariance matrix of the fusion result becomes $\mathbf{C}'_\mathrm{fus} = \mathrm{E}\left[\tilde{\underline{\mathbf{x}}}_\mathrm{EI}\tilde{\underline{\mathbf{x}}}_\mathrm{EI}^\mathrm{T}\right]$ which is also shown in Fig. 1 in form of a covariance ellipsoid. The magnified section in the plot points out that EI underestimates this error matrix. Hence, **EI is not consistent**.

In consequence, the question of how an optimal solution to Problem 4 should look like remains to be discussed.

4.2 *The Missing Piece*

The preceding considerations have revealed that CI is too conservative and EI is not conservative enough—the solution obviously lies between those two. We reapproach the problem by considering the ellipsoids related to the inverse covariance matrices. Fig. 2 shows that $\mathbf{\Gamma}_{\mathrm{EI}}^{-1}$ refers to a maximum inner ellipsoid within the intersection $\mathcal{E}(\underline{0}, \mathbf{C}_{\mathsf{A}}^{-1}) \cap \mathcal{E}(\underline{0}, \mathbf{C}_{\mathsf{B}}^{-1})$. Furthermore, a number of 100 arbitrarily chosen admissible $\mathbf{\Gamma}^{-1}$ have been drawn in Fig. 2, which results in the shaded area. We can easily observe that the intersection tightly bounds the covariance ellipsoids of all admissible $\mathbf{\Gamma}^{-1}$. This leads to the conclusion that $\mathbf{\Gamma}_{\mathrm{EI}}^{-1}$ is chosen too small to account for all admissible $\mathbf{\Gamma}^{-1}$. Instead, the intersection of the ellipsoids related to the inverse covariance matrices $\mathbf{C}_{\mathsf{A}}^{-1}$ and $\mathbf{C}_{\mathsf{B}}^{-1}$ has to be considered.

5 Inverse Covariance Intersection

This section presents a novel fusion rule that solves Problem 4. It commences with a brief excursion to ellipsoidal calculus, which confirms our observation in the previous section with respect to the intersection of inverse covariance ellipsoids. Based on these considerations, consistency and tightness of the proposed concept are then proven. A comparison with CI unveils that the novel approach reports a lower mean squared error and, hence, is less conservative. An important result is that the method can be applied successively to multiple estimates without impairing consistency or tightness.

5.1 Derivation of the Fusion Rule

As expected and already indicated in our previous discussions, relationship (11) is strongly linked to the intersection of ellipsoids. In particular, the inclusion $\mathcal{E}(\underline{0}, \mathbf{\Omega}) \subseteq \mathcal{E}(\underline{0}, \mathbf{A}) \cap \mathcal{E}(\underline{0}, \mathbf{B})$ is equivalent to $\mathbf{\Omega} \leq \mathbf{A}$ and $\mathbf{\Omega} \leq \mathbf{B}$ for positive definite matrices $\mathbf{\Omega}$, $\mathbf{A}$, and $\mathbf{B}$. Additionally, for every $\underline{x} \in \mathcal{E}(\underline{0}, \mathbf{A}) \cap \mathcal{E}(\underline{0}, \mathbf{B})$, there exists an $\mathbf{\Omega} \leq \mathbf{A}, \mathbf{B}$ such that $\underline{x} \in \mathcal{E}(\underline{0}, \mathbf{\Omega})$—see Kahan (1968). Hence, the union of all ellipsoids $\mathcal{E}(\underline{0}, \mathbf{\Omega})$ with $\mathbf{\Omega} \leq \mathbf{A}, \mathbf{B}$ completely fills out the intersection $\mathcal{E}(\underline{0}, \mathbf{A}) \cap \mathcal{E}(\underline{0}, \mathbf{B})$.

The intersection of two ellipsoids generally does not yield an ellipsoid. As stated, for instance, by Kahan (1968), each ellipsoid $\mathcal{E}(\underline{0}, \mathbf{\Lambda})$ that tightly encloses the intersection, i.e.,

$$\mathcal{E}(\underline{0}, \mathbf{A}) \cap \mathcal{E}(\underline{0}, \mathbf{B}) \subseteq \mathcal{E}(\underline{0}, \mathbf{D}) \subseteq \mathcal{E}(\underline{0}, \mathbf{\Lambda}) \Longrightarrow \mathbf{D} = \mathbf{\Lambda} \ ,$$

has the parameterization $\mathbf{\Lambda}[\omega] = (\omega \mathbf{A}^{-1} + (1-\omega)\mathbf{B}^{-1})^{-1}$ with $\omega \in [0, 1]$. Together with $\mathbf{\Omega} \leq \mathbf{A}, \mathbf{B}$, it follows that the inequality

$$\mathbf{\Omega} \leq (\omega\, \mathbf{A}^{-1} + (1-\omega)\mathbf{B}^{-1})^{-1} \tag{14}$$

holds for all admissible $\mathbf{\Omega}$ and is tight.

By setting $\mathbf{A} := \mathbf{C}_{\mathsf{A}}^{-1}$, $\mathbf{B} := \mathbf{C}_{\mathsf{B}}^{-1}$, and $\mathbf{\Omega} := \mathbf{\Gamma}^{-1}$ and considering (14), the excursion to ellipsoidal calculus reveals that the right-hand side in

$$\mathbf{\Gamma}^{-1} \leq \left(\omega \mathbf{C}_{\mathsf{A}} + (1-\omega)\mathbf{C}_{\mathsf{B}}\right)^{-1} \tag{15}$$

is a tight outer bound for all admissible $\mathbf{\Gamma}^{-1}$ from Def. 2. In particular, this bound is related to the intersection of inverse covariance ellipsoids, which verifies the observation in Fig. 2.

Remark 5. Except for trivial cases $\mathcal{E}(\underline{0}, \mathbf{A}) \subseteq \mathcal{E}(\underline{0}, \mathbf{B})$, $\mathcal{E}(\underline{0}, \mathbf{B}) \subseteq \mathcal{E}(\underline{0}, \mathbf{A})$, or scalar $\mathbf{x}$, $\mathcal{E}(\underline{0}, \mathbf{\Lambda}[\omega])$ is a tight bound for each $\omega \in [0, 1]$. By considering a cost function J that satisfies $\mathbf{M} \leq \mathbf{N} \Longrightarrow J(\mathbf{M}) \leq J(\mathbf{N})$ for positive definite $\mathbf{M}$ and $\mathbf{N}$, the parameter $\omega^* = \arg\min J(\mathbf{\Lambda}[\omega])$ always provides a tight bound. As for the CI method, typical cost functions are the trace or determinant.

Lemma 6. *Every fusion result* $(\hat{\underline{\mathbf{x}}}_{\mathrm{fus}}, \mathbf{C}_{\mathrm{fus}})$ *that solves Problem 4 reports a tight covariance matrix in the form*

$$\mathbf{C}_{\mathrm{fus}}^{-1} = \mathbf{C}_{\mathsf{A}}^{-1} + \mathbf{C}_{\mathsf{B}}^{-1} - \left(\omega \mathbf{C}_{\mathsf{A}} + (1-\omega)\mathbf{C}_{\mathsf{B}}\right)^{-1} \ . \tag{16}$$

Proof. Since (15) is tight for any $\omega \in [0, 1]$, also

$$-\mathbf{\Gamma}^{-1} \geq -\left(\omega \mathbf{C}_{\mathsf{A}} + (1-\omega)\mathbf{C}_{\mathsf{B}}\right)^{-1}$$

holds and hence,

$$\mathbf{C}_{\mathsf{A}}^{-1} + \mathbf{C}_{\mathsf{B}}^{-1} - \mathbf{\Gamma}^{-1} \geq \mathbf{C}_{\mathsf{A}}^{-1} + \mathbf{C}_{\mathsf{B}}^{-1} - \left(\omega \mathbf{C}_{\mathsf{A}} + (1-\omega)\mathbf{C}_{\mathsf{B}}\right)^{-1}$$

is tight for all admissible $\mathbf{\Gamma}$. Consequently, this also holds for

$$\left(\mathbf{C}_{\mathsf{A}}^{-1} + \mathbf{C}_{\mathsf{B}}^{-1} - \mathbf{\Gamma}^{-1}\right)^{-1} \leq \mathbf{C}_{\mathrm{fus}} \ ,$$

where the left-hand side corresponds to the covariance matrix of the optimal fusion result (8) for known $\mathbf{\Gamma}$. □

Lemma 6 confirms the observation in Sec. 4.2. The following theorem presents a fusion method that actually attains (16) as a bound on the error covariance matrix and thereby constitutes a solution to Problem 4.

Theorem 7 (Inverse Covariance Intersection)**.** *Given Problem 4, a consistent combination of the estimates* $(\hat{\underline{\mathbf{x}}}_{\mathsf{A}}, \mathbf{C}_{\mathsf{A}})$ *and* $(\hat{\underline{\mathbf{x}}}_{\mathsf{B}}, \mathbf{C}_{\mathsf{B}})$ *is provided by* $(\hat{\underline{\mathbf{x}}}_{\mathrm{ICI}}, \mathbf{C}_{\mathrm{ICI}})$ *with*

$$\hat{\underline{\mathbf{x}}}_{\mathrm{ICI}} = \mathbf{K}_{\mathrm{ICI}}\, \hat{\underline{\mathbf{x}}}_{\mathsf{A}} + \mathbf{L}_{\mathrm{ICI}}\, \hat{\underline{\mathbf{x}}}_{\mathsf{B}} \tag{17a}$$

and

$$\mathbf{C}_{\mathrm{ICI}}^{-1} = \mathbf{C}_{\mathsf{A}}^{-1} + \mathbf{C}_{\mathsf{B}}^{-1} - \left(\omega \mathbf{C}_{\mathsf{A}} + (1-\omega)\mathbf{C}_{\mathsf{B}}\right)^{-1} \tag{17b}$$

for any $\omega \in [0, 1]$*. The gains in* (17a) *are given by*

$$\mathbf{K}_{\mathrm{ICI}} = \mathbf{C}_{\mathrm{ICI}} \cdot \left(\mathbf{C}_{\mathsf{A}}^{-1} - \omega(\omega \mathbf{C}_{\mathsf{A}} + (1-\omega)\mathbf{C}_{\mathsf{B}})^{-1}\right) \ , \tag{18a}$$

$$\mathbf{L}_{\mathrm{ICI}} = \mathbf{C}_{\mathrm{ICI}} \cdot \left(\mathbf{C}_{\mathsf{B}}^{-1} - (1-\omega)(\omega \mathbf{C}_{\mathsf{A}} + (1-\omega)\mathbf{C}_{\mathsf{B}})^{-1}\right) \ . \tag{18b}$$

Remark 8. The fusion result (17) is also tight if the parameter ω is determined according to Rem. 5. In particular, ω can be chosen to minimize an optimality criterion such as the trace or determinant of the bounding error covariance matrix (17b). Like CI, the proposed fusion rule offers a family of estimates that is parameterized by $\omega \in [0,1]$. A simple Matlab implementation can be downloaded from www.bennoack.net/ICI.

Proof of Theorem 7. We consider an arbitrary linear combination $\mathbf{K}\,\underline{\hat{\mathbf{x}}}_\mathsf{A} + \mathbf{L}\,\underline{\hat{\mathbf{x}}}_\mathsf{B}$ of the estimates (3) with gains $\mathbf{K}, \mathbf{L} \in \mathbb{R}^{n\times n}$. The gains that minimize the MSE after combining unbiased estimates have to fulfill $\mathbf{I} = \mathbf{K} + \mathbf{L}$, and the error covariance matrix becomes

$$\begin{aligned}
&\mathrm{E}\left[\left(\mathbf{K}\,\underline{\hat{\mathbf{x}}}_\mathsf{A} + \mathbf{L}\,\underline{\hat{\mathbf{x}}}_\mathsf{B} - \underline{\mathbf{x}}\right)\cdot\left(\mathbf{K}\,\underline{\hat{\mathbf{x}}}_\mathsf{A} + \mathbf{L}\,\underline{\hat{\mathbf{x}}}_\mathsf{B} - \underline{\mathbf{x}}\right)^\mathrm{T}\right] \\
&= \mathrm{E}\left[\left(\mathbf{K}\,\underline{\tilde{\mathbf{x}}}_\mathsf{A} + \mathbf{L}\,\underline{\tilde{\mathbf{x}}}_\mathsf{B}\right)\cdot\left(\mathbf{K}\,\underline{\tilde{\mathbf{x}}}_\mathsf{A} + \mathbf{L}\,\underline{\tilde{\mathbf{x}}}_\mathsf{B}\right)^\mathrm{T}\right] \\
&= \mathbf{K}\mathbf{C}_\mathsf{A}\mathbf{K}^\mathrm{T} + \mathbf{L}\mathbf{C}_\mathsf{B}\mathbf{L}^\mathrm{T} \\
&\quad + \mathbf{K}\mathbf{C}_\mathsf{A}\mathbf{\Gamma}^{-1}\mathbf{C}_\mathsf{B}\mathbf{L}^\mathrm{T} + \mathbf{L}\mathbf{C}_\mathsf{B}\mathbf{\Gamma}^{-1}\mathbf{C}_\mathsf{A}\mathbf{K}^\mathrm{T} \ ,
\end{aligned} \tag{19}$$

where the estimation errors $\underline{\tilde{\mathbf{x}}}_\mathsf{A}$ and $\underline{\tilde{\mathbf{x}}}_\mathsf{B}$ are given by (6), and the cross-covariance matrix in (7) has been used. We derive a bound on (19) with the help of the auxiliary random vector

$$\underline{\tilde{\mathbf{v}}} := \frac{1}{\sqrt{\lambda}}\mathbf{K}\,\mathbf{C}_\mathsf{A}\mathbf{C}_\mathsf{B}^{-1}\underline{\tilde{\mathbf{x}}}_\mathsf{B} - \sqrt{\lambda}\mathbf{L}\,\mathbf{C}_\mathsf{B}\mathbf{C}_\mathsf{A}^{-1}\underline{\tilde{\mathbf{x}}}_\mathsf{A}$$

for an arbitrary $\lambda > 0$. The corresponding error covariance matrix yields

$$\begin{aligned}
\mathrm{E}\left[\underline{\tilde{\mathbf{v}}}\underline{\tilde{\mathbf{v}}}^\mathrm{T}\right] = &\frac{1}{\lambda}\mathbf{K}\mathbf{C}_\mathsf{A}\mathbf{C}_\mathsf{B}^{-1}\mathbf{C}_\mathsf{A}\mathbf{K}^\mathrm{T} + \lambda\mathbf{L}\mathbf{C}_\mathsf{B}\mathbf{C}_\mathsf{A}^{-1}\mathbf{C}_\mathsf{B}\mathbf{L}^\mathrm{T} \\
&- \mathbf{K}\mathbf{C}_\mathsf{A}\mathbf{\Gamma}^{-1}\mathbf{C}_\mathsf{B}\mathbf{L}^\mathrm{T} - \mathbf{L}\mathbf{C}_\mathsf{B}\mathbf{\Gamma}^{-1}\mathbf{C}_\mathsf{A}\mathbf{K}^\mathrm{T} \ .
\end{aligned}$$

Since $\mathrm{E}\left[\underline{\tilde{\mathbf{v}}}\underline{\tilde{\mathbf{v}}}^\mathrm{T}\right] > \mathbf{0}$, we obtain the inequality

$$\begin{aligned}
&\frac{1}{\lambda}\mathbf{K}\mathbf{C}_\mathsf{A}\mathbf{C}_\mathsf{B}^{-1}\mathbf{C}_\mathsf{A}\mathbf{K}^\mathrm{T} + \lambda\mathbf{L}\mathbf{C}_\mathsf{B}\mathbf{C}_\mathsf{A}^{-1}\mathbf{C}_\mathsf{B}\mathbf{L}^\mathrm{T} \\
&\geq \mathbf{K}\mathbf{C}_\mathsf{A}\mathbf{\Gamma}^{-1}\mathbf{C}_\mathsf{B}\mathbf{L}^\mathrm{T} + \mathbf{L}\mathbf{C}_\mathsf{B}\mathbf{\Gamma}^{-1}\mathbf{C}_\mathsf{A}\mathbf{K}^\mathrm{T} \ .
\end{aligned}$$

This inequality can now be utilized to derive a bound on (19), i.e.,

$$\begin{aligned}
&\mathrm{E}\left[\left(\mathbf{K}\,\underline{\tilde{\mathbf{x}}}_\mathsf{A} + \mathbf{L}\,\underline{\tilde{\mathbf{x}}}_\mathsf{B}\right)\cdot\left(\mathbf{K}\,\underline{\tilde{\mathbf{x}}}_\mathsf{A} + \mathbf{L}\,\underline{\tilde{\mathbf{x}}}_\mathsf{B}\right)^\mathrm{T}\right] \\
&\leq \mathbf{K}\mathbf{C}_\mathsf{A}\mathbf{K}^\mathrm{T} + \mathbf{L}\mathbf{C}_\mathsf{B}\mathbf{L}^\mathrm{T} \\
&\quad + \frac{1}{\lambda}\mathbf{K}\mathbf{C}_\mathsf{A}\mathbf{C}_\mathsf{B}^{-1}\mathbf{C}_\mathsf{A}\mathbf{K}^\mathrm{T} + \lambda\mathbf{L}\mathbf{C}_\mathsf{B}\mathbf{C}_\mathsf{A}^{-1}\mathbf{C}_\mathsf{B}\mathbf{L}^\mathrm{T} \\
&= \mathbf{K}\big(\mathbf{C}_\mathsf{A} + \frac{1}{\lambda}\mathbf{C}_\mathsf{A}\mathbf{C}_\mathsf{B}^{-1}\mathbf{C}_\mathsf{A}\big)\mathbf{K}^\mathrm{T} \\
&\quad + \mathbf{L}\big(\mathbf{C}_\mathsf{B} + \lambda\mathbf{C}_\mathsf{B}\mathbf{C}_\mathsf{A}^{-1}\mathbf{C}_\mathsf{B}\big)\mathbf{L}^\mathrm{T}
\end{aligned}$$

which holds for every possible $(\underline{\hat{\boldsymbol{\gamma}}}, \mathbf{\Gamma})$ and each choice of $\mathbf{K}$ and $\mathbf{L}$. Since the bound is independent of $\mathbf{\Gamma}$, the trace of the bound becomes minimal when the gains are chosen as in (9). This leads us to the fused covariance matrix

$$\begin{aligned}
\mathbf{C}_\text{fus}^{-1} &\overset{(9b)}{=} \big(\mathbf{C}_\mathsf{A} + \frac{1}{\lambda}\mathbf{C}_\mathsf{A}\mathbf{C}_\mathsf{B}^{-1}\mathbf{C}_\mathsf{A}\big)^{-1} + \big(\mathbf{C}_\mathsf{B} + \lambda\mathbf{C}_\mathsf{B}\mathbf{C}_\mathsf{A}^{-1}\mathbf{C}_\mathsf{B}\big)^{-1} \\
&= \mathbf{C}_\mathsf{A}^{-1} - \big(\lambda\mathbf{C}_\mathsf{B} + \mathbf{C}_\mathsf{A}\big)^{-1} + \mathbf{C}_\mathsf{B}^{-1} - \big(\frac{1}{\lambda}\mathbf{C}_\mathsf{A} + \mathbf{C}_\mathsf{B}\big)^{-1} \\
&= \mathbf{C}_\mathsf{A}^{-1} + \mathbf{C}_\mathsf{B}^{-1} - (1+\lambda)\big(\lambda\mathbf{C}_\mathsf{B} + \mathbf{C}_\mathsf{A}\big)^{-1} \\
&= \mathbf{C}_\mathsf{A}^{-1} + \mathbf{C}_\mathsf{B}^{-1} - \Big(\frac{\lambda}{(1+\lambda)}\mathbf{C}_\mathsf{B} + \frac{1}{(1+\lambda)}\mathbf{C}_\mathsf{A}\Big)^{-1} \ ,
\end{aligned}$$

where the Woodbury identity has been applied. We set $\omega = \frac{1}{(1+\lambda)}$ and arrive at

$$\mathbf{C}_\text{fus}^{-1} = \mathbf{C}_\mathsf{A}^{-1} + \mathbf{C}_\mathsf{B}^{-1} - \big(\omega\mathbf{C}_\mathsf{A} + (1-\omega)\mathbf{C}_\mathsf{B}\big)^{-1} \ ,$$

which proves (17b). The formulas (18) for the gains $\mathbf{K}_\text{ICI}$ and $\mathbf{L}_\text{ICI}$ follow from (9). □

5.2 Comparison with Covariance Intersection

The ICI fusion method only renders a viable alternative to CI if it provides us with less conservative fusion results. In essence, we have to compare the covariance matrices (10b) and (17b), both of which are dependent upon a scalar parameter. The result of the comparison is summarized in the following lemma.

Lemma 9. *For each $\omega^* \in [0,1]$, there is a parameter $\omega^+ \in [0,1]$ with $\mathbf{C}_\text{ICI}[\omega^+] \leq \mathbf{C}_\text{CI}[\omega^*]$.*

Proof. By setting $\omega^+ = 1 - \omega^*$, we obtain

$$\begin{aligned}
&\mathbf{C}_\text{ICI}^{-1}[1-\omega^*] - \mathbf{C}_\text{CI}^{-1}[\omega^*] \\
&= \mathbf{C}_\mathsf{A}^{-1} + \mathbf{C}_\mathsf{B}^{-1} - \big((1-\omega^*)\mathbf{C}_\mathsf{A} + \omega^*\mathbf{C}_\mathsf{B}\big)^{-1} \\
&\quad - \big(\omega^*\mathbf{C}_\mathsf{A}^{-1} + (1-\omega^*)\mathbf{C}_\mathsf{B}^{-1}\big) \\
&= (1-\omega^*)\mathbf{C}_\mathsf{A}^{-1} + \omega^*\mathbf{C}_\mathsf{B}^{-1} - \big((1-\omega^*)\mathbf{C}_\mathsf{A} + \omega^*\mathbf{C}_\mathsf{B}\big)^{-1},
\end{aligned} \tag{20}$$

The joint transformation (12) applied to (20) yields

$$\begin{aligned}
\bar{\mathbf{D}} &:= \mathbf{T}\big(\mathbf{C}_\text{ICI}^{-1}[1-\omega^*] - \mathbf{C}_\text{CI}^{-1}[\omega^*]\big)\mathbf{T}^\mathrm{T} \\
&= (1-\omega^*)\mathbf{D}_\mathsf{A}^{-1} + \omega^*\mathbf{D}_\mathsf{B}^{-1} - \big((1-\omega^*)\mathbf{D}_\mathsf{A} + \omega^*\mathbf{D}_\mathsf{B}\big)^{-1} \ .
\end{aligned}$$

With $d_\mathsf{A}^i := (\mathbf{D}_\mathsf{A})_{ii}$ and $d_\mathsf{B}^i := (\mathbf{D}_\mathsf{B})_{ii}$, the diagonal entries are

$$(\bar{\mathbf{D}})_{ii} = (1-\omega^*)\frac{1}{d_\mathsf{A}^i} + \omega^*\frac{1}{d_\mathsf{B}^i} - \frac{1}{(1-\omega^*)d_\mathsf{A}^i + \omega^* d_\mathsf{B}^i} \geq 0 \ .$$

By multiplying with $(1-\omega^*)d_\mathsf{A}^i + \omega^* d_\mathsf{B}^i$, which is a positive value, we obtain

$$(1-\omega^*)^2 + (\omega^*)^2 + \omega^*(1-\omega^*)\Big(\frac{d_\mathsf{B}^i}{d_\mathsf{A}^i} + \frac{d_\mathsf{A}^i}{d_\mathsf{B}^i}\Big) - 1 \geq 0 \ .$$

From the inequality $(a^2+b^2) \geq 2|ab|$, it follows

$$\left(\frac{d^i_{\mathsf{B}}}{d^i_{\mathsf{A}}} + \frac{d^i_{\mathsf{A}}}{d^i_{\mathsf{B}}}\right) = \frac{(d^i_{\mathsf{A}})^2 + (d^i_{\mathsf{B}})^2}{d^i_{\mathsf{A}} \cdot d^i_{\mathsf{B}}} \geq 2$$

and hence

$$\begin{aligned}(1-\omega^*)^2 + (\omega^*)^2 + \omega^*(1-\omega^*)\left(\frac{d^i_{\mathsf{B}}}{d^i_{\mathsf{A}}} + \frac{d^i_{\mathsf{A}}}{d^i_{\mathsf{B}}}\right) - 1 \\ \geq (1-\omega^*)^2 + (\omega^*)^2 + 2\omega^*(1-\omega^*) - 1 = 0 \ .\end{aligned}$$

The last inequality finally implies that each diagonal component $(\underline{\mathbf{D}})_{ii}$ is positive, and therefore also the difference (20) is positive definite. This leads to the inequality $\mathbf{C}^{-1}_{\mathrm{ICI}}[1-\omega^*] \geq \mathbf{C}^{-1}_{\mathrm{CI}}[\omega^*]$ and, in particular, to $\mathbf{C}_{\mathrm{ICI}}[1-\omega^*] \leq \mathbf{C}_{\mathrm{CI}}[\omega^*]$. □

Consequently, the ICI approach provides more accurate fusion results than CI. In particular, for any ω_{CI} that minimizes the trace or determinant of $\mathbf{C}_{\mathrm{CI}}[\omega_{\mathrm{CI}}]$, the matrix $\mathbf{C}_{\mathrm{ICI}}[1-\omega_{\mathrm{CI}}]$ is attributed to an even smaller trace or determinant, respectively. However, the parameter ω_{ICI} that minimizes the same criterion for $\mathbf{C}_{\mathrm{ICI}}[\omega_{\mathrm{ICI}}]$ can be different to $1-\omega_{\mathrm{CI}}$. In this case, the difference $\mathbf{C}_{\mathrm{CI}}[\omega_{\mathrm{CI}}] - \mathbf{C}_{\mathrm{ICI}}[\omega_{\mathrm{ICI}}]$ may be indefinite.

5.3 Transitivity of Common Information

From an application-oriented point of view, a fusion method is supposed to be applied iteratively to multiple estimates without impairing consistency. Given the situation that the estimates $(\hat{\underline{\mathbf{x}}}_{\mathsf{A}}, \mathbf{C}_{\mathsf{A}})$ and $(\hat{\underline{\mathbf{x}}}_{\mathsf{B}}, \mathbf{C}_{\mathsf{B}})$ both share the common estimate $(\hat{\underline{\gamma}}, \mathbf{\Gamma})$, we study the question of how the common information is affected after applying ICI to $(\hat{\underline{\mathbf{x}}}_{\mathsf{A}}, \mathbf{C}_{\mathsf{A}})$ and $(\hat{\underline{\mathbf{x}}}_{\mathsf{B}}, \mathbf{C}_{\mathsf{B}})$. With the gains (18) and the decompositions (3), the common estimate is transformed according to

$$\begin{aligned}\hat{\underline{\mathbf{x}}}_{\mathrm{ICI}} &\overset{(17a)}{=} \mathbf{K}_{\mathrm{ICI}}\hat{\underline{\mathbf{x}}}_{\mathsf{A}} + \mathbf{L}_{\mathrm{ICI}}\hat{\underline{\mathbf{x}}}_{\mathsf{B}} \\ &= \mathbf{K}_{\mathrm{ICI}}\mathbf{C}_{\mathsf{A}}(\mathbf{C}^{\mathrm{I}}_{\mathsf{A}})^{-1}\hat{\underline{\mathbf{x}}}^{\mathrm{I}}_{\mathsf{A}} + \mathbf{L}_{\mathrm{ICI}}\mathbf{C}_{\mathsf{B}}(\mathbf{C}^{\mathrm{I}}_{\mathsf{B}})^{-1}\hat{\underline{\mathbf{x}}}^{\mathrm{I}}_{\mathsf{B}} \\ &\quad + \mathbf{K}_{\mathrm{ICI}}\mathbf{C}_{\mathsf{A}}\mathbf{\Gamma}^{-1}\hat{\underline{\gamma}} + \mathbf{L}_{\mathrm{ICI}}\mathbf{C}_{\mathsf{B}}\mathbf{\Gamma}^{-1}\hat{\underline{\gamma}}\end{aligned} \tag{21}$$

The last sum can be rewritten as

$$(\mathbf{K}_{\mathrm{ICI}}\mathbf{C}_{\mathsf{A}} + \mathbf{L}_{\mathrm{ICI}}\mathbf{C}_{\mathsf{B}})\,\mathbf{\Gamma}^{-1}\hat{\underline{\gamma}} = \mathbf{C}_{\mathrm{ICI}}\left(\bar{\mathbf{K}}\mathbf{C}_{\mathsf{A}} + \bar{\mathbf{L}}\mathbf{C}_{\mathsf{B}}\right)\mathbf{\Gamma}^{-1}\hat{\underline{\gamma}} \tag{22}$$

with

$$\begin{aligned}\bar{\mathbf{K}} &:= \mathbf{C}^{-1}_{\mathsf{A}} - \omega(\omega\mathbf{C}_{\mathsf{A}} + (1-\omega)\mathbf{C}_{\mathsf{B}})^{-1} \ , \\ \bar{\mathbf{L}} &:= \mathbf{C}^{-1}_{\mathsf{B}} - (1-\omega)(\omega\mathbf{C}_{\mathsf{A}} + (1-\omega)\mathbf{C}_{\mathsf{B}})^{-1} \ .\end{aligned}$$

With the identity

$$\begin{aligned}\bar{\mathbf{K}}\mathbf{C}_{\mathsf{A}} + \bar{\mathbf{L}}\mathbf{C}_{\mathsf{B}} &= 2\cdot\mathbf{I} - \omega(\omega\mathbf{C}_{\mathsf{A}} + (1-\omega)\mathbf{C}_{\mathsf{B}})^{-1}\mathbf{C}_{\mathsf{A}} \\ &\quad - (1-\omega)(\omega\mathbf{C}_{\mathsf{A}} + (1-\omega)\mathbf{C}_{\mathsf{B}})^{-1}\mathbf{C}_{\mathsf{B}} \\ &= \mathbf{I} \ ,\end{aligned}$$

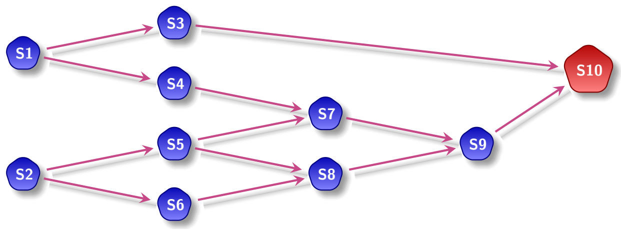


Fig. 3. Communication scheme.

the vector (22) reduces to $\mathbf{C}_{\mathrm{ICI}}\mathbf{\Gamma}^{-1}\hat{\underline{\gamma}}$. For this reason, the fused estimate (21) has the decomposition

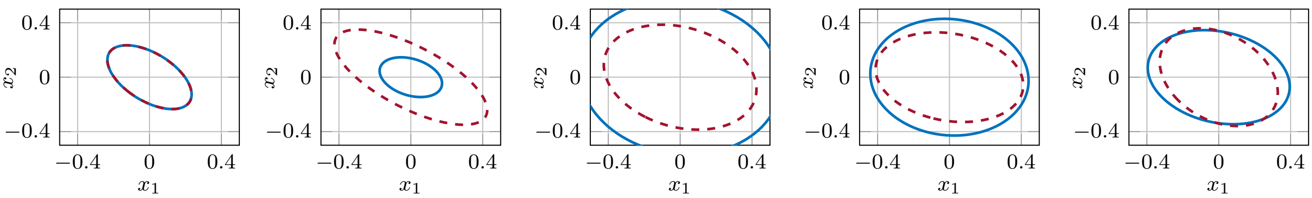
$$\hat{\underline{\mathbf{x}}}_{\mathrm{ICI}} = \mathbf{C}_{\mathrm{ICI}}\left((\mathbf{C}^{\mathrm{I}}_{\mathrm{ICI}})^{-1}\hat{\underline{\mathbf{x}}}^{\mathrm{I}}_{\mathrm{ICI}} + \mathbf{\Gamma}^{-1}\hat{\underline{\gamma}}\right) \ ,$$

which resembles (3). Analogously, the covariance matrix (17b) can be decomposed in the same fashion. The independent parts in (21) have been summarized in $(\hat{\underline{\mathbf{x}}}^{\mathrm{I}}_{\mathrm{ICI}}, \mathbf{C}^{\mathrm{I}}_{\mathrm{ICI}})$. In conclusion, the ICI fusion method preserves the common estimate and the problem structure in Sec. 3.

6 Discussion and Example

The full potential of the ICI fusion methodology can be exploited in sensor networks where double-counting of data poses a severe problem. The communication paths shown in Fig. 3 display several examples of critical nodes that have to take care of common information. As such, node S8 obtains estimates from S5 and S6 that share the information received from S2. Of course, in order to prevent double-counting, simple strategies can be pursued. For instance, each node can forward the received data to the subsequent node on the path instead of computing a fusion result, which may lead to an unacceptably high data volume. Another possibility is to keep track of common information, but this requires bookkeeping and cannot account for changing network topologies. Therefore, an on-site fusion of received data renders the most flexible solution as neither a specific topology nor any additional bookkeeping is required. Concepts like split covariance intersection that rely on an explicit separation of dependent and independent parts are difficult to employ as independent parts cannot be identified.

In order to compare the performance of the different fusion methods, Fig. 4 shows the result of a Monte-Carlo simulation over 50 000 runs, where the communication path in Fig. 3 is used. In each run, each node is initialized with $(\hat{\underline{\mathbf{x}}}_{\mathsf{SX}}, \mathbf{C}_{\mathsf{SX}}) = \left(\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right], \left[\begin{smallmatrix}2&0\\0&2\end{smallmatrix}\right]\right)$, which also corresponds to the uncertainty of the true state. Each node has a local Kalman filter and measures the state with measurement matrix $\mathbf{H}_{\mathsf{SX}} = [\sin(\frac{\pi}{2}\cdot\frac{\mathsf{X}}{10}), \cos(\frac{\pi}{2}\cdot\frac{\mathsf{X}}{10})]^{\mathrm{T}}$ and zero-mean noise $\mathbf{v}_{\mathsf{SX}}$ with variance $\mathbf{C}_v = 0.2$. Each initial estimate is locally updated with a measurement, before it is sent along the path shown in Fig. 3 and fused with other estimates. For node S10, Fig. 4 compares the covariance



(a) Opt. fusion. (b) Naive fusion. (c) CI fusion. (d) ICI fusion. (e) EI fusion.

Fig. 4. Fusion results at node S10 in Fig. 3. Comparison of covariance matrices $\mathbf{C}$ (—) reported by fusion methods and actual error covariance matrices $\widetilde{\mathbf{C}}$ (- -) depicted as ellipsoids. $\widetilde{\mathbf{C}}$ is computed as sample covariance matrix of the estimation errors $\underline{\tilde{\mathbf{x}}}_{\mathsf{S10}} = \underline{\hat{\mathbf{x}}}_{\mathsf{S10}} - \underline{\mathbf{x}}$ after 50 000 runs, where $\underline{\hat{\mathbf{x}}}_{\mathsf{S10}}$ has been computed by the different fusion methods. A fusion method is consistent if the actual error is bounded by the reported error ellipsoid, i.e., $\mathbf{C} \geq \widetilde{\mathbf{C}}$.

matrices $\mathbf{C}$ that are reported by the fusion methods with the actual error matrices $\widetilde{\mathbf{C}} = \mathrm{E}[(\underline{\hat{\mathbf{x}}} - \underline{\mathbf{x}})(\underline{\hat{\mathbf{x}}} - \underline{\mathbf{x}})^{\mathrm{T}}]$. The latter matrices have been computed based on the Monte-Carlo simulation. As expected, for known $(\hat{\underline{\gamma}}, \mathbf{\Gamma})$ shared by the estimates, $\mathbf{C}_{\mathbf{\Gamma}}$ in (8) equals $\widetilde{\mathbf{C}}_{\mathbf{\Gamma}}$. With naive fusion, $\mathbf{C}_{\mathrm{in}}$ severely underestimates the actual $\widetilde{\mathbf{C}}_{\mathrm{in}}$ because of double-counting. Both CI and ICI report consistent estimates while ICI is far tighter, and also the actual error of ICI is smaller. EI is not consistent as the difference $\mathbf{C}_{\mathrm{EI}} - \widetilde{\mathbf{C}}_{\mathrm{EI}}$ is not positive semi-definite.

The example underpins the optimality of ICI in problem setups, where information is double counted and which lie in the scope of Problem 4. ICI provides results with proven consistency when cycles in the network cause dependencies and, in general, other sources of dependencies are absent or negligible. Further studies will particularly focus on the effect of common process noise, which may lead to a parameterization different from (7).

7 Conclusions

A novel concept for decentralized data fusion has been derived by studying the problem of combining estimates that share unknown common information. It employs a decomposition of each estimate into a common and an independent part, which has first been introduced for the ellipsoidal intersection method. For the considered fusion problem, covariance intersection and ellipsoidal intersection are possible candidates. However, a detailed analysis of both concepts has revealed that covariance intersection is too conservative and ellipsoidal intersection may lead to inconsistent fusion results. By identifying the missing pieces, a conservative and tight bound on unknown common information has been derived. This bound is closely related to the intersection of inverse covariance ellipsoids.

The main contribution of this work is the inverse covariance intersection method that represents an optimal—consistent and tight—solution to the problem of fusing estimates that share unknown common information. Important results are that inverse covariance intersection is less conservative than covariance intersection and can recursively be applied to fuse estimates.

Acknowledgements

This work was supported by the German Research Foundation (DFG) under grant NO 1133/1-1.

References

Bar-Shalom, Y. and Li, X.R. (1995). *Multitarget Multisensor Tracking: Principles and Techniques*. YBS Publishing.

Battistelli, G., Chisci, L., Mugnai, G., Farina, A., and Graziano, A. (2015). Consensus-Based Linear and Nonlinear Filtering. *IEEE Transactions on Automatic Control*, 60(5), 1410–1415.

Cattivelli, F.S. and Sayed, A.H. (2010). Diffusion Strategies for Distributed Kalman Filtering and Smoothing. *IEEE Transactions on Automatic Control*, 55(9), 2069–2084.

Chong, C.Y., Mori, S., Govaers, F., and Koch, W. (2014). Comparison of Tracklet Fusion and Distributed Kalman Filter for Track Fusion. In *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*. Salamanca, Spain.

Deng, Z., Zhang, P., Qi, W., Liu, J., and Gao, Y. (2012). Sequential Covariance Intersection Fusion Kalman Filter. *Information Sciences*, 189(15), 293–309.

Grime, S. and Durrant-Whyte, H.F. (1994). Data Fusion in Decentralized Sensor Networks. *Control Engineering Practice*, 2(5), 849–863.

Hall, D.L., Chong, C.Y., Llinas, J., and Liggins II, M.E. (eds.) (2013). *Distributed Data Fusion for Network-Centric Operations*. The Electrical Engineering and Applied Signal Processing Series. CRC Press.

Hu, J., Xie, L., and Zhang, C. (2012). Diffusion Kalman Filtering Based on Covariance Intersection. *IEEE Transactions on Signal Processing*, 60(2), 891–902.

Julier, S.J. (2008). Fusion without Independence (keynote abstract). In *Proceedings of the IET Seminar on Tracking and Data Fusion: Algorithms and Applications*, 1–5. Birmingham, United Kingdom.

Julier, S.J. and Uhlmann, J.K. (1997). A Non-divergent Estimation Algorithm in the Presence of Unknown Correlations. In *Proceedings of the IEEE American Control Conference (ACC 1997)*, volume 4, 2369–2373. Albuquerque, New Mexico, USA.

Kahan, W. (1968). Circumscribing an Ellipsoid about the Intersection of Two Ellipsoids. *Canadian Mathematical Bulletin*, 11(3), 437 – 441.

Reinhardt, M., Noack, B., Arambel, P.O., and Hanebeck, U.D. (2015). Minimum Covariance Bounds for the Fusion under Unknown Correlations. *IEEE Signal Processing Letters*, 22(9), 1210–1214.

Sijs, J. and Lazar, M. (2012). State-fusion with Unknown Correlation: Ellipsoidal Intersection. *Automatica*, 48(8), 1874–1878.