Learning and Exploiting Partial Knowledge in Distributed Estimation

Susanne Radtke, Jiří Ajgl, Ondřej Straka, and Uwe D. Hanebeck

***Abstract*— In distributed estimation, several sensor nodes provide estimates of the same underlying dynamic process. These estimates are correlated but due to local processing, the correlations are only partially known or even unknown. For a consistent fusion of the local estimates, the correlation needs to be properly treated. Many methods provide consistent but overly conservative fusion results. In this paper, we propose to learn partial knowledge about the correlation in the form of correlation sets and exploit this knowledge to provide less conservative estimates. We use a simple numerical example to demonstrate the advantages of the proposed approach in terms of quality and consistency and how the quality of the fused estimate increases with time.**

I. Introduction

Considered Problem: Distributed sensor networks can cooperatively perform various tasks [1]. While centralized processing of measurements can be done optimally, the local processing into state estimates has proven to be more robust, flexible, and scalable [2]. However, the fusion of state estimates suffers from the track-to-track correlation problem [3] that needs to be addressed to ensure consistent fusion results. This paper considers the problem of fusing the state estimates from two discrete-time linear Gaussian systems with two completely synchronized sensors with linear Gaussian observations.

State-of-the-Art: Over the last four decades, several methods for distributed estimation have been proposed [4], such as the information form of them Kalman filter [5], [6], [7]. When fusing state estimates, correlations between the local tracks have to be accounted for [8] to ensure consistency of the fused estimate. Furthermore, [9] pointed out that the fusion configuration plays a vital role in the quality of the fused estimate. Covariance intersection [10], [11], [12] is commonly used in distributed estimation tasks because it ensures consistent fusion results in all circumstances. The relationship between this upper bound and the Minkowski sums of ellipsoids is discussed in [13]. Since covariance intersection tends to produce estimates that are often too conservative, other approaches propose closer bounds [14], [15], [16], to exploit implicit knowledge about partial knowledge [17], or to exploit known independent knowledge [18]. Inverse covariance intersection [19], [20] is another method that aims to exploit partial knowledge to achieve tighter bounds. Recently, several methods to reconstruct cross-covariances have been proposed that include methods using an ensemble [21], samples [22], [23], [24] or square-root decompositions [25], [26]. However, these methods suffer from increased computation and communication requirements, which is prohibitive in many distributed estimation tasks. Nonetheless, reconstruction methods can find patterns in repetitive fusion tasks and help learn correlations between state estimates over time.

Contribution: This paper explores the learning of partial knowledge about correlated estimation errors of state estimates in distributed estimation. We focus on learning the bounds of the sets of correlations and using this partial information to improve the quality of the fused estimates. We introduce an analytic and a simulation approach to learn the correlation coefficients suitable for different types of uncertainty about the correlation. Furthermore, we discuss bounding techniques based on [14], [15] that can provide tight bounds for the estimation errors. See also [27], where similar approaches for the exploitation of partial knowledge are proposed. We focus on the fusion of two state estimates and a linear system description. The numerical evaluation example is chosen to allow for an intuitive understanding of the learned correlation coefficients.

Outline: Sec. II formulates the problem of fusing two state estimates with unknown correlation. In Sec. III, we consider different methods to learn correlated estimation errors between distributed estimators and possible ways to exploit this knowledge to improve the fusion result. We highlight the proposed approach by using a simple numerical experiment in Sec. IV. Finally, our findings are summarized in Sec. V.

II. Problem Formulation

Consider the discrete-time linear time-variant stochastic dynamic system with time index k

$$\underline{x}_{k+1} = \mathbf{A}_k \underline{x}_k + \underline{w}_k\,, \qquad \underline{w}_k \sim \mathcal{N}(\underline{0}, \mathbf{Q}_k)\,, \tag{1}$$

$$\underline{z}_k^i = \mathbf{H}_k^i \underline{x}_k + \underline{v}_k^i\,, \qquad \underline{v}_k^i \sim \mathcal{N}(\underline{0}, \mathbf{R}_k^i)\,, \tag{2}$$

where $\underline{x}_k$ is the system state, $\underline{w}_k$ is the Gaussian zero-mean process noise, $\underline{z}_k^i$ is the measurement from the i-th sensor node, $\underline{v}_k^i$ is the Gaussian zero-mean measurement noise. $\mathbf{A}_k$ and $\mathbf{H}_k^i$ are matrices of corresponding dimensions and the initial condition is $\underline{x}_0 \sim \mathcal{N}(\hat{\underline{x}}_{0|-1}, \mathbf{P}_{0|-1})$. The covariance

This work was supported by the German Research Foundation (DFG) under grant HA 3789/20-1 and by the Czech Science Foundation under grant GA20-06054J.

Susanne Radtke and Uwe D. Hanebeck are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, susanne.radtke@kit.edu, uwe.hanebeck@ieee.org.

Jiří Ajgl and Ondřej Straka are with the European Centre of Excellence – New Technologies for Information Society, Department of Cybernetics, Faculty of Applied Sciences, University of West Bohemia, Pilsen, Czech Republic, jiriajgl@kky.zcu.cz, straka30@kky.zcu.cz.

matrices $\mathbf{Q}_k$ and $\mathbf{R}_k^i$ are assumed to be known and positive definite. The noises are white, mutually independent, and independent of the initial condition. The goal of state estimation is to infer the state $\underline{x}_k$ given all the information up to time instant k consisting of the available measurements.

A. Estimation in Local Sensor Nodes

In distributed estimation, each sensor node processes its own measurements to obtain the local state estimate $\underline{\hat{x}}_{k|k}^i$ and corresponding error covariance matrix $\mathbf{P}_{k|k}^i$. The estimates are communicated between nodes or sent to the fusion center, where the estimates are fused to obtain better estimates.

For the linear system (1) and (2), the estimate $\underline{\hat{x}}_{k|k}^i$ is computed by the Kalman filter by performing a prediction step

$$\begin{aligned}\underline{\hat{x}}_{k|k-1}^i &= \mathbf{A}_k \underline{\hat{x}}_{k-1|k-1}^i\,,\\ \mathbf{P}_{k|k-1}^i &= \mathbf{A}_k \mathbf{P}_{k-1|k-1}^i (\mathbf{A}_k)^\top + \mathbf{Q}_k\end{aligned}$$

and incorporating a local measurement

$$\begin{aligned}\underline{\hat{x}}_{k|k}^i &= \underline{\hat{x}}_{k|k-1}^i + \mathbf{K}_k^i(\underline{z}_k^i - \underline{\hat{x}}_{k|k-1}^i),\\ \mathbf{P}_{k|k}^i &= (\mathbf{I}-\mathbf{K}_k^i\mathbf{H}_k^i)\mathbf{P}_{k|k-1}^i(\mathbf{I}-\mathbf{K}_k^i\mathbf{H}_k^i)^\top + \mathbf{K}_k^i\mathbf{R}_k^i(\mathbf{K}_k^i)^\top,\end{aligned}$$

where the Kalman gain $\mathbf{K}_k^i$ is given by

$$\mathbf{K}_k^i = \mathbf{P}_{k|k-1}^i\mathbf{H}_k^i\left(\mathbf{H}_k^i\mathbf{P}_{k|k-1}^i(\mathbf{H}_k^i)^\top + \mathbf{R}_k^i\right)^{-1}.$$

This local estimate can be improved by fusing it with the estimate of another sensor node.

B. Fusion of State Estimates

Once a sensor node receives the estimate and the corresponding covariance matrix of another sensor node, they are fused with the local estimate to increase its quality. The two estimates $\underline{\hat{x}}^i$ and $\underline{\hat{x}}^j$ have joint error covariance

$$\mathbf{J} = \begin{bmatrix}\mathbf{P}^i & \mathbf{P}^{ij}\\ \mathbf{P}^{ji} & \mathbf{P}^j\end{bmatrix},$$

where the cross-covariance $\mathbf{P}^{ij} = (\mathbf{P}^{ji})^\top$ is due to common prior information, common process noise, and double counting of measurement information. Note that the time indices were dropped for the convenience purposes. The estimates are fused using the linear combination

$$\underline{\hat{x}}^{\mathrm{f}} = \mathbf{F}^i\underline{\hat{x}}^i + \mathbf{F}^j\underline{\hat{x}}^j \quad \text{with} \quad \mathbf{F}^i + \mathbf{F}^j = \mathbf{I},$$

where $\mathbf{F}^i$ and $\mathbf{F}^j$ are fusion gain matrices and $\mathbf{I}$ is the identity matrix.

The optimal fusion gain can be calculated according to the Bar-Shalom/Campo (BSC) formula [3]

$$\mathbf{F}^j = (\mathbf{P}^i - \mathbf{P}^{ij})(\mathbf{P}^i + \mathbf{P}^j - \mathbf{P}^{ij} - \mathbf{P}^{ij})^{-1}, \qquad (3)$$

where the cross-covariance $\mathbf{P}^{ij}$ has to be known. It is usually non-zero and generally hard to keep track of it.

C. Cross-Correlation

Instead of tracking the cross-covariance exactly, one can strive to obtain some partial knowledge about it, e.g., in the form of a bounded set. A conservative natural bound is given by positive semi-definiteness of the joint error covariance $\mathbf{J}$, which is used by CI [13]. However, in this paper we are interested in a smaller subset for bounding the possible cross-covariances.

For easier representation and learning of the joint error covariance $\mathbf{J}$, one may wish to work with its normalized version in the form of the cross-correlation matrix obtained by a decomposition

$$\mathbf{J} = \begin{bmatrix}\mathbf{S}^i & \mathbf{0}\\ \mathbf{0} & \mathbf{S}^j\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{\Omega}\\ \mathbf{\Omega}^\top & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{S}^i & \mathbf{0}\\ \mathbf{0} & \mathbf{S}^j\end{bmatrix}^\top, \qquad (4)$$

where $\mathbf{S}^i$ and $\mathbf{S}^j$ are obtained by the Cholesky decompositions of the covariances $\mathbf{P}^i$ and $\mathbf{P}^j$, respectively. The matrix $\mathbf{\Omega}$ containing the correlation coefficients (CCs) satisfies the natural bound

$$\mathbf{\Omega}\,\mathbf{\Omega}^\top \leq \mathbf{I}$$

for (4) being positive definite and equals to

$$\mathbf{\Omega} = (\mathbf{S}^i)^{-1}\mathbf{P}^{ij}(\mathbf{S}^j)^{-\top}. \qquad (5)$$

The advantage of the CC is that it provides intuitive information about the correlation. Furthermore, compared to the cross-covariance, a single CC value is naturally bounded between $[-1, 1]$. This paper focuses on learning a set of CCs smaller than the set given by the natural bounds. The aim is to utilize this information in the sensor nodes for the fusion together with the error covariances sent by the sensor nodes. As a result, the quality of the fused estimate should become better compared to CI using just the natural bounds.

III. Learning and Exploiting Sets of Correlation Coefficients

This section discusses i) the reasons for the CCs being unknown (in addition to the difficulty of keeping track of them), ii) the means to learn some information about the CC and derive an upper bound, and iii) exploiting the learned information for fusing local estimates.

A. Uncertainty of Correlation Coefficient

In this paper, the information about the CC will exclusively be considered in the form of its lower and upper bounds. The reason for the CC uncertainty may lie either in the *system model* uncertainty or in the uncertainty of the *data processing*. The system model uncertainty can be caused by unknown values of the dynamics or measurement model parameters or unknown initial conditions. Let M denote a system model consisting of the matrices $\mathbf{A}_k$, $\mathbf{H}_k$, $\mathbf{Q}_k$, $\mathbf{R}_k$, and $\mathbf{P}_0$ and let $\mathcal{M}$ denote the set of all models that are admissible, i.e., the models that correspond to the limited knowledge of the designer. Then, $M \in \mathcal{M}$ expresses the system model uncertainty. The set $\mathcal{M}$ is thus specified implicitly. Note that the system model uncertainty could

also emerge if a node uses a linear model obtained by linearization of a nonlinear one, and the linearization point is unknown to the other nodes. The CC uncertainty due to data processing can be caused by communication problems such as packet losses when a node may not receive the estimate from its neighbors or by uncertain/unknown network topology. The CC uncertainty can also be used to conveniently express the accuracy of the approximations involved in the fusion/estimation design. If a time-invariant model is considered, for example, the fusion can be tuned for a steady-state value of CC while it is also used in the transition process. The CC uncertainty can be used as an efficient tool to reflect this discrepancy.

B. Learning Correlation Coefficient

In principle, two approaches can be followed to obtain a bound for the CC. The first approach is based on an *analysis* of CC values for all system models, while the second one uses Monte Carlo *simulations* of the system.

Analytical approach: In the analytical approach, for each system model $M \in \mathcal{M}$, the cross-covariance matrices of the prediction estimation error $\mathbf{P}^{ij}_{k|k-1}$,

$$\begin{aligned}\mathbf{P}^{ij}_{k|k-1} &= E[(\underline{\hat{x}}^i_{k|k-1} - \underline{x}_k)(\underline{\hat{x}}^j_{k|k-1} - \underline{x}_k)^\top] \\ &= \mathbf{A}_{k-1}\mathbf{P}^{ij}_{k-1|k-1}\mathbf{A}^\top_{k-1} + \mathbf{Q}_{k-1}, \end{aligned} \tag{6}$$

and the cross-covariance matrices of the filtering estimation error $\mathbf{P}^{ij}_{k|k}$,

$$\begin{aligned}\mathbf{P}^{ij}_{k|k} &= E[(\underline{\hat{x}}^i_{k|k} - \underline{x}_k)(\underline{\hat{x}}^j_{k|k} - \underline{x}_k)^\top] \\ &= (\mathbf{I} - \mathbf{K}^i_k\mathbf{H}^i_k)\mathbf{P}^{ij}_{k|k-1}(\mathbf{I} - \mathbf{K}^j_k\mathbf{H}^j_k)^\top, \end{aligned} \tag{7}$$

are calculated. Then, the information about the CC matrix is obtained from (5).

Note that if the set $\mathcal{M}$ is infinite or too large, the bounds can be obtained approximately over a convenient subset (e.g., grid) of models. Such an approximation may lead to a slightly inconsistent fusion. Further, note that the techniques proposed in [24], [25], [26] belong to this approach. This approach is particularly suitable for the system model uncertainty of the CC.

Simulation approach: This approach consists of performing Monte Carlo simulations of the whole system, including the local estimation up to the next fusion step. The situation when a node does not receive the estimate from its neighbor can be efficiently modeled by an uncertain initial condition. The information about the CC is learned from the knowledge of the simulated state and the simulated measurement that leads to simulated estimates and corresponding estimate errors.

Let $\varepsilon^i(\ell)$ denote an estimation error of the n-th element of the state at the ℓ-th Monte Carlo simulation

$$\varepsilon^i(\ell) = \mathbf{e}_n^\top \left(\underline{x}_k(\ell) - \underline{\hat{x}}^i_{k|k}(\ell)\right),$$

where $\mathbf{e}_n$ is the n-th column of the identity matrix $\mathbf{I}$ of the corresponding dimension. Note that the element index n and time index k were dropped in $\varepsilon^i(\ell)$ for convenient purposes. Then, given the samples of the errors of the i-th and j-th state estimates, their sample (co-)variances $s^i = \frac{1}{L}\sum_{\ell=1}^{L}[\varepsilon^i(\ell)]^2$, $s^j = \frac{1}{L}\sum_{\ell=1}^{L}[\varepsilon^j(\ell)]^2$, and $s^{i,j} = \frac{1}{L}\sum_{\ell=1}^{L}\varepsilon^i(\ell)\varepsilon^j(\ell)$ can be used to obtain the sample CC as

$$c^{i,j} = \frac{s^{i,j}}{\sqrt{s^i s^j}}. \tag{8}$$

The uncertainty of the sample CC estimate in terms of the confidence interval can be calculated using the Fisher z-transform [28]. Since the probability distribution of the correlation becomes extremely skewed towards the natural bounds of the correlation coefficients, the Fisher z-transformation reduces this skewed curve to approximate a normal distribution. The limits of the confidence intervals are then treated as the CC bounds. Note that for large values of L, the sample CC estimate variance is almost identical [29] to

$$\sigma^2_{i,j} = \frac{\left(1 - (c^{i,j})^2\right)^2}{L-2}. \tag{9}$$

While (9) is simpler for specification of confidence intervals, it does lead to probability mass outside of the admissible area of CC. The Fisher z-transform does not suffer from this problem.

The simulation approach described above can be used for learning the CC matrix using confidence sets element-wise only. This is a significant difference from the analytical approach, which can naturally work with whole CC matrices.

The simulation approach is expensive in terms of communication and computing power. The approach is suitable for CC uncertainty due to data processing and can also be employed for a generalization of the CC learning and exploiting concept to nonlinear system estimation. Note that the learning procedures proposed in [22], [23] stay on the borderland between the analytical and simulation approaches.

C. Exploiting Correlation Coefficient

While the previous sections dealt with implicit specification of the admissible set $\mathcal{M}$, this section assumes that the implicit specification has already been transformed to an explicit one. At this point, it is irrelevant whether the explicit specification of the admissible set has been obtained from the model or by a learning procedure.

Assume the learned information about the CC matrix is available in the form

$$\mathbf{\Omega} \in \left\{\mathbf{T}^{ij} + \mathbf{T}^i\Lambda(\mathbf{T}^j)^\top \,\middle|\, \mathbf{I} - \Lambda\Lambda^\top \geq 0\right\}, \tag{10}$$

where Λ, $\mathbf{T}^i$, $\mathbf{T}^j$, and $\mathbf{T}^{ij}$ are matrices of appropriate dimensions and the matrices $\mathbf{T}^i$ and $\mathbf{T}^j$ are full column rank. Also, suppose the decompositions $\mathbf{S}^i$ and $\mathbf{S}^j$ of error covariance matrices $\mathbf{P}^i$ and $\mathbf{P}^j$, respectively, are provided by the sensor nodes. The parametrization of the set (10) corresponds to the split CI fusion [18], which assumes the joint error covariance $\mathbf{J}$ equals to a sum of error covariance term with known correlation denoted as $\mathbf{J}^k$ and error covariance with unknown correlation denoted as $\mathbf{J}^u$,

$$\mathbf{J} = \mathbf{J}^k + \mathbf{J}^u,$$

where

$$\mathbf{J}^k = \begin{bmatrix} \mathbf{P}^i - \mathbf{S}^i\mathbf{T}^i(\mathbf{T}^i)^\top(\mathbf{S}^i)^\top & \mathbf{S}^i\mathbf{T}^{ij}(\mathbf{S}^j)^\top \\ \mathbf{S}^j(\mathbf{T}^{ij})^\top(\mathbf{S}^i)^\top & \mathbf{P}^j - \mathbf{S}^j\mathbf{T}^j(\mathbf{T}^j)^\top(\mathbf{S}^j)^\top \end{bmatrix}$$

and

$$\mathbf{J}^u = \begin{bmatrix} \mathbf{S}^i\mathbf{T}^i(\mathbf{S}^i\mathbf{T}^i)^\top & \mathbf{S}^i\mathbf{T}^i\Lambda(\mathbf{T}^j)^\top(\mathbf{S}^j)^\top \\ \mathbf{S}^j\mathbf{T}^j\Lambda^\top(\mathbf{T}^i)^\top(\mathbf{S}^i)^\top & \mathbf{S}^j\mathbf{T}^j(\mathbf{S}^j\mathbf{T}^j)^\top \end{bmatrix},$$

where $\mathbf{I} - \Lambda\Lambda^\top \geq 0$. The upper bound for the joint error-covariance is then parameterized by $\mu > 0$ as

$$\begin{bmatrix} \mathbf{P}^i + \mu\mathbf{S}^i\mathbf{T}^i(\mathbf{T}^i)^\top(\mathbf{S}^i)^\top & \mathbf{S}^i\mathbf{T}^{ij}(\mathbf{S}^j)^\top \\ \mathbf{S}^j(\mathbf{T}^{ij})^\top(\mathbf{S}^i)^\top & \mathbf{P}^j + \frac{1}{\mu}\mathbf{S}^j\mathbf{T}^j(\mathbf{T}^j)^\top(\mathbf{S}^j)^\top \end{bmatrix}. \quad (11)$$

The upper bound (11) is then used instead of the unknown joint error covariance $\mathbf{J}$ and its elements are used for the fusion (3).

A special case of (10) has been used in [15] considering

$$\boldsymbol{\Omega} \in \left\{ \mathbf{T}^{ij} + \rho\Lambda \middle| \mathbf{I} - \Lambda\Lambda^\top \geq 0 \right\},$$

where $0 \leq \rho \leq 1$ is a known parameter. Then, the upper bound of the joint error covariance $\mathbf{J}$ is

$$\begin{bmatrix} \mathbf{P}^i(1+\rho\mu) & \mathbf{T}^{ij} \\ (\mathbf{T}^{ij})^\top & \mathbf{P}^j(1+\rho\frac{1}{\mu}) \end{bmatrix}.$$

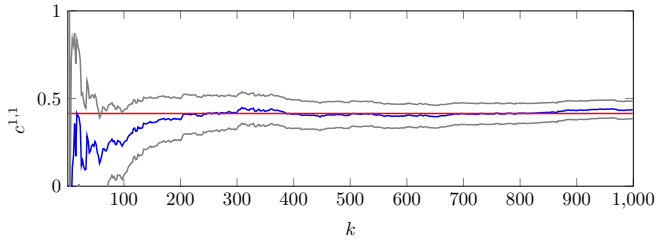
Because of the centering matrix $\mathbf{T}^{ij}$, this method can be seen as an asymmetric bound. The scalar value ρ is bounding the set of admissible CC matrices.

The above considerations dealt with the fusion of two estimates. For fusion of more than two estimates, the following aspects have to be investigated. A CC matrix $\boldsymbol{\Omega}$ for each pair of estimates has to be specified. Some combinations of the CC matrices need not lead to a positive semidefinite joint matrix $\mathbf{J}$. Construction of the optimal upper bounds for $\mathbf{J}$ is unclear when more than two estimates are fused.

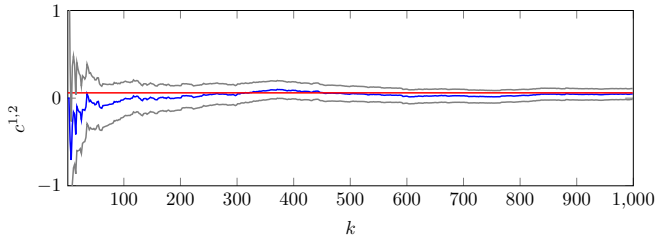
IV. Evaluation

We now discuss a numerical example of learning the CC matrix using Monte Carlo simulations. Kalman filters perform local estimation and then communicate their estimates to a central fusion node that performs the fusion. We will use the Markov assumption about the estimators in which both estimators' output and the correlation of the state estimates only depend on the input of the local filters, which is the initial covariance matrix. Furthermore, we assume that the employed local system and measurement models of system set $\mathcal{M}$ are unknown. Therefore, we can only estimate the correlation of the state estimates based on the output of the local estimators and the estimation error.

We will first examine the problem of estimating the unknown CC matrix between two state estimates over time using only a single known input covariance matrix $\mathbf{P}_0$. This procedure is then extended to learning the CC matrix when applying different input covariance matrices generated with a random scaling factor in a known range.



(a) Estimated $c^{1,1}$.



(b) Estimated $c^{1,2}$.

Fig. 1: Estimated correlation coefficients $c^{1,1}$ and $c^{1,2}$ (blue) and the estimated 95% confidence interval ρ (gray) with increasing number of training covariances compared to the analytically calculated correlation coefficients $\varrho^{1,1}$ and $\varrho^{1,2}$ (red) for a single training covariance.

A. Learning Correlations for Single Input Covariance

We consider two sensor nodes A and B that both estimate the state of a discrete-time time-invariant linear stochastic dynamic system with

$$\mathbf{A} = \begin{bmatrix} 1 & \Delta T \\ 0 & 1 \end{bmatrix}, \ \mathbf{Q} = 0.1 \begin{bmatrix} \frac{\Delta T^3}{3} & \frac{\Delta T^2}{2} \\ \frac{\Delta T^2}{2} & \Delta T \end{bmatrix}, \ \Delta T = 0.1\,.$$

Both sensor nodes use a linear measurement model with measurement matrices $\mathbf{H}^A = \mathbf{H}^B = \mathbf{I}$. Each observation is corrupted by additive white Gaussian noise with variances

$$\mathbf{R}^A = \mathrm{diag}(0.5^2, 0.5^2)\,, \ \mathbf{R}^B = \mathrm{diag}(5^2, 1^2)\,.$$

Both Kalman filters of the local sensor nodes are initialized with the same input covariance matrix $\mathbf{P}_0 = [1, 0.5; 0.5, 1]$. Afterward, they execute one prediction step and one measurement update. The local estimates are then sent back to the fusion center to be fused. Sensor node A sends the full state estimate $\hat{\underline{x}}^A$ and covariance $\mathbf{P}^A$ to the fusion center while node B sends only its first state $\hat{\underline{x}}_1^B$ and the belonging variance $\mathbf{P}_{11}^B$. The joint covariance matrix as already discussed in (4) is given as

$$\mathbf{J} = \begin{bmatrix} \mathbf{S}^A & 0 \\ 0 & S^B \end{bmatrix} \left[\begin{array}{cc:c} 1 & \varrho^A & \varrho^{1,1} \\ \varrho^A & 1 & \varrho^{1,2} \\ \hdashline \varrho^{1,1} & \varrho^{1,2} & 1 \end{array}\right] \begin{bmatrix} \mathbf{S}^A & 0 \\ 0 & S^B \end{bmatrix}^\top,$$

where $\mathbf{S}^A$ is the square-root decomposition of the elements on the main diagonal of $\mathbf{P}^A$ and S^B is the square-root of $\mathbf{P}_{11}^B$. The CC ϱ^A is given by the covariance matrix $\mathbf{P}^A$. The unknown CCs $\varrho^{1,1}$ and $\varrho^{1,2}$ describe how the two state estimates are correlated. In the beginning, no

knowledge about the fusion is available to the fusion center, and, therefore, covariance intersection (CI) is used to obtain consistent fusion results. Using the simulated ground truth and the state estimates, we can estimate the sample CC over time by using (8). Further, the uncertainty of every CC estimate $\sigma^2_{i,j}$ is calculated using (9). The resulting estimated CCs $c^{1,1}$ and $c^{1,2}$ and the 95% confidence interval of the estimate is shown in Fig. 1. In the beginning, the uncertainty is large and reduces over time. However, because of the naïve calculation of the uncertainty, the values violate the natural bounds of the CC matrix in the beginning, when only a small number of estimations are available.

For the fusion, we decided to exploit the partial knowledge by using the asymmetric bounds described by [15] as already discussed in Sec. III-C. This asymmetric bound requires the centering matrix $\mathbf{T}^{AB}$ and the scalar parameter ρ. Because we estimate two CCs, ρ is equal to a circle that includes all admissible CCs. We calculate ρ by using the largest singular value of the CC estimates

$$\rho = 1.96\sqrt{\max(\mathrm{svd}\left(\sigma^2\right))}$$

and calculating the 95% confidence interval. In our case this is equal to taking the largest variance. The estimated CCs $c^{1,1}$ and $c^{1,2}$ are the entries of the off-diagonal matrix $\mathbf{T}^{AB} = [c^{1,1}, c^{1,2}]^\top$.

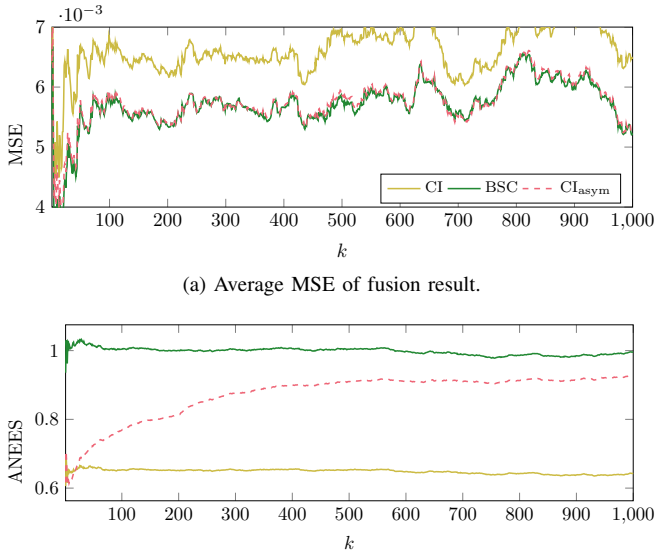
The improvement of the fusion result over time after $L = 50$ Monte Carlo simulations using the estimated CCs is shown in Fig. 2. The mean squared error (MSE) in Fig. 2a drops immediately to a value close to the optimal fusion using the Bar-Shalom/Campo (BSC) formula with the analytically calculated cross-covariance using (6) and (7). Covariance Intersection (CI) shows a slightly higher MSE. The consistency of the fused estimate is evaluated using the averaged normalized estimation error squared (ANEES) [30], where the ANEES of time step k and Monte Carlo simulation ℓ is defined as

$$\bar{\epsilon}_k = \frac{1}{nL}\sum_{\ell=1}^{L}\left(\underline{x}_k(\ell) - \underline{\hat{x}}^{\mathrm{f}}_k(\ell)\right)^\top \mathbf{P}^{\mathrm{f}}(\ell)^{-1}\left(\underline{x}_k(\ell) - \underline{\hat{x}}^{\mathrm{f}}_k(\ell)\right),$$

with the system state $\underline{x}_k$, and the fused estimate $\underline{\hat{x}}^{\mathrm{f}}$ and covariance matrix $\mathbf{P}^{\mathrm{f}}$ of state dimension n. The ANEES indicates whether the estimated uncertainty matches the actual MSE. Therefore, values above 1 indicate an inconsistent estimator which underestimates the uncertainty and values below 1 indicate an conservative estimator which overestimates the uncertainty. In Fig. 2b we can see that CI shows a considerably lower ANEES, which indicates that the uncertainty is overestimated. The fusion using the asymmetric bound ($\mathrm{CI}_{\mathrm{asym}}$) shows over-bounding in the beginning when the uncertainty is outside of the natural bounds of the CC matrix but improves significantly over time when ρ becomes smaller.

B. Learning Correlations for Several Input Covariances

In the second example, we want to explore learning the CC matrix with different input covariances and then interpolate between the learned data points. We generate a set of covariance matrices with linear scaling factor $\mathbf{P}_0 = p[1, 0.5; 0.5, 1]$, where p is randomly chosen between $p_{\min} = 0.01$ and $p_{\max} = 1$. As in the first example, the covariance matrix is used to initialize the local Kalman filters, which again execute one prediction and one filtering step and send their state estimates back to the fusion center. For each simulation run, we save the scaling value p and the estimation errors. For the estimation of $c^{1,1}$ and $c^{1,2}$, we use a search radius of $r = 0.01$ to find data points with similar scaling factor p and use the acquired estimation errors for the estimation of the CCs. Over time, more data is available, and thus the uncertainty of the CC estimates decreases, as can be seen in Fig. 4. The plots show that the CC estimates are very close to the analytically calculated CCs.



(a) Average MSE of fusion result.

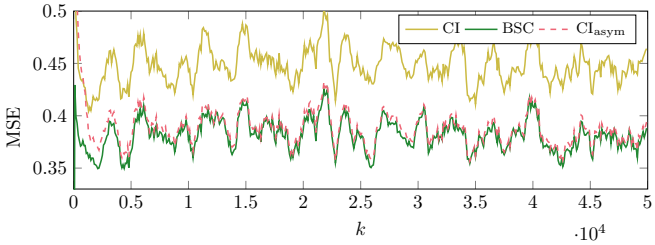
(b) Average ANEES of fusion result.

Fig. 2: Comparison of the fusion results of different algorithms for 50 Monte Carlo simulations, moving average over the last 200 values.

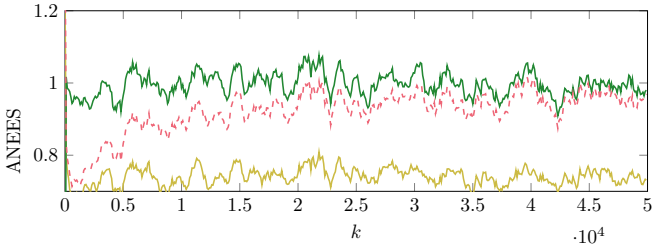
We also simulate the improvement of the fusion result over time which can be seen in Fig. 3. CI does show a higher MSE and is more conservative than the optimal fusion and the asymmetric bounds using the estimated CCs. These bounds are close to the performance of CI in the beginning but approach the fusion result of the optimal fusion over time as more data points are available.

V. Conclusion

This paper explored different approaches to learning partial knowledge about correlations in distributed estimation and proposed methods to bound sets of possible CCs to obtain consistent estimates. A natural progression of this work is to investigate the estimation of CC matrices to ensure that the natural bounds are not violated. A possible way to achieve this might be estimating CC matrices directly in the correlation domain or a suitable space that naturally only allows admissible values.



(a) Average MSE of fusion result.



(b) Average ANEES of fusion result.

Fig. 3: Comparison of the fusion results of different algorithms with improved partial knowledge over time, moving average over the last 1000 values.

This paper also showed that the proposed bounding technique using one scalar value is limited in its performance and should be developed further to provide even tighter bounds. The topic of transforming the set of admissible models to a set of correlations should also be investigated in the future. Future research should be conducted to investigate how to use the sets of CC matrices directly instead of bounding them, providing even less conservative results. Finally, the topic of transforming the set of admissible models to a set of correlations should also be investigated in the future.

References

[1] D. L. Hall, C.-Y. Chong, J. Llinas, and M. E. Liggins II, Eds., *Distributed Data Fusion for Network-Centric Operations*, ser. The Electrical Engineering and Applied Signal Processing Series. Boca Raton, Florida, USA: CRC Press, 2013.

[2] S. Grime and H. F. Durrant-Whyte, "Data Fusion in Decentralized Sensor Networks," *Control Engineering Practice*, vol. 2, no. 5, pp. 849–863, Oct. 1994.

[3] Y. Bar-Shalom and L. Campo, "On the Track-to-Track Correlation Problem," *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 571–572, Apr. 1981.

[4] C. Chong, K. Chang, and S. Mori, "A Review of Forty Years of Distributed Estimation," in *Proceedings of the 21st International Conference on Information Fusion (Fusion 2018)*, Cambridge, United Kingdom, Jul. 2018, pp. 1–8.

[5] A. G. O. Mutambara, *Decentralized Estimation and Control for Multisensor Systems*. Boca Raton, Florida, USA: CRC Press, 1998.

[6] A. G. O. Mutambara and H. F. Durrant-Whyte, "Nonlinear Information Space: A Practical Basis for Decentralization," in *Proceedings of SPIE's International Symposium on Photonics for Industrial Applications, Sensor Fusion VII*, P. S. Schenker, Ed., vol. 2355, no. 1. SPIE, 1994, pp. 97–105.

[7] B. Noack, D. Lyons, M. Nagel, and U. D. Hanebeck, "Nonlinear Information Filtering for Distributed Multisensor Data Fusion," in *Proceedings of the 2011 American Control Conference (ACC 2011)*, San Francisco, California, USA, Jun. 2011, pp. 4846–4852.

[8] C.-Y. Chong and S. Mori, "Convex Combination and Covariance Intersection Algorithms in Distributed Fusion," in *Proceedings of the 4th International Conference on Information Fusion (Fusion 2001)*, Montréal, Québec, Canada, Aug. 2001.

[9] J. Ajgl and O. Straka, "Covariance Intersection in Track-to-Track Fusion: Comparison of Fusion Configurations," *IEEE Transactions On Industrial Informatics*, vol. 14, no. 3, pp. 1127–1136, March 2018.

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

[11] L. Chen, P. O. Arambel, and R. K. Mehra, "Fusion under Unknown Correlation: Covariance Intersection Revisited," *IEEE Transactions on Automatic Control*, vol. 47, no. 11, pp. 1879–1882, Nov. 2002.

[12] M. Reinhardt, B. Noack, P. O. Arambel, and U. D. Hanebeck, "Minimum Covariance Bounds for the Fusion under Unknown Correlations," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1210–1214, Sep. 2015.

[13] J. Ajgl and O. Straka, "A Geometrical Perspective on Fusion under Unknown Correlations based on Minkowski Sums," in *Proceedings of the 20th International Conference on Information Fusion*, Xi'an, China, July 2017, pp. 725–732.

[14] U. D. Hanebeck, K. Briechle, and J. Horn, "A Tight Bound for the Joint Covariance of Two Random Vectors with Unknown but Constrained Cross–Correlation," in *Proceedings of the 2001 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2001)*, Baden–Baden, Germany, Aug. 2001, pp. 85–90.

[15] S. Reece and S. Roberts, "Robust, Low-Bandwidth, Multi-Vehicle Mapping," in *2005 7th International Conference on Information Fusion*, vol. 2, 2005, p. 8.

[16] Z. Wu, Q. Cai, and M. Fu, "Covariance Intersection for Partially Correlated Random Vectors," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 619–629, 2018.

[17] J. Ajgl and O. Straka, "Comparison of Fusions Under Unknown and Partially Known Correlations," in *IFAC-PapersOnLine, 7th IFAC Workshop on Distributed Estimation and Control in Networked Systems (NecSys18)*, vol. 51, no. 23, Groningen, The Netherlands, August 2018, pp. 295–300.

[18] S. J. Julier and J. K. Uhlmann, *Handbook of Multisensor Data Fusion: Theory and Practice*, 2nd ed. Boca Raton, Florida, USA: CRC Press, 2009, ch. General Decentralized Data Fusion with Covariance Intersection, pp. 319–343.

[19] B. Noack, J. Sijs, M. Reinhardt, and U. D. Hanebeck, "Decentralized Data Fusion with Inverse Covariance Intersection," *Automatica*, vol. 79, pp. 35–41, May 2017.

[20] B. Noack, J. Sijs, and U. D. Hanebeck, "Inverse Covariance Intersection: New Insights and Properties," in *Proceedings of the 20th International Conference on Information Fusion (Fusion 2017)*, Xi'an, China, Jul. 2017, pp. 1–8.

[21] J. Čurn, D. Marinescu, N. O'Hara, and V. Cahill, "Data Incest in Cooperative Localisation with the Common Past-Invariant Ensemble Kalman Filter," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, Istanbul, Turkey, Jul. 2013, pp. 68–76.

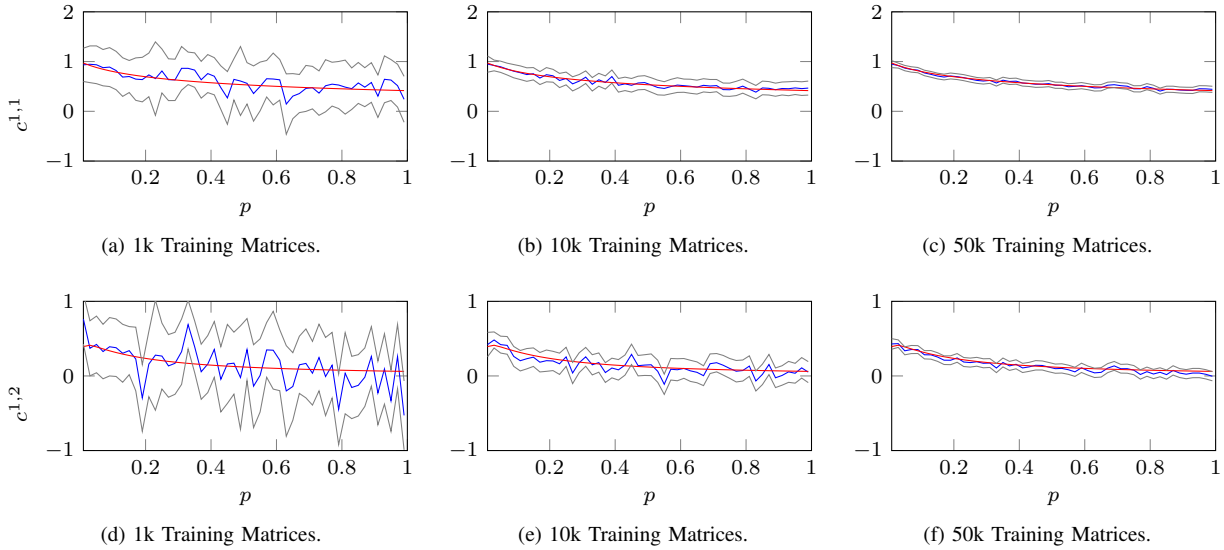
[22] M. Reinhardt, B. Noack, and U. D. Hanebeck, "Reconstruction of Joint Covariance Matrices in Networked Linear Systems," in *Proceedings of the 48th Annual Conference on Information Sciences and Systems (CISS 2014)*, Princeton, New Jersey, USA, Mar. 2014, pp. 1–6.

[23] J. Steinbring, B. Noack, M. Reinhardt, and U. D. Hanebeck, "Optimal Sample-Based Fusion for Distributed State Estimation," in *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*, Heidelberg, Germany, Jul. 2016, pp. 1600–1607.

[24] S. Radtke, B. Noack, U. D. Hanebeck, and O. Straka, "Reconstruction of Cross-Correlations with Constant Number of Deterministic Samples," in *Proceedings of the 21st International Conference on Information Fusion (Fusion 2018)*, Cambridge, United Kingdom, Jul. 2018, pp. 1638–1645.

[25] M. Reinhardt, S. Kulkarni, and U. D. Hanebeck, "Generalized Covariance Intersection based on Noise Decomposition," in *Proceedings of the 2014 IEEE International Conference on Multisensor Fusion and Information Integration (MFI 2014)*, Beijing, China, Sep. 2014, pp. 1–8.

[26] S. Radtke, B. Noack, and U. D. Hanebeck, "Distributed Estimation using Square Root Decompositions of Dependent Information," in *Proceedings of the 22nd International Conference on Information Fusion (Fusion 2019)*, Ottawa, Canada, Jul. 2019, pp. 1–8.



(a) 1k Training Matrices. (b) 10k Training Matrices. (c) 50k Training Matrices.

(d) 1k Training Matrices. (e) 10k Training Matrices. (f) 50k Training Matrices.

Fig. 4: Estimated correlation coefficients $c^{1,1}$ and $c^{1,2}$ (blue) and the estimated 95% confidence interval ρ (gray) with increasing number of training covariances compared with the analytically calculated correlation coefficients $\varrho^{1,1}$ and $\varrho^{1,2}$ (red) for random training covariances with different scaling factor p.

[27] S. J. Julier, "Estimating and Exploiting the Degree of Independent Information in Distributed Data Fusion," in *Proceedings of the 12th International Conference on Information Fusion (Fusion 2009)*, Seattle, Washington, USA, Jul. 2009.

[28] R. A. Fisher, "On the "Probable Error" of a Coefficient of Correlation Deduced from a Small Sample," *Metron*, vol. 1, pp. 3–32, 1921.

[29] A. L. Bowley, "The Standard Deviation of the Correlation Coefficient," *Journal of the American Statistical Association*, vol. 23, no. 161, pp. 31–34, 1928.

[30] X. R. Li, Z. Zhao, and V. P. Jilkov, "Practical Measures and Test for Credibility of an Estimator," in *Proceedings of Workshop on Estimation, Tracking, and Fusion - A Tribute to Yaakov Bar-Shalom*, 2001.