Generalized Covariance Intersection based on Noise Decomposition

Marc Reinhardt[*], **Sanjeev Kulkarni**[†], and **Uwe D. Hanebeck**[*]

[*]Intelligent Sensor-Actuator-Systems Laboratory (ISAS)
Institute for Anthropomatics and Robotics
Karlsruhe Institute of Technology (KIT), Germany
marc.reinhardt@ieee.org, uwe.hanebeck@ieee.org

[†]Department of Electrical Engineering
Princeton University, USA
kulkarni@princeton.edu

***Abstract*—In linear decentralized estimation, several nodes concurrently aim to estimate the state of a common phenomenon by means of local measurements and data exchanges. In this contribution, an efficient algorithm for consistent estimation of linear systems in sensor networks is derived. The main theorems generalize Covariance Intersection by means of an explicit consideration of individual noise terms. We apply the results to linear decentralized estimation and obtain covariance bounds with a scalable precision between the exact covariances and the bounds provided by Covariance Intersection.**

I. Introduction

With the development of new algorithms and technologies for sensing and actuating systems, the complexity of technical systems has reached a point in which decentralized processing of data is quite often advantageous or even the only feasible solution [1]. In decentralized architectures, a variety of new challenges arise. The independent processing of sensed values that describe related information, e.g., the position estimate of multi-axial robot arms, constitutes one of the key problems.

Although many real-world estimation problems are inherently nonlinear, in particular for decentralized systems, filter algorithms are often based on linear approaches such as the Kalman filter (KF) [2]. Primarily, this is due to the possibility to process data efficiently in a recursive manner and to quantify the estimation quality in terms of the mean-square-error (MSE) in closed form.

In linear decentralized estimation, either (transformed) measurements or estimates are communicated. The idea of the former is to process sensed values on the receiver side. In the seminal work [3], it was shown that even in the presence of stochastic communication, the MSE minimizing processing can be derived. Many results on that topic have followed since, considering an arbitrary number of sensors, different communication networks, or other constraints (e.g., [4]).

The major drawback of exchanging measurements is the need to communicate all sensed information separately. Especially for communication-restricted networks or in the presence of packet losses, the performance of the sensor-network suffers. Hence, it is preferable to recursively combine measurements in local estimates and process these quantities in the network. Popular approaches that guarantee asymptotic properties for linear time-invariant (LTI) systems are Kalman-Consensus filters [5] and diffusion strategies [6].

The general decentralized estimation problem, which involves time-variant systems and arbitrary communication networks, still is a challenging research topic. Methods for the local processing of estimates, i.e., the filter strategy, have been derived [7], [8] that outperform local KFs [2]. However, the optimal fusion of exchanged estimates encompasses correlations between estimates [9]–[11] and neglecting these dependencies, e.g., as proposed for the convex combination [12], may lead to diverging estimates.

While different solutions to approximate the cross-covariances in a central node have been presented [13], [14], a decentralized technique to calculate and maintain correlations has not been proposed yet. The lack of viable methods has given rise to the development of fusion methods that operate under unknown correlations [15]–[20].

Covariance Intersection (CI) [15] bypasses the explicit reconstruction of cross-covariances by artificially inflating the (known) local covariances with scalar factors. This approach was originally developed in the context of set theoretic estimation [21] and guarantees that the true fused covariance is bounded, i.e., the true covariance is overestimated in the positive semi-definite sense. The technique has been combined with various linear filters, e.g., recently with the diffusion Kalman filter [6] in [22].

Although CI provides the smallest bound in a large class of algorithms [17], it is often considered as too conservative. Split-CI has been proposed to shrink the bound when independent information is included in the fusion process [18]. An extension of Split CI for the estimation with more than two nodes has been proposed in [23]. In [19], [20], local filters have been derived that allow the explicit consideration of correlations in the bounding procedure. However, for all of these generalizations, cross-covariances must be assumed unknown after the fusion of estimates in order to guarantee consistent results.

In this contribution, we provide a consistent estimator with

adjustable precision. To this end, we combine a technique for the explicit reconstruction of cross-covariances based on the decomposition of errors into noise terms with covariance bounding ideas for decentralized estimation.

We derive a novel bounding technique in Sec. III, which is applicable to arbitrary linear estimators. No assumptions are made about communication, remote model knowledge, and other sensor network parameters. It generalizes (basic) CI in terms of the number of considered estimates and the inclusion of explicit cross-covariance information. In contrast to other generalizations of CI [18]–[20], the approach provides detailed covariance information even after the fusion of estimates and is therefore well-suited for decentralized estimation. The bounding technique is then applied to decentralized estimation in Sec. IV. To this end, the local processing at the sensors is discussed and optimal fusion gains are derived. In the course of proving consistency and optimality for the derived approach, these attributes are also derived for CI with more than two sensors. Finally, the proposed algorithm is evaluated in Sec. V.

Notation

If not otherwise stated, vectors are assumed to be column vectors and are characterized by an underscore $\underline{x}$. Random vectors are bold typed. Matrices are depicted as bold capital letters $\mathbf{M}$. Relations between matrices are to be understood in the positive semi-definite sense if not otherwise stated. We denote the real numbers as $\mathbb{R}$, positive integers as $\mathbb{N}^+$, and the set of symmetric positive (semi-) definite matrices of dimension n as $\mathbb{S}_+^{n\times n}$ ($\mathbb{S}^{n\times n}$). The transpose of vectors and matrices is denoted by $\mathbf{M}^\top$ and the inverse by $(\mathbf{M})^{-1}$. The trace operator is denoted by $\mathrm{tr}\{\mathbf{M}\}$. We define $(\underline{x})^2 = (\underline{x})(\underline{x})^\top$ and denote the identity matrix concatenation $(\mathbf{I} \quad \dots \quad \mathbf{I})^\top$ as $\underline{\mathbf{1}}$.

II. Problem Formulation

We consider a sensor-network with sensors $s \in \mathcal{S}$, $S = |\mathcal{S}|$ that aim to estimate the state of a common phenomenon. The system state $\underline{\boldsymbol{x}}_k \in \mathbb{R}^n$ evolves according to the linear time-discrete system

$$\underline{\boldsymbol{x}}_{k+1} = \mathbf{A}_k \underline{\boldsymbol{x}}_k + \underline{\boldsymbol{w}}_k \ , \mathbf{A}_k \in \mathbb{R}^{n\times n} \ . \tag{1}$$

The sensors observe the system state by means of measurements that are obtained according to

$$\underline{z}_k^s = \mathbf{H}_k^s \underline{\boldsymbol{x}}_k + \underline{\boldsymbol{v}}_k^s \ , \underline{z}_k^s \in \mathbb{R}^{m_s} \ , \mathbf{H}_k^s \in \mathbb{R}^{m_s \times n} \ . \tag{2}$$

All noise terms $\underline{\boldsymbol{w}}_k \in \mathbb{R}^n, \underline{\boldsymbol{v}}_k^s \in \mathbb{R}^{m_s}$ are assumed to be zero-mean and independent of each other. The probability distribution of the noise terms is arbitrary with covariances $\mathrm{E}\{\underline{\boldsymbol{w}}_k(\underline{\boldsymbol{w}}_k)^\top\} = \mathbf{C}_k^w$ and $\mathrm{E}\{\underline{\boldsymbol{v}}_k^s(\underline{\boldsymbol{v}}_k^s)^\top\} = \mathbf{C}_k^{v_s}$. The sensors are assumed to have full knowledge about the process model (1) and exclusive knowledge about their respective local sensor model (2). By communicating with their neighbors, sensor information is spread through the network.

The objective of this paper is to provide consistent estimators that outperform state-of-the-art methods in terms of accuracy and concurrently keep communication and computational effort limited.

Let $\hat{\underline{x}}^s$ denote a sensor estimate with true covariance $\mathbf{P}^s = \mathrm{E}\{(\hat{\underline{x}}^s - \underline{\boldsymbol{x}})^2\}$. Due to common prior information [12] and common process noise [9], correlations emerge that cannot be calculated without sensor models and estimator transformations of remote sensors from all past time steps [13]. As not only the optimal fusion gains [9]–[11] but also the covariance of the fused estimate depend on these unknown correlations, the fusion of estimates poses a great challenge in sensor networks.

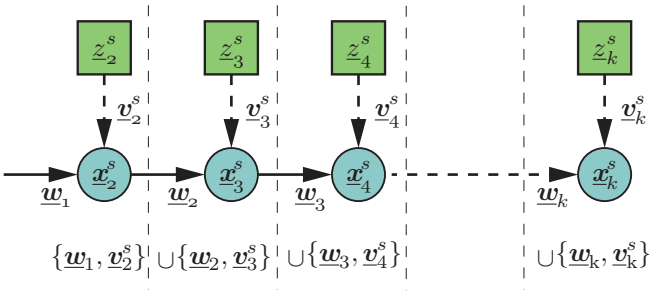

Figure 1. The emerging noise terms due to prediction and filtering at sensor s for different time steps.

However, it has been shown in [18], [24] that consistent estimation can still be realized when covariance bounds with $\bar{\mathbf{P}}^s \geq \mathbf{P}^s$ are processed for all estimates $\hat{\underline{x}}^s$. Therefore, two challenges arise in the course of providing efficient and consistent estimators in sensor networks. In a first step, covariance bounds must be obtained in a (preferably recursive manner) at the sensors without knowledge about the models and processing at remote nodes. In a second step, the filter and fusion processing is to be optimized so that the covariance bounds are minimized.

III. Consistent Bounds for Linear Estimators

In this section, we derive covariance bounds by means of a noise decomposition of errors. All necessary calculations are carried out at the sensors subject to the information constraints described in Sec. II, i.e., independent from models of remote sensors.

A naïve idea to calculate joint covariances in sensor networks is to store and exchange the estimator operations and noise covariances of all sensors. Based on this information, techniques from [13] can be utilized to obtain (cross-)covariances. In a first step, we improve this naïve processing by deriving formulas for the recursive calculation of necessary quantities.

Linear estimators are defined by their processing in initialization, prediction, filter, and fusion steps. As it is depicted in Fig. 1, the estimation error for linear estimators is composed of measurement and process noise terms. Let $\mathcal{I}$ denote a set of unique identifiers for all noise terms occurring during the estimation process in the sensor network and let $\mathcal{I}^s \subseteq \mathcal{I}$ denote those terms that affect the estimate of sensor s. The dimension of noise term $i \in \mathcal{I}$ is denoted by $m_i \in \mathbb{N}^+$. Then, the estimation error of linear estimators has the following form.

Lemma 1 *For the considered system* (1) *with measurement models* (2)*, the error of a linear estimator* $\hat{\underline{e}}^s = \hat{\underline{x}}^s - \underline{\boldsymbol{x}}$ *is given as a linear combination of independent noise terms* $\hat{\underline{e}}^s = \sum_{i\in\mathcal{I}^s} \mathbf{B}_i^s \underline{\boldsymbol{n}}_i$*, where* $\mathbf{B}_i^s \in \mathbb{R}^{n\times m_i}$ *denote sensor-specific transformation matrices.*

PROOF. The claim is satisfied at initialization with $\mathbf{B}_i^s = \mathbf{I}$ and by noise terms $\underline{\boldsymbol{n}}_i$ that reflect the initial uncertainties. In the prediction step, noise terms from previous operations are linearly transformed with $\mathbf{A}_k$ and a new noise term $\underline{\boldsymbol{w}}_k$ is included. Let $\mathbf{K}_k^s$ denote the filter gain in sensor s. Then, in the filter step, noise terms from previous operations are transformed with $\mathbf{I} - \mathbf{K}_k^s \mathbf{H}_k^s$ and $\mathbf{K}_k^s \underline{\boldsymbol{v}}_k^s$ is added as a new noise term.

The fusion of estimates $\hat{\underline{x}}^s$, $s \in \mathcal{S}' \subseteq \mathcal{S}$ according to $\hat{\underline{x}}^f = \sum_{s \in \mathcal{S}'} \mathbf{F}^s \hat{\underline{x}}^s$ with fusion gains $\mathbf{F}^s$ yields the fused error $\hat{\underline{e}}^f = \sum_{s \in \mathcal{S}'} \mathbf{F}^s \hat{\underline{e}}^s$. Let $\mathbf{B}_i^s = \mathbf{0}$ when $\underline{\boldsymbol{n}}_i$ is not comprised in the estimation error of sensor s. Then, the fused error is given by the linear combination of noise terms $\hat{\underline{e}}^f = \sum_{i \in \mathcal{I}} \mathbf{B}_i^f \underline{\boldsymbol{n}}_i$, with $\mathbf{B}_i^f = \sum_{s \in \mathcal{S}'} \mathbf{F}^s \mathbf{B}_i^s$. □

The matrices $\mathbf{B}_i^s$ entail the linear transformations that have been applied to the estimate since the inclusion of the respective noise term $\underline{\boldsymbol{n}}_i \in \{\underline{\boldsymbol{w}}_t, \underline{\boldsymbol{v}}_t^s | t \leq k \ , s \in \mathcal{S}\}$. For example in the filter step, the estimate is transformed with the gain $\mathbf{I} - \mathbf{K}_k^s \mathbf{H}_k^s$ and a new noise term $\mathbf{K}_k^s \underline{\boldsymbol{v}}_k^s$ is included. More details on the transformations are given in the proof of Lemma 1.

The error processes define covariances according to $\mathbf{P}^{s_1 s_2} = \mathrm{E}\{\hat{\underline{e}}^{s_1} (\hat{\underline{e}}^{s_2})^\top\}$. We exploit the independence of noise terms $\underline{\boldsymbol{n}}_i$ and $\underline{\boldsymbol{n}}_j$, $i \neq j \in \mathcal{I}$ and obtain the following lemma.

Lemma 2 *Let $\mathbf{P}_i^n \in \mathbb{S}^{m_i \times m_i}$ denote the covariances of independent noise terms $\underline{\boldsymbol{n}}_i$, $i \in \mathcal{I}$, and let $\mathbf{B}_i^s$ denote the estimator specific transformation matrices from Lemma 1. Then, it holds*

$$\mathbf{P}^{s_1 s_2} = \sum_{i \in \mathcal{I}^{s_1} \cap \mathcal{I}^{s_2}} \mathbf{B}_i^{s_1} \mathbf{P}_i^n (\mathbf{B}_i^{s_2})^\top \ .$$

PROOF. With the independence of noise terms $i \neq j \in \mathcal{I}$, it directly follows $\mathrm{E}\{\underline{\boldsymbol{n}}_i (\underline{\boldsymbol{n}}_j)^\top\} = \mathbf{0}$, and thus,

$$\mathbf{P}^{s_1 s_2} = \sum_{i \in \mathcal{I}^{s_1} \cap \mathcal{I}^{s_2}} \mathrm{E}\{\mathbf{B}_i^{s_1} \underline{\boldsymbol{n}}_i (\underline{\boldsymbol{n}}_i)^\top (\mathbf{B}_i^{s_2})^\top\} = \sum_{i \in \mathcal{I}^{s_1} \cap \mathcal{I}^{s_2}} \mathbf{B}_i^{s_1} \mathbf{P}_i^n (\mathbf{B}_i^{s_2})^\top \ .$$

□

Hence, estimators can store the square roots[1] of noise terms

$$\sqrt{\mathbf{P}_i^s} = \mathbf{B}_i^s \sqrt{\mathbf{P}_i^n} \ , i \in \mathcal{I}^s \ , \tag{3}$$

along with the respective noise index in sensor-specific sets $\mathcal{P}^s$, the (cross-)covariances can be reconstructed by means of Lemma 2. Note that the matrices $\sqrt{\mathbf{P}_i^s} \in \mathbb{R}^{n \times m_i}$ can be processed recursively in conformity with the linear transformations of the estimator so that matrices $\mathbf{B}_i^s$ describe the transformations from Lemma 1. When memory and communication capabilities at the sensors allow the storage of all relevant noise terms, (cross-)covariance terms are obtained exactly and optimal fusion methods [9]–[11] can be applied.

However, the number of stored matrices grows with every operation that involves noise terms. In particular, after the fusion of estimates, all noise terms stored in at least one of the participating sensors must be maintained subsequently as depicted in Fig. 2.

[1]For positive definite $\mathbf{P}_i^n$, the Cholesky decomposition can be used.

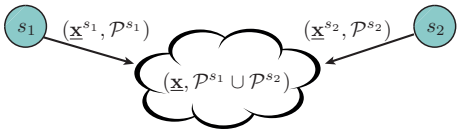

Figure 2. The fusion of estimates s_1 and s_2. As the error of the fused estimate is a composition of noise terms from $\hat{\underline{x}}^{s_1}$ and $\hat{\underline{x}}^{s_2}$, the union of noise term indices must be considered when the estimate is processed further.

This stands in stark contrast to (local) Kalman filtering where it is sufficient to store a tuple of estimate and covariance $(\hat{\underline{x}}^s, \mathbf{P}^s)$ at the sensors. The solution proposed in this paper is to find a middle ground, i.e., to process a small set of relevant noise terms individually and to aggregate the remaining noise quantities in a residual term. More precisely, each sensor processes the triple $(\hat{\underline{x}}^s, \mathcal{P}^s, \mathbf{R}^s)$ with the set of noise terms

$$\mathcal{P}^s = \{\sqrt{\mathbf{P}_i^s} \mid i \in \mathcal{I}_P^s \subseteq \mathcal{I}^s\} \tag{4}$$

and the residual

$$\mathbf{R}^s \geq \sum_{i \in \mathcal{I}^s \setminus \mathcal{I}_P^s} \sqrt{\mathbf{P}_i^s} (\sqrt{\mathbf{P}_i^s})^\top \ , \tag{5}$$

where $\mathcal{I}_P^s \subseteq \mathcal{I}^s$ denotes an arbitrary set of noise indices. Note that the residual is supposed to provide a bound in the positive definite sense and not necessarily the true value. For the sake of a simple notation, we define the part of the covariance that is obtained from stored noise terms as $||\mathcal{P}^s|| = \sum_{\sqrt{\mathbf{P}_i^s} \in \mathcal{P}^s} \sqrt{\mathbf{P}_i^s} (\sqrt{\mathbf{P}_i^s})^\top$.

In the following, we aim to provide covariance bounds by means of the individually stored noise and the residuals. To this end, we extend the notion of consistency from tuples $(\hat{\underline{x}}^s, \bar{\mathbf{P}}^s)$ to the quantities considered in this paper.

Definition 1 *The triple $(\hat{\underline{x}}^s, \mathcal{P}^s, \mathbf{R}^s)$ is **consistent** when $\sum_{i \in \mathcal{I}^s \setminus \mathcal{I}_P^s} \sqrt{\mathbf{P}_i^s} (\sqrt{\mathbf{P}_i^s})^\top \leq \mathbf{R}^s$ in the positive semi-definite sense.*

It directly follows from the definition of $\mathcal{P}^s$ that

$$\mathrm{E}\{(\hat{\underline{x}}^s - \underline{\boldsymbol{x}})^2\} = \mathrm{E}\{(\hat{\underline{e}}^s)^2\} \leq ||\mathcal{P}^s|| + \mathbf{R}^s \ . \tag{6}$$

Consistency as defined in Def. 1 pertains only to the local covariance and thus, is not violated when terms from $\mathcal{P}^s$ are summed up.

Proposition 1 *Let $(\hat{\underline{x}}^s, \mathcal{P}^s, \mathbf{R}^s)$ be consistent, then*

$$(\hat{\underline{x}}^s, \mathcal{P}^s \backslash \tilde{\mathcal{P}}^s, \mathbf{R}^s + ||\tilde{\mathcal{P}}^s||) \tag{7}$$

with $\tilde{\mathcal{P}}^s \subseteq \mathcal{P}^s$ is consistent.

Hence, the number of stores matrices in $\mathcal{P}^s$ can be reduced at any time by adding the discarded noise to the residual. However, as the composition of the error is stored in $\mathcal{P}^s$, information about the cross-covariances is lost when noise terms are aggregated in the residual.

The processing in the prediction and filter steps follows along the lines of the proof of Lemma 1 and results in the following formulas.

Proposition 2 *Let $(\hat{\underline{x}}^s, \mathcal{P}^s, \mathbf{R}^s)$ be consistent. Then*

$$(\mathbf{A}\hat{\underline{x}}^s, \mathbf{A}\mathcal{P}^s \cup \{\sqrt{\mathbf{C}^w}\}, \mathbf{A}\mathbf{R}^s(\mathbf{A})^\top) \text{ and} \tag{8}$$

$$(\mathbf{L}\hat{\underline{x}}^s + \mathbf{K}\underline{z}^s, \mathbf{L}\mathcal{P}^s \cup \{\sqrt{\mathbf{C}^{v_s}}\}, \mathbf{L}\mathbf{R}^s(\mathbf{L})^\top) \ , \tag{9}$$

with $\mathbf{L} = \mathbf{I} - \mathbf{K}\mathbf{H}$ and $\mathbf{T}\mathcal{P}^s := \{\mathbf{T}\sqrt{\mathbf{P}_i^s} \mid \sqrt{\mathbf{P}_i^s} \in \mathcal{P}^s\}$, $\mathbf{T} \in \mathbb{R}^{n\times n}$ are consistent estimates in prediction and filter step respectively.

PROOF. See Appendix. □

As exact cross-covariances are obtained only if all noise terms from the intersection $\mathcal{I}^{s_1} \cap \mathcal{I}^{s_2}$ are stored in both sensors (c.f., Lemma 2), in general, no optimal fusion algorithms under known correlations can be applied and even local covariances become inexact as the fused covariances depend on cross-covariances. Therefore, the unknown correlations between residuals must be bounded in the fusion operation. For this purpose, the following consistency lemma is utilized.

Lemma 3 (Joint Covariance Consistency) *Let ω^s denote weights with $\sum_{s\in\mathcal{S}'} \omega^s = 1$, $\omega^s \geq 0$, $\mathcal{S}' \subseteq \mathcal{S}$ and $\mathbf{P} \in \mathbb{S}_+^{S'n\times S'n}$ a positive definite (joint covariance) matrix. Then,*

$$\begin{bmatrix} \mathbf{P}^{s_1} & \dots & \mathbf{P}^{s_1 s_{S'}} \\ \vdots & \ddots & \vdots \\ \mathbf{P}^{s_{S'} s_1} & \dots & \mathbf{P}^{s_{S'}} \end{bmatrix} \leq \begin{bmatrix} \frac{1}{\omega^{s_1}}\mathbf{P}^{s_1} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \frac{1}{\omega^{s_{S'}}}\mathbf{P}^{s_{S'}} \end{bmatrix}$$

in the positive semi-definite sense.

PROOF. See Appendix. □

Note that Lemma 3 is a general result that applies to basic CI literature as well. Let $\mathbf{P}$ denote the true joint covariance. It is well-known that the true covariance of the fused estimate which is obtained with gains $\mathbf{F} = (\mathbf{F}^{s_1} \quad \dots \quad \mathbf{F}^{s_{S'}}) \in \mathbb{R}^{n\times S'n}$ is given by $\mathbf{P}^f = \mathbf{F}\mathbf{P}(\mathbf{F})^\top$. For joint covariance bounds $\bar{\mathbf{P}}$ it follows with basic linear algebra

$$\bar{\mathbf{P}} \geq \mathbf{P} \Rightarrow \mathbf{F}\bar{\mathbf{P}}(\mathbf{F})^\top \geq \mathbf{F}\mathbf{P}(\mathbf{F})^\top \ . \tag{10}$$

Therefore, it is possible to obtain covariance bounds for the fused estimate by means of joint covariance bounds. The opposite direction of (10) is not true as $\mathbf{F}$ is singular and thus, Lemma 3 provides bounds that are more general than bounds that apply only to the fused estimate. This and its applicability to $S' > 2$ estimates make the statement slightly stronger than the consistency statements from [15], [17], [21], [25]. The convex combination utilized in CI is a special case with $\mathbf{F}^s = \left(\sum_{\tilde{s}\in\mathcal{S}'} (\mathbf{P}^{\tilde{s}})^{-1}\omega^{\tilde{s}}\right)^{-1}(\mathbf{P}^s)^{-1}\omega^s$, $s \in \mathcal{S}' \subseteq \mathcal{S}$.

By means of Lemma 3, we ensure consistency in the fusion step.

Theorem 1 *Let $(\hat{\underline{x}}^s, \mathcal{P}^s, \mathbf{R}^s)$ be consistent for $s \in \mathcal{S}' \subseteq \mathcal{S}$. We define the sets $\mathcal{P}_\cap^s = \{\sqrt{\mathbf{P}_i^s} \mid \sqrt{\mathbf{P}_i^s} \in \mathcal{P}^s \ , i \in \bigcap_{s\in\mathcal{S}'} \mathcal{I}_P^s\} \subseteq \mathcal{P}^s$, which comprise those noise terms that are in the intersection of the stored terms. Then, $(\hat{\underline{x}}^f, \mathcal{P}^f, \mathbf{R}^f)$ with*

$$\begin{aligned} \hat{\underline{x}}^f &= \sum_{s\in\mathcal{S}'} \mathbf{F}^s\hat{\underline{x}}^s \\ \mathcal{P}^f &= \Big\{ \sum_{s\in\mathcal{S}'} \mathbf{F}^s\sqrt{\mathbf{P}_i^s} \mid \sqrt{\mathbf{P}_i^s} \in \mathcal{P}_\cap^s \Big\} \\ \mathbf{R}^f &= \sum_{s\in\mathcal{S}'} (\omega^s)^{-1}\mathbf{F}^s\big(\mathbf{R}^s + ||\mathcal{P}^s\backslash\mathcal{P}_\cap^s||\big)(\mathbf{F}^s)^\top \ , \end{aligned} \tag{11}$$

is consistent for $\omega^s > 0$, $\sum_{s\in\mathcal{S}'} \omega^s = 1$ and fusion gains $\mathbf{F}^s$ with $\sum_{s\in\mathcal{S}'} \mathbf{F}^s = \mathbf{I}$.

PROOF. See Appendix. □

With Propositions 1 and 2 and Theorem 1, it is feasible to obtain covariance bounds when the estimator processing, i.e., the calculation procedure for filter and fusion gains, is fixed or constant. The technique can for example be utilized to obtain bounds for known estimators such as local Kalman filters, the consensus Kalman filter [5], [26], and many others. In contrast to the bounds that are provided by other CI generalizations [18]–[20], the fused noise decomposition consists of an explicitly known part $\mathcal{P}^f$ that can be exploited in subsequent fusion operations. Also, the precision, i.e., the difference between true covariance and bound in (6), can be tuned by means of the number of matrices stored in $\mathcal{P}^s$.

IV. Consistent Estimation in Sensor Networks

In this section, we utilize the bounding tools from Sec. III to derive an efficient estimation algorithm for sensor networks. To this end, we discuss which noise terms should be stored, derive the optimal fusion gains for the bound specified in Theorem 1, and propose a filter processing for the local sensors.

For a meaningful application, it is necessary to select the noise terms to be individually stored in $\mathcal{P}^s$. Fortunately, for stable estimators, the eigenvalues of combined prediction, filter, and fusion matrices $\mathbf{F}^s(\mathbf{I} - \mathbf{K}^s\mathbf{H}^s)\mathbf{A}$ are not larger than one and so, the influence of individual noise terms fades with time. As a consequence, only a subset of $\mathcal{I}^s$ consisting of the latest noise terms determines the (cross-)covariances.

The precision of the bound in Theorem 1 depends on which common noise terms are stored at the sensors $s \in \mathcal{S}'$, i.e., the intersection $\bigcap_{s\in\mathcal{S}'} \mathcal{I}_P^s$. The more noise terms are in the intersection, the better the reconstruction precision as only the remaining terms are bounded.

We propose to store all noise terms that are added in prediction, filter, and fusion steps in $\mathcal{P}^s$ and to aggregate only those terms in the residual that date back τ or more time steps. Therefore, a noise aggregation step consists of the summation of entries according to Proposition 1 with

$$\tilde{\mathcal{P}}^s = \{\sqrt{\mathbf{P}_i^s} \in \mathcal{P}^s \mid \underline{\boldsymbol{n}}_i = \underline{\boldsymbol{w}}_t \vee \underline{\boldsymbol{n}}_i = \underline{\boldsymbol{v}}_t^{\tilde{s}} \ , t \leq k-\tau, \tilde{s} \in \mathcal{S}\} \ . \tag{12}$$

Then, sensors have the implicit knowledge that all noise terms from time steps $[k-\tau+1, k]$ are either stored in $\mathcal{P}^s$ or do not affect the estimate of sensor s, i.e., $\mathbf{B}_i^s = \mathbf{0}$ in Lemma 1. This property can be exploited in Theorem 1 by setting

$$\mathcal{I}_P^s = \{i \in \mathcal{I} \mid \underline{\boldsymbol{n}}_i = \underline{\boldsymbol{w}}_t \vee \underline{\boldsymbol{n}}_i = \underline{\boldsymbol{v}}_t^{\tilde{s}} \ , t > k-\tau \ , \tilde{s} \in \mathcal{S}\}$$

with $\sqrt{\mathbf{P}_i^s} = \mathbf{0}$ if no term is available in $\mathcal{P}^s$. By this means, the joint covariance for noise terms from the period $[k-\tau+1, k]$ is reconstructed exactly and older noise terms are bounded in the residual $\mathbf{R}^s$.

However, typically, fusion gains $\mathbf{F}$ are calculated so that the MSE of the fused estimate is minimized. We define the joint covariance for the intersection of noise terms as

$$\mathbf{P}_\cap = \sum_{i \in \bigcap \mathcal{I}_P^s} \begin{bmatrix} \sqrt{\mathbf{P}_i^{s_1}} \\ \vdots \\ \sqrt{\mathbf{P}_i^{s_{S'}}} \end{bmatrix} \begin{bmatrix} \sqrt{\mathbf{P}_i^{s_1}}^\top & \dots & \sqrt{\mathbf{P}_i^{s_{S'}}}^\top \end{bmatrix}$$

and the joint covariance for the remaining terms as the block diagonal matrix $\mathbf{P}_\omega$ with elements $(\omega^s)^{-1}\mathbf{R}^s_{\setminus\cap}$. Then, the challenge boils down to finding optimal fusion gains for a joint covariance matrix that is the sum of two independent parts. Hence, the following theorem also provides the (missing) fusion method for the problem setting considered in [19], [20] for an arbitrary number of sensors.

Theorem 2 *Let the assumptions be the same as in Theorem 1 and $\mathbf{P}_\cap + \mathbf{P}_\omega$ be positive definite. The fusion gains that optimize the trace of the bound* (6) *provided by Theorem 1 are given by*

$$\mathbf{F} = \left(\underline{\mathbf{1}}^\top \left(\mathbf{P}_\cap + \mathbf{P}_\omega\right)^{-1} \underline{\mathbf{1}}\right)^{-1} \underline{\mathbf{1}}^\top \left(\mathbf{P}_\cap + \mathbf{P}_\omega\right)^{-1} , \tag{13}$$

with $\mathbf{F} = \begin{pmatrix} \mathbf{F}^{s_1} & \dots & \mathbf{F}^{s_{S'}} \end{pmatrix}$ *subject to*

$$\min_{\omega^s, s\in\mathcal{S}'} \operatorname{tr}\{\left(\underline{\mathbf{1}}^\top \left(\mathbf{P}_\cap + \mathbf{P}_\omega\right)^{-1} \underline{\mathbf{1}}\right)^{-1}\}, \sum_{s\in\mathcal{S}'} \omega^s = 1 \ , \omega^s > 0 \ . \tag{14}$$

PROOF. See Appendix. □

When the joint covariance bound $\mathbf{P}_\cap + \mathbf{P}_\omega$ is singular, the optimal fusion gains are obtained by replacing the inverses with pseudoinverses [11]. As a special case of Theorem 2 we obtain the CI optimization with $\mathbf{P}_\cap = \mathbf{0}$. When $\mathbf{P}_\cap \geq \mathbf{0}$, the choice of ω affects a smaller part of the overall covariance bound compared to CI, i.e., only the residual, so that the computational effort of a numerical optimization is not always pertinent. Therefore, we propose to use the simple procedure from [25] that finds an approximate solution for the weighting factors according to

$$\omega^s = \frac{1/\operatorname{tr}\{\mathbf{R}^s + ||\mathcal{P}^s \setminus \mathcal{P}^s_\cap||\}}{\sum_{\tilde{s}\in\mathcal{S}'} 1/\operatorname{tr}\{(\mathbf{R}^{\tilde{s}} + ||\mathcal{P}^{\tilde{s}} \setminus \mathcal{P}^{\tilde{s}}_\cap||)\}} \ . \tag{15}$$

Then, (13) can be solved in closed form.

The result of the bounding procedure from Theorem 1 with fusion gains (13) for fixed $\omega^{s_1} = 0.3$ and different degrees of stored noise information are depicted in Fig. 3. When no correlation is stored, the CI result is obtained. With additional information about correlations, the bound becomes tighter with respect to the corresponding true covariance. Indeed, this insight also transfers to the fusion gains. Due to the optimization with partially or fully known cross-covariances, the true covariance shrinks with increasing information and reaches the optimal fusion result under exactly known correlations.

In order to pursue decentralized estimation, it is moreover necessary to define the filter processing at the sensors. The proposed technique is compatible with any forms of unbiased filters that provide linear transformation matrices. In particular, it is viable to apply local KFs with

$$\mathbf{K}_k^s = \left(\left(\mathbf{P}_k^s\right)^{-1} + \left(\mathbf{H}_k^s\right)^\top \left(\mathbf{C}_k^{v_s}\right)^{-1} \mathbf{H}_k^s\right)^{-1} \left(\mathbf{H}_k^s\right)^\top \left(\mathbf{C}_k^{v_s}\right)^{-1}, \tag{16}$$

and $\mathbf{L}_k^s = (\mathbf{I} - \mathbf{K}_k^s \mathbf{H}_k^s)$.

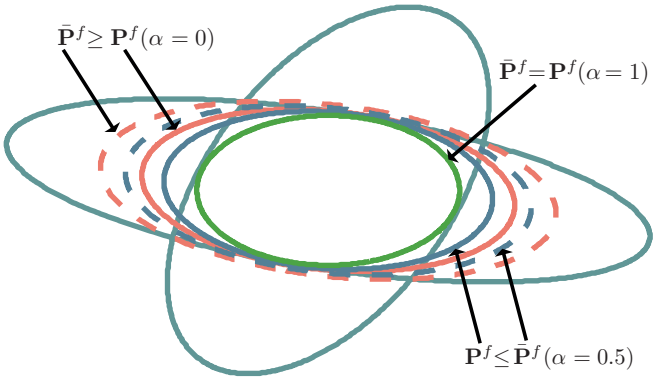

Figure 3. The covariance ellipsoids for the true input covariances $\mathbf{P}^{s_1}$ and $\mathbf{P}^{s_2}$ (solid turquoise), the true fused covariances (solid), and the covariance bounds $\bar{\mathbf{P}}^f$ (dashed) for different degrees of information about correlations are given for an illustrative example. The parameter α^f denotes the relation between known and unknown correlations.

The necessary processing of an estimation algorithm that employs the composition technique is summarized in Alg. 1. The main difference to a standard linear estimator is that instead of the tuple $(\hat{\underline{x}}_k^s, \mathbf{P}_k^s)$ the triple $\mathcal{X}_k^s := (\hat{\underline{x}}_k^s, \mathcal{P}_k^s, \mathbf{R}_k^s)$ is maintained at the sensors.

Algorithm 1 Sensor Processing

1: Initialize $\mathcal{X}_1^s = (\hat{\underline{x}}_1^s, \{\mathbf{P}_1^s\}, \mathbf{0})$
2: **for** $k = 2 \ ; \ \dots \ ; \ k = k+1$ **do**
3: Predict $\mathcal{X}_{k-1}^s$ with (8)
4: Calculate filter gains, e.g., with (16)
5: Filter $\mathcal{X}_k^s$ with (9)
6: Exchange $\mathcal{X}_{k|k}^s$ with all neighbors
7: Calculate fusion gains, e.g., with (13)
8: Fuse $\mathcal{X}_{k|k}^{s_i}$ with (11)
9: Aggregate $\mathcal{X}_k^s$ with (7), $\mathcal{B}^s$ from (12)
10: **end for**

In order to cope with distinguishing challenges in communication and computation, it is feasible to extract a small set $\tilde{\mathcal{X}}_k^s$ from $\mathcal{X}_k^s$ by means of (7) and exchange only the smaller set while concurrently storing the original $\mathcal{X}_k^s$ for further processing.

In summary, the noise decomposition is easily integrated in decentralized estimation algorithms and allows for the application of an efficient fusion step. The advantages of maintaining cross-covariances explicitly are demonstrated in the next section.

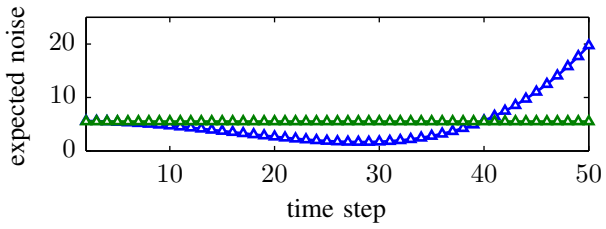

Figure 4. The average expected measurement noise in blue and the time average of the quantity in green.

V. Evaluation

We examine the performance in a one-dimensional (constant acceleration) white noise jerk model [27] for 100 randomly generated communication patterns. Let three sensors be positioned at $p^s = s$, $s = 1, \dots, 3$ observing the position of a target with initial state vector $\underline{\boldsymbol{x}}_1 = (0 \quad 0 \quad 0.5)^\top$, where the states denote (position, velocity, acceleration). We obtain $\mathbf{H}^s = [1 \quad 0 \quad 0]$ as measurement matrix while the measurement noise $\mathbf{C}_k^{v_s} = ([\underline{\boldsymbol{x}}_k]_1 - p^s)^2 + 1$ depends quadratically on the difference between target and sensor position p^s. The initial estimates are assumed to be independent with uncertainty $\mathbf{P}_1^s = 3 \cdot \mathbf{I}$. A communication between two sensors is established with probability 0.3.

The state model parameters are given by

$$\mathbf{A} = \begin{bmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{C}^w = q \cdot \begin{bmatrix} t^5/20 & t^4/8 & t^3/6 \\ t^4/8 & t^3/3 & t^2/2 \\ t^3/6 & t^2/2 & t \end{bmatrix}.$$

For the evaluation we set $q = 1$ and $t = 0.1$.

In order to allow an easy comparison of different runs, the measurement noises are approximated based on a deterministic movement of the target with initial state $\underline{\boldsymbol{x}}_1$. The expected average measurement noise per sensor $\bar{\mathbf{C}}^v = \frac{1}{\mu} \sum_{s \in \mathcal{S}} \mathbf{C}_k^{v_s}$ with $\mu^s = 1 + 0.3 \cdot (3 - 1) = 1.6$ is depicted in Fig. 4 for different time steps.

We evaluate the proposed noise decomposition approach (ND) from Sec. IV for different horizons τ. Note that for $\tau = 0$, ND corresponds to CI so that the improvements due to the proposed generalization are demonstrated as well. The filter gains are obtained with (16) where $\mathbf{P}_k^s$ is substituted with the provided covariance bound. The fusion gains are optimized with Theorem 2 where ω^s are obtained with the simple approximation in (15).

As a baseline, we use KFs that calculate filter gains (16) and optimal fusion gains [10] by means of the true covariances (OPT). The true covariances are actually unknown to the sensors but are supposed to be available to the baseline estimator in order to provide an achievable performance in linear decentralized estimation.

In Fig. 5, the covariance bounds after the fusion step of CI, OPT, and ND for time horizons $\tau = 1, 3$ are depicted. After the period in the beginning that is distorted by the uncertainties of the initial estimates, the covariance bounds of all estimators follow the shape of the average noise depicted in Fig. 4. The bounds derived by ND decline with the considered time horizon τ. For $\tau = 0$, i.e., CI, the bound is on average over all time

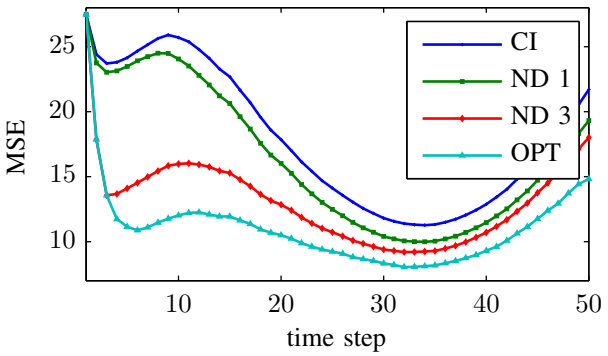

Figure 5. The covariance bounds for CI, the noise decomposition (ND) approach with time horizons 1 and 3, and for the baseline (OPT).

steps about 62 percent larger than the true MSE of the baseline while for $\tau = 3$, the bound is on average only 19 percent worse than the baseline. Additional increases of τ (not depicted in Fig. 5) improve the bound only marginally until eventually for $\tau = 50$ the baseline results are yielded. This agrees with the intuition that noise terms carry the more significant information the newer they are.

The number of scalars to transmit for ND depends on the communication probability and the dimension of the measurement noise. In the present scenario, the stored measurement noise matrices have dimension 3×1. The average number of scalars to communicate for the calculation of the bound is 12 for $\tau = 1$ and 40.6 for $\tau = 3$ compared to 9 for CI[2].

In summary, the noise decomposition allows the reconstruction of the most important dependencies between estimates and therefore, considerably improves the true MSE and the precision of the bound compared to CI. Usually, it suffices to individually store noise terms of a small horizon τ to achieve the demonstrated improvements, which makes the proposed processing efficiently applicable in realistic sensor network contexts.

VI. Conclusions

In this paper, we tackled the problem of providing covariance bounds in linear decentralized estimation. By means of the derived techniques, sensors are enabled to consistently filter measurements and fuse estimates based on locally available model information only.

The key idea is to partially reconstruct cross-covariances based on a decomposition of errors into their noise terms and to combine this technique with a generalized version of Covariance Intersection in order to provide consistent estimates. The precision of the resulting bound is scalable in the number of stored noise terms.

When all (relevant) noise terms are considered individually, the exact cross-covariances are reconstructed. The other extreme is to store only the aggregation of noise terms in a residual. Then, the (conservative) Covariance Intersection bound is obtained. We propose to store a selection of individual noise

[2]For the sake of simplicity, we ignored possible compression techniques, e.g., by exploiting the symmetry of covariances.

terms and bound only the remaining terms. By this means, the relevant dependencies are reconstructed exactly and almost optimal fusion gains can be derived.

Future research will aim to develop more sophisticated techniques for the selection of noise terms that are stored individually and to reduce the data amount that must be exchanged in the fusion step.

ACKNOWLEDGMENT

This research was supported in part by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-0939370, by the Karlsruhe House of Young Scientists (KHYS), and by support from Deutsche Telekom T-Labs.

APPENDIX

A. PROOF OF PROPOSITION 2.

The true covariance of $\mathbf{A}\hat{\underline{x}}^s - \mathbf{A}\underline{\boldsymbol{x}} - \underline{\boldsymbol{w}}$ is given by

$$\mathbf{A}\,\mathrm{E}\{(\hat{\underline{x}}^s - \underline{\boldsymbol{x}})^2\}(\mathbf{A})^\top + \mathbf{C}^w \ .$$

With $\mathbf{P}^1 \leq \mathbf{P}^2 \Rightarrow \mathbf{M}\mathbf{P}^1(\mathbf{M})^\top \leq \mathbf{M}\mathbf{P}^2(\mathbf{M})^\top$ for positive semi-definite $\mathbf{P}^i \in \mathbb{S}^{m_i \times m_i}$ and $\mathbf{M} \in \mathbb{R}^{n\times n}$, it follows

$$\mathbf{A}\,\mathrm{E}\{(\hat{\underline{x}}^s - \underline{\boldsymbol{x}})^2\}(\mathbf{A})^\top + \mathbf{C}^w \leq \mathbf{A}||\mathcal{P}^s||(\mathbf{A})^\top + \mathbf{A}\mathbf{R}^s(\mathbf{A})^\top + \mathbf{C}^w = ||\mathbf{A}\mathcal{P}^s|| + \sqrt{\mathbf{C}^w}(\sqrt{\mathbf{C}^w})^\top + \mathbf{A}\mathbf{R}^s(\mathbf{A})^\top \ ,$$

which proves consistency of the prediction step processing.

It holds $\mathbf{L}\hat{\underline{x}}^s - \mathbf{K}\underline{z}^s - \underline{\boldsymbol{x}} = \mathbf{L}(\hat{\underline{x}}^s - \underline{\boldsymbol{x}}) + \mathbf{K}\underline{\boldsymbol{v}}^s$ and thus the true covariance is given by

$$\mathbf{L}\,\mathrm{E}\{(\hat{\underline{x}}^s - \underline{\boldsymbol{x}})^2\}(\mathbf{L})^\top + \mathbf{K}\mathbf{C}^{v_s}(\mathbf{K})^\top \ .$$

The remaining part of the proof follows along the lines of the prediction step. □

B. PROOF OF LEMMA 3.

Let $\mathbf{P}_k$ and $\bar{\mathbf{P}}_k$ denote the joint covariance matrix and its bound of the first $k \leq S'$ entries

$$\mathbf{P}_k = \begin{bmatrix} \mathbf{P}^1 & \dots & \mathbf{P}^{1k} \\ \vdots & \ddots & \vdots \\ \mathbf{P}^{k1} & \dots & \mathbf{P}^k \end{bmatrix} \text{ and } \bar{\mathbf{P}}_k = \begin{bmatrix} \frac{1}{\omega^1}\mathbf{P}^1 & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \frac{1}{\omega^k}\mathbf{P}^k \end{bmatrix} \ .$$

We show by induction that $\bar{\omega}_k\bar{\mathbf{P}}_k \geq \mathbf{P}_k$ with $\bar{\omega}_k = \sum_{s=1}^k \omega^s$. Then, the claim follows immediately with $\bar{\omega}_{S'} = 1$.

For $k = 1$, we have $\bar{\omega}_1\bar{\mathbf{P}}_1 = \frac{\omega^1}{\omega^1}\mathbf{P}^1 \geq \mathbf{P}_1$. Let us assume that the claim holds for $k-1$. Then,

$$\bar{\omega}_k\bar{\mathbf{P}}_k - \mathbf{P}_k = \bar{\omega}_k \begin{bmatrix} \frac{1}{\bar{\omega}_{k-1}}\bar{\omega}_{k-1}\bar{\mathbf{P}}_{k-1} & \mathbf{0} \\ \mathbf{0} & \frac{1}{\omega^k}\mathbf{P}^k \end{bmatrix} - \mathbf{P}_k = \begin{bmatrix} \frac{\bar{\omega}_k}{\bar{\omega}_{k-1}}\bar{\omega}_{k-1}\bar{\mathbf{P}}_{k-1} - \mathbf{P}_{k-1} & -\mathbf{P}^{(k-1)k} \\ -\mathbf{P}^{k(k-1)} & \frac{\bar{\omega}_k}{\omega^k}\mathbf{P}^k - \mathbf{P}^k \end{bmatrix} \ ,$$

where $\mathbf{P}^{(k-1)k}$ denotes the corresponding cross-covariances. According to the induction hypothesis, it holds $\bar{\omega}_{k-1}\bar{\mathbf{P}}_{k-1} \geq \mathbf{P}_{k-1}$ and thus, we obtain with $\frac{\bar{\omega}_k}{\bar{\omega}_{k-1}} \geq 0$

$$\bar{\omega}_k\bar{\mathbf{P}}_k - \mathbf{P}_k \geq \begin{bmatrix} \frac{\bar{\omega}_k - \bar{\omega}_{k-1}}{\bar{\omega}_{k-1}}\mathbf{P}_{k-1} & -\mathbf{P}^{(k-1)k} \\ -\mathbf{P}^{k(k-1)} & \frac{\bar{\omega}_k - \omega^k}{\omega^k}\mathbf{P}^k \end{bmatrix} \ .$$

According to Theorem 7.7.6 from [28], the covariance difference is positive definite if and only if

$$(\mathbf{P}^{k(k-1)})^\top \frac{\bar{\omega}_{k-1}}{\bar{\omega}_k - \bar{\omega}_{k-1}}(\mathbf{P}_{k-1})^{-1}\mathbf{P}^{(k-1)k} < \frac{\bar{\omega}_k - \omega^k}{\omega^k}\mathbf{P}^k \ .$$

With $\bar{\omega}_k = \bar{\omega}_{k-1} + \omega^k$, we obtain $\frac{\bar{\omega}_{k-1}}{\bar{\omega}_k - \bar{\omega}_{k-1}} = \frac{\bar{\omega}_{k-1}}{\omega^k} = \frac{\bar{\omega}_k - \omega^k}{\omega^k}$ and thus, $\bar{\omega}_k\bar{\mathbf{P}}_k \geq \mathbf{P}_k$ for positive definite $\mathbf{P}^k$. □

C. PROOF OF THEOREM 1.

From Lemma 1 we know that errors are linear combinations of noise terms. In the following, we bound the joint covariance

$$\mathbf{P} = \sum_{i\in\mathcal{I}} \begin{bmatrix} \mathbf{B}_i^{s_1}\mathbf{P}_i^n(\mathbf{B}_i^{s_1})^\top & \dots & \mathbf{B}_i^{s_1}\mathbf{P}_i^n(\mathbf{B}_i^{s_{S'}})^\top \\ \vdots & \ddots & \vdots \\ \mathbf{B}_i^{s_{S'}}\mathbf{P}_i^n(\mathbf{B}_i^{s_1})^\top & \dots & \mathbf{B}_i^{s_{S'}}\mathbf{P}_i^n(\mathbf{B}_i^{s_{S'}})^\top \end{bmatrix} \ ,$$

with $\mathbf{B}_i^{s_i} = \mathbf{0}$ for $i \notin \mathcal{I}^{s_i}$.

We represent the covariances as sums of independent components each. Let $\mathbf{P}^s$ denote the true covariance of sensor s. It follows from (6)

$$\mathbf{P}^s \leq ||\mathcal{P}^s_\cap|| + ||\mathcal{P}^s \backslash \mathcal{P}^s_\cap|| + \mathbf{R}^s \ . \tag{17}$$

For noise terms collected in the first term of the sum, the cross-covariances are given by Lemma 2. We define the respective joint covariance as $\mathbf{P}_\cap$ and note that $\mathbf{P}_\cap$ depicts the sum terms of $\mathbf{P}$ with $i \in \bigcap \mathcal{I}^s_P$.

As different sensors can represent different noise terms in the residuals, the cross-covariances between the latter two terms in (17) are unknown. However, it holds $\mathbf{R}^s \geq \sum_{i\in\mathcal{I}^s\backslash\mathcal{I}^s_P} \sqrt{\mathbf{P}_i^{s_i}}(\sqrt{\mathbf{P}_i^{s_i}})^\top = \sum_{i\in\mathcal{I}\backslash\mathcal{I}^s_P} \sqrt{\mathbf{P}_i^{s_i}}(\sqrt{\mathbf{P}_i^{s_i}})^\top$ and thus, $||\mathcal{P}^s\backslash\mathcal{P}^s_\cap|| + \mathbf{R}^s$ bound the block diagonal matrices of $\mathbf{P} - \mathbf{P}_\cap$. We define the matrix

$$\mathbf{P}_\omega = \begin{bmatrix} \frac{1}{\omega^{s_1}}\mathbf{R}^{s_1}_{\backslash\cap} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \frac{1}{\omega^{s_{S'}}}\mathbf{R}^{s_{S'}}_{\backslash\cap} \end{bmatrix} \ ,$$

with $\mathbf{R}^s_{\backslash\cap} = ||\mathcal{P}^s\backslash\mathcal{P}^s_\cap|| + \mathbf{R}^s$, which is a joint covariance bound according to Lemma 3. Therefore $\mathbf{P}_\cap + \mathbf{P}_\omega \geq \mathbf{P}$ and consistency of the fused covariance follows with (10). □

D. PROOF OF THEOREM 2.

The bound provided by Theorem 1 is the result of a fusion operation with joint covariance $\bar{\mathbf{P}} = \mathbf{P}_\cap + \mathbf{P}_\omega$. The fusion gains that provide the smallest fused covariance in the positive definite order have been derived for $S = 2$ estimates [9] and for $S \geq 2$ estimates [10]. An application of the formulas from [10] leads to the fused covariance bound

$$\bar{\mathbf{P}}^f = \left(\underline{\mathbf{1}}^\top\left(\mathbf{P}_\cap + \mathbf{P}_\omega\right)^{-1}\underline{\mathbf{1}}\right)^{-1}$$

and the optimal fusion gains as stated in the Theorem. The given optimization problem for ω^s derives the trace optimal weights. □

REFERENCES

[1] M. Proetzsch, T. Luksch, and K. Berns, "Development of Complex Robotic Systems using the Behavior-based Control Architecture iB2C," Robotics and Autonomous Systems, vol. 58, no. 1, pp. 46–67, 2010.

[2] R. E. Kalman and Others, "A New Approach to Linear Filtering and Prediction Problems," Journal of Basic Engineering, vol. 82, no. 1, pp. 35–45, 1960.

[3] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman Filtering with Intermittent Observations," IEEE Transactions on Automatic Control, vol. 49, no. 9, pp. 1453–1464, 2004.

[4] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of Control and Estimation over Lossy Networks," Proceedings of the IEEE, vol. 95, no. 1, pp. 163–187, 2007.

[5] R. Olfati-Saber, "Kalman-Consensus Filter: Optimality, Stability, and Performance," in Proceedings of the 48th IEEE Conference on Decision and Control (CDC 2009), 2009, pp. 7036–7042.

[6] F. S. Cattivelli, S. Member, and A. H. Sayed, "Diffusion Strategies for Distributed Kalman Filtering and Smoothing," IEEE Transactions on Automatic Control, vol. 55, no. 9, pp. 2069–2084, 2010.

[7] N. A. Carlson, "Federated Square Root Filter for Decentralized Parallel Processors," IEEE Transactions on Aerospace and Electronic Systems, vol. 26, no. 3, pp. 517–525, 1990.

[8] F. Govaers and W. Koch, "An Exact Solution to Track-to-Track-Fusion at Arbitrary Communication Rates," IEEE Transactions on Aerospace and Electronic Systems, vol. 48, no. 3, 2012.

[9] Y. Bar-Shalom and L. Campo, "The Effect of the Common Process Noise on the Two-Sensor Fused-Track Covariance," IEEE Transactions on Aerospace and Electronic Systems, vol. 22, no. 6, pp. 803–805, Nov. 1986.

[10] S.-l. Sun, "Multi-sensor optimal information fusion kalman filters with applications," Aerospace Science and Technology, vol. 8, no. 1, pp. 57–62, 2004.

[11] J. Zhou, Y. Zhu, Z. You, and E. Song, "An Efficient Algorithm for Optimal Linear Estimation Fusion in Distributed Multisensor Systems," IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, vol. 36, no. 5, pp. 1000–1009, 2006.

[12] 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), 2001.

[13] K. H. Kim, "Development of Track to Track Fusion Algorithms," in Proceedings of the American Control Conference 1994, vol. 1. IEEE, 1994, pp. 1037–1041.

[14] 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.

[15] S. J. Julier and J. K. Uhlmann, "A Non-divergent Estimation Algorithm in the Presence of Unknown Correlations," in Proceedings of the American Control Conference 1997, vol. 4, 1997, pp. 2369–2373.

[16] 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.

[17] L. Chen, P. O. Arambel, and R. K. Mehra, "Fusion under Unknown Correlation - Covariance Intersection as a Special Case," in Proceedings of the 5th International Conference on Information Fusion (FUSION 2002), Annapolis, Maryland, USA, 2002.

[18] S. J. Julier and J. K. Uhlmann, "Simultaneous Localisation and Map Building Using Split Covariance Intersection," in Proceedings of the International Conference on Intelligent Robots and Systems (IROS 2001), vol. 3, 2001, pp. 1257–1262.

[19] U. D. Hanebeck and J. Horn, "An Efficient Method for Simultaneous Map Building and Localization," in Proceedings of SPIE, Vol. 4385, AeroSense Symposium, Orlando, Florida, USA, Apr. 2001.

[20] B. Noack, M. Baum, and U. D. Hanebeck, "Automatic Exploitation of Independencies for Covariance Bounding in Fully Decentralized Estimation," in Proceedings of the 18th IFAC World Congress (IFAC 2011), Milan, Italy, Aug. 2011.

[21] F. C. Schweppe, Uncertain Dynamic Systems. Prentice-Hall Englewood Cliffs, 1973, vol. 160.

[22] J. Hu, L. Xie, and C. Zhang, "Diffusion Kalman Filtering based on Covariance Intersection," IEEE Transactions on Signal Processing, vol. 60, no. 2, pp. 891–902, 2012.

[23] P. O. Arambel, "Structured Pedigree Information for Distributed Fusion Systems," in SPIE Signal Processing, Sensor Fusion, and Target Recognition XVI. Society of Photo-Optical Instrumentation Engineers, 2008, pp. 69 680X–1.

[24] J. K. Uhlmann, "Covariance Consistency Methods for Fault-tolerant Distributed Data Fusion," Information Fusion, vol. 4, no. 3, pp. 201–215, 2003.

[25] W. Niehsen, "Information Fusion Based on Fast Covariance Intersection Filtering," in Proceedings of the 5th International Conference on Information Fusion (FUSION 2002), vol. 2. IEEE, 2002, pp. 901–904.

[26] R. Olfati-Saber, "Distributed Kalman Filter with Embedded Consensus Filters," in Proceedings of the 44th IEEE Conference on Decision and Control (CDC-ECC 2005). IEEE, 2005, pp. 8179–8184.

[27] R. Li and V. P. Jilkov, "Survey of Maneuvering Target Tracking. Part I. Dynamic Models," IEEE Transactions on Aerospace and Electronic Systems, vol. 39, no. 4, pp. 1333–1364, 2003.

[28] R. A. Horn and C. R. Johnson, Matrix Analysis. Cambridge University Press, 2005.