Fusion Strategies for Unequal State Vectors in Distributed Kalman Filtering

Benjamin Noack [*], **Joris Sijs** [**], **and Uwe D. Hanebeck** [*]

[*] *Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics, Karlsruhe Institute of Technology (KIT), Germany. (e-mail: benjamin.noack@ieee.org, uwe.hanebeck@ieee.org)*
[**] *TNO Technical Sciences, Den Haag, The Netherlands. (e-mail: joris.sijs@tno.nl)*

Abstract: Distributed implementations of state estimation algorithms generally have in common that each node in a networked system computes an estimate on the entire global state. Accordingly, each node has to store and compute an estimate of the same state vector irrespective of whether its sensors can only observe a small part of it. In particular, the task of monitoring large-scale phenomena renders such distributed estimation approaches impractical due to the sheer size of the corresponding state vector. In order to reduce the workload of the nodes, the state vector to be estimated is subdivided into smaller, possibly overlapping parts. In this situation, fusion does not only refer to the computation of an improved estimate but also to the task of reassembling an estimate for the entire state from the locally computed estimates of unequal state vectors. However, existing fusion methods require equal state representations and, hence, cannot be employed. For that reason, a fusion strategy for estimates of unequal and possibly overlapping state vectors is derived that minimizes the mean squared estimation error. For the situation of unknown cross-correlations between local estimation errors, also a conservative fusion strategy is proposed.

Keywords: Estimation theory, filtering techniques, Kalman filters, distributed estimation, data fusion

1. INTRODUCTION

State estimation methods such as Kalman filters (Kalman, 1960) are applied for the purpose of providing useful information about an unknown quantity. An estimate is dynamically computed based on prior information, a process model, and measurements coming from sensor devices.

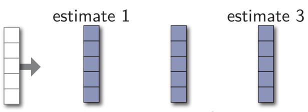


Fig. 1. Local estimates.

In many network-based applications, it is often not a single instance that computes an estimate but many autonomous nodes each of which is equipped with its own state estimation system. Networked estimation systems offer several advantages over monolithic systems including scalability, robustness to failures, and distribution of computational resources. Some situations simply require the use of sensor networks, e.g., in order to monitor a large-scale phenomenon. Collecting all measurements at a central system is in general not an option due to high requirements regarding frequency and volume of data transfers. Instead, distributed state estimation approaches (Liggins II et al., 2009) are the method of choice, where each node computes an estimate based on its sensor readings. State estimation then has to address the additional question of how to fuse estimates stemming from different sources in order to obtain a single informative estimate.

Distributed estimation principles are in the focus of many studies—see, for instance, Hall et al. (2013), Olfati-Saber (2007), Carli et al. (2008), and Govaers and Koch (2010). The most significant contributions have been achieved in the context of target tracking applications (Bar-Shalom and Campo, 1986; Bar-Shalom et al., 2001; Chong et al., 2000), where different tracks of the same target are to be fused. A continuing challenge in developing distributed tracking algorithms is the treatment of cross-correlations between the tracks to be fused (Chang et al., 1997). Since bookkeeping of cross-correlations between all local estimators is often too cumbersome, conservative fusion strategies (Julier and Uhlmann, 1997; Noack et al., 2011) can be pursued. If fully decentralized estimation architectures are aspired, it is even not possible to keep track of correlations.

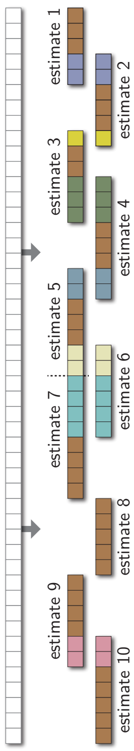


Fig. 2.

As illustrated in Figure 1, approaches towards distributed state estimation usually have in common that each locally computed estimate refers to the same state, for instance, to the position of a single target. However, a high-dimensional state vector renders such a distributed estimation approach rather impractical. As shown in Figure 2, a more desirable approach is then to divide the global state vector into smaller parts for each of which a local estimate is computed on a corresponding node. The local estimates may thereby refer to overlapping parts of the state vector, as indicated by the different colors in Figure 2. For instance, if a large-scale physical process is to be monitored, a single sensor device is only in the position to cover a small area of the phenomenon. In a sensor network deployed to monitor this phenomenon, it is computationally too demanding for a node to maintain an estimate of the entire state vector. Instead, a more effective approach is to let each node only compute an estimate of that part of the state that is affected by its sensor observations. This approach not only reduces the computational pressure of the nodes but also reduces the data volume to be transferred by each node. Furthermore, it is an easy task to increase or decrease the monitored area, i.e., to add elements to or delete elements from the global state vector. Such a scaling only affects neighboring nodes with overlapping parts whereas the usual approach, as shown in Figure 1, implies that each local estimate has to be adapted if the state vector is altered.

In the situation of unequal state representations, the role of state fusion is not only to provide an improved estimate but also to assemble an estimate for the entire state from the locally computed lower-dimensional estimates. The distributed estimation methods discussed in the preceding paragraphs cannot be applied here for fusion. For consensus-based estimation techniques, a method for merging estimates of overlapping parts has been proposed by Stanković et al. (2009). However, a consensus, in general, does not take into account the covariance matrices, i.e., uncertainties, of the locally computed estimates and hence does not systematically reduce the overall estimation error. Against the background of decentralized control problems, Ikeda et al. (1981) consider a similar situation of overlapping information. In our initial studies—see Sijs et al. (2013)—we have investigated an empirical method for fusing estimates of overlapping state vectors that is particularly suited to fully decentralized estimation problems. In the following, we derive a linear minimum mean squared error fusion method that exploits all available information, such as cross-correlations. If this information is not available, a suboptimal fusion strategy can be pursued.

2. PRELIMINARIES

Underlined variables $\underline{x}$ denote vectors or vector-valued functions, and lowercase boldface letters $\underline{\boldsymbol{x}}$ are used for random quantities. Matrices are written in uppercase boldface letters $\mathbf{C}$. By $(\hat{\underline{x}}, \mathbf{C})$, we denote an estimate with mean $\hat{\underline{x}}$ and covariance matrix $\mathbf{C}$. The notation $\hat{\underline{x}}$ is used for the mean of a random variable, an estimate of an uncertain quantity, or an observation. The matrix $\mathbf{I}$ denotes the identity matrix of appropriate dimension.

3. THE CONSIDERED SETUP

Linear estimation problems are studied. The state to be estimated is characterized by a discrete-time process model

$$\underline{\boldsymbol{x}}_{k+1} = \mathbf{A}_k \, \underline{\boldsymbol{x}}_k + \underline{\boldsymbol{w}}_k \tag{1}$$

with system matrix $\mathbf{A}_k$ and zero-mean white process noise $\underline{\boldsymbol{w}}_k$, which has the error covariance matrix $\mathbf{C}_k^w$. Each sensor device provides observations that are related to the state through a linear model

$$\underline{\boldsymbol{z}}_k = \mathbf{H}_k \, \underline{\boldsymbol{x}}_k + \underline{\boldsymbol{v}}_k \ , \tag{2}$$

where $\mathbf{H}_k$ is the measurement matrix and $\underline{\boldsymbol{v}}_k$ denotes a zero-mean white sensor noise with error covariance matrix $\mathbf{C}_k^v$. The index k denotes the time step.

We consider a network of several state estimation systems each of which provides an estimate of an arbitrarily sized part of the global state vector. On each node, a Kalman filter is employed to compute an estimate, where local versions of the models (1) and (2) are used. In particular, (2) refers to the node's sensor equipment. The results of the local estimators are sent to a data sink, where an estimate for the global state vector is to be derived. Since fusion usually involves only estimates referring to the same time step, the index k is omitted in most of the following discussions.

4. JOINT STATE SPACE FORMULATION OF ESTIMATION AND FUSION

As a first step, we revisit common estimation and fusion principles by availing ourselves appropriate joint state space reformulations. Li et al. (2003) have demonstrated that fusion can be stated in terms of a *weighted least squares* (WLS) problem. The WLS method provides a solution $\hat{\underline{x}}^{\mathrm{e}}$ to the minimization problem

$$\hat{\underline{x}}^{\mathrm{e}} = \arg\min_{\underline{x}} \, (\hat{\underline{z}} - \mathbf{H}\,\underline{x})^{\mathrm{T}} (\mathbf{C}^z)^{-1} (\hat{\underline{z}} - \mathbf{H}\,\underline{x}) \ , \tag{3}$$

where all observations $\hat{\underline{z}}_A, \hat{\underline{z}}_B, \ldots$ of an uncertain quantity $\underline{\boldsymbol{x}}$ and according measurement models $\mathbf{H}_A, \mathbf{H}_B, \ldots$ are summarized in

$$\hat{\underline{z}} = \begin{bmatrix} \hat{\underline{z}}_A \\ \hat{\underline{z}}_B \\ \vdots \end{bmatrix} \text{ and } \mathbf{H} = \begin{bmatrix} \mathbf{H}_A \\ \mathbf{H}_B \\ \vdots \end{bmatrix} \ ,$$

respectively. The observation noise terms $\underline{\boldsymbol{v}}_A, \underline{\boldsymbol{v}}_B, \ldots$ have the joint covariance matrix $\mathbf{C}^v$ and are possibly correlated. In this original form, the WLS formulation does not incorporate any prior knowledge on $\underline{\boldsymbol{x}}$. However, as shown by Li et al. (2003), this method can be generalized to a fusion rule that also takes into account prior information.

The solution of problem (3) can be computed by means of the gain

$$\mathbf{K} = \left(\mathbf{H}^{\mathrm{T}} (\mathbf{C}^v)^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^{\mathrm{T}} (\mathbf{C}^v)^{-1} \ . \tag{4}$$

As it is usually done, we assume that $\mathbf{C}^v$ is positive definite, and $\mathbf{H}^{\mathrm{T}} (\mathbf{C}^v)^{-1} \mathbf{H}$ is regular. The optimal estimate then yields

$$\hat{\underline{x}}^{\mathrm{e}} = \mathbf{K}\,\hat{\underline{z}} \tag{5}$$

with

$$\mathbf{C}^{\mathrm{e}} = \mathbf{K} (\mathbf{C}^v) \mathbf{K}' = \left(\mathbf{H}^{\mathrm{T}} (\mathbf{C}^v)^{-1} \mathbf{H}\right)^{-1} \ . \tag{6}$$

The matrix $\mathbf{C}^{\mathrm{e}}$ denotes the *mean squared error* (MSE) matrix [1]

$$\mathbf{C}^{\mathrm{e}} = \mathrm{E}\left[(\hat{\underline{x}}^{\mathrm{e}} - \underline{\boldsymbol{x}})(\hat{\underline{x}}^{\mathrm{e}} - \underline{\boldsymbol{x}})^{\mathrm{T}}\right] \ .$$

In the considered setup, WLS estimates are identical to maximum-likelihood estimates. The following subsections consider the most common fusion problems that are reformulated in terms of WLS problems.

4.1 Bar-Shalom/Campo Fusion

Bar-Shalom and Campo (1986) consider the problem of fusing two tracks of a target. The estimates to be fused are given by $\hat{\underline{x}}_A$ and $\hat{\underline{x}}_B$ with the corresponding MSE matrices

$$\mathbf{C}_A = \mathrm{E}\left[(\hat{\underline{x}}_A - \underline{\boldsymbol{x}})(\hat{\underline{x}}_A - \underline{\boldsymbol{x}})^{\mathrm{T}}\right]$$

[1] Note that trace($\mathbf{C}^{\mathrm{e}}$) is the MSE.

and

$$\mathbf{C}_B = \mathrm{E}\left[(\underline{\hat{x}}_B - \underline{\boldsymbol{x}})(\underline{\hat{x}}_B - \underline{\boldsymbol{x}})^{\mathrm{T}}\right] ,$$

where the cross-covariance matrix

$$\mathbf{C}_{AB} = \mathrm{E}\left[(\underline{\hat{x}}_A - \underline{\boldsymbol{x}})(\underline{\hat{x}}_B - \underline{\boldsymbol{x}})^{\mathrm{T}}\right]$$

is in general nonzero because the same process model (1) is used for both estimates. Hence, the same process noise is modeled twice and leads to correlations. The Bar-Shalom/Campo formulas allow to exploit these dependencies in order to compute a minimum MSE fusion result. They are derived by determining optimal gains for the linear combination

$$\underline{\hat{x}}^{\mathrm{e}} = \mathbf{K}_A \, \underline{\hat{x}}_A + \mathbf{K}_B \, \underline{\hat{x}}_B \ .$$

The same estimate $\underline{\hat{x}}^{\mathrm{e}}$ can be obtained when this fusion problem is redefined to a WLS problem. For this purpose, the estimates $\underline{\hat{x}}_A$ and $\underline{\hat{x}}_B$ are interpreted as observations by considering the identities

$$\begin{aligned} \underline{\hat{x}}_A &= \underline{\boldsymbol{x}} + (\underline{\hat{x}}_A - \underline{\boldsymbol{x}}) = \underline{\boldsymbol{x}} + \underline{\tilde{\boldsymbol{x}}}_A \text{ and} \\ \underline{\hat{x}}_B &= \underline{\boldsymbol{x}} + (\underline{\hat{x}}_B - \underline{\boldsymbol{x}}) = \underline{\boldsymbol{x}} + \underline{\tilde{\boldsymbol{x}}}_B \ , \end{aligned} \tag{7}$$

where $\underline{\tilde{\boldsymbol{x}}} := [\underline{\tilde{\boldsymbol{x}}}_A, \underline{\tilde{\boldsymbol{x}}}_B]^{\mathrm{T}}$ represents the observation noise. The corresponding observation model then reads

$$\begin{bmatrix} \underline{\hat{x}}_A \\ \underline{\hat{x}}_B \end{bmatrix} = \mathbf{H} \, \underline{\boldsymbol{x}} + \underline{\tilde{\boldsymbol{x}}} \quad \text{with } \mathbf{H} = \begin{bmatrix} \mathbf{I} \\ \mathbf{I} \end{bmatrix} \ . \tag{8}$$

The measurement noise $\underline{\tilde{\boldsymbol{x}}}$ has the covariance matrix

$$\begin{aligned} \mathbf{C}^x &= \mathrm{E}\left[\left(\begin{bmatrix} \underline{\hat{x}}_A \\ \underline{\hat{x}}_B \end{bmatrix} - \begin{bmatrix} \mathbf{I} \\ \mathbf{I} \end{bmatrix} \underline{\boldsymbol{x}} \right) \left(\begin{bmatrix} \underline{\hat{x}}_A \\ \underline{\hat{x}}_B \end{bmatrix} - \begin{bmatrix} \mathbf{I} \\ \mathbf{I} \end{bmatrix} \underline{\boldsymbol{x}} \right)^{\mathrm{T}} \right] \\ &= \begin{bmatrix} \mathbf{C}_A & \mathbf{C}_{AB} \\ \mathbf{C}_{BA} & \mathbf{C}_B \end{bmatrix} \ . \end{aligned} \tag{9}$$

With these parameters, the WLS gain (4) can now be computed that enters into equations (5) and (6) for the estimate and error covariance matrix, respectively.

We demonstrate the relation to the Bar-Shalom/Campo fusion rule by considering the error covariance matrix (6), i.e.,

$$\mathbf{C}^{\mathrm{e}} = \left(\mathbf{H}^{\mathrm{T}} (\mathbf{C}^x)^{-1} \mathbf{H} \right)^{-1} ,$$

where $\mathbf{C}^x$, given by (9), can be inverted blockwise according to

$$(\mathbf{C}^x)^{-1} = \begin{bmatrix} \mathbf{C}_A^{-1} + \mathbf{C}_A^{-1}\mathbf{C}_{AB}\mathbf{W}\mathbf{C}_{BA}\mathbf{C}_A^{-1} & -\mathbf{C}_A^{-1}\mathbf{C}_{AB}\mathbf{W} \\ -\mathbf{W}\mathbf{C}_{BA}\mathbf{C}_A^{-1} & \mathbf{W} \end{bmatrix}^{-1}$$

with $\mathbf{W} = (\mathbf{C}_B - \mathbf{C}_{BA}\mathbf{C}_A^{-1}\mathbf{C}_{AB})^{-1}$. The transformation with $\mathbf{H}$ then yields the sum

$$\begin{aligned} \mathbf{H}^{\mathrm{T}} (\mathbf{C}^x)^{-1} \mathbf{H} &= \mathbf{C}_A^{-1} + \mathbf{C}_A^{-1}\mathbf{C}_{AB}\mathbf{W}\mathbf{C}_{BA}\mathbf{C}_A^{-1} \\ &\quad - \mathbf{C}_A^{-1}\mathbf{C}_{AB}\mathbf{W} - \mathbf{W}\mathbf{C}_{BA}\mathbf{C}_A^{-1} + \mathbf{W} \\ &= \mathbf{C}_A^{-1} + (\mathbf{C}_A^{-1}\mathbf{C}_{AB} - \mathbf{I})\mathbf{W}(\mathbf{C}_{BA}\mathbf{C}_A^{-1} - \mathbf{I}) \end{aligned}$$

and finally, with the Woodbury matrix identity, we have

$$\begin{aligned} \mathbf{C}^{\mathrm{e}} &= \mathbf{C}_A - \mathbf{C}_A(\mathbf{C}_A^{-1}\mathbf{C}_{AB} - \mathbf{I}) \\ &\quad \left(\mathbf{W}^{-1} + (\mathbf{C}_{BA}\mathbf{C}_A^{-1} - \mathbf{I})\mathbf{C}_A(\mathbf{C}_A^{-1}\mathbf{C}_{AB} - \mathbf{I}) \right)^{-1} \\ &\quad (\mathbf{C}_{BA}\mathbf{C}_A^{-1} - \mathbf{I})\mathbf{C}_A \\ &= \mathbf{C}_A - (\mathbf{C}_A - \mathbf{C}_{AB}) \\ &\quad (\mathbf{C}_A + \mathbf{C}_B - \mathbf{C}_{AB} - \mathbf{C}_{BA})^{-1}(\mathbf{C}_A - \mathbf{C}_{BA}) \ . \end{aligned}$$

This result equals the Bar-Shalom/Campo covariance matrix (see Bar-Shalom and Campo (1986)).

It is important to point out that the Bar-Shalom/Campo combination of estimates provides us with a minimum MSE fusion result but not with a minimum MSE estimate given all available measurements. More precisely, a central processing of all measurements from devices A and B generally yields better estimates than local computations of estimates with subsequent fusion. This issue is, for instance, discussed by Chang et al. (1997).

4.2 Measurement Fusion

The WLS fusion formalism can also be utilized to redevelop the standard Kalman filtering step, i.e., the fusion of prior information with an observation. A measurement $\underline{\hat{z}}_k$ at time step k is related to the state by the sensor model (2). As it has been done in (7), the prior estimate $(\underline{\hat{x}}_k^{\mathrm{p}}, \mathbf{C}_k^{\mathrm{p}})$ is now also regarded as a measurement, and we obtain the joint measurement model

$$\begin{bmatrix} \underline{\hat{x}}_k^{\mathrm{p}} \\ \underline{\hat{z}}_k \end{bmatrix} = \begin{bmatrix} \underline{\boldsymbol{x}} + (\underline{\hat{x}}_k^{\mathrm{p}} - \underline{\boldsymbol{x}}) \\ \mathbf{H}_k \, \underline{\boldsymbol{x}} + \underline{\boldsymbol{v}}_k \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \mathbf{H}_k \end{bmatrix} \underline{\boldsymbol{x}} + \underline{\tilde{\boldsymbol{x}}} \ , \tag{10}$$

where $\underline{\tilde{\boldsymbol{x}}}$ is characterized by the joint noise matrix

$$\mathbf{C}^x = \begin{bmatrix} \mathbf{C}_k^{\mathrm{p}} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_k^v \end{bmatrix} \ .$$

The measurement is assumed to be conditionally independent of the current estimate as it is common practice. According to (6), the covariance matrix of the WLS result can now be computed by

$$\begin{aligned} \mathbf{C}^{\mathrm{e}} &= \left([\mathbf{I} \ \mathbf{H}_k] \, (\mathbf{C}^x)^{-1} \begin{bmatrix} \mathbf{I} \\ \mathbf{H}_k \end{bmatrix} \right)^{-1} \\ &= \left((\mathbf{C}_k^{\mathrm{p}})^{-1} + \mathbf{H}_k^{\mathrm{T}} (\mathbf{C}_k^v)^{-1} \mathbf{H}_k \right)^{-1} \ . \end{aligned}$$

This result is identical to the information filter formulas (Mutambara, 1998), which are an algebraical reformulation of the Kalman filter algorithm.

This subsection has demonstrated that prior information can also be incorporated by the WLS fusion method. In this case, even a linear minimum MSE (LMMSE) estimator is obtained. As already stated by Li et al. (2003), the WLS formulation can be regarded as a universal tool for fusion problems.

4.3 Results and the Next Step

Sections 4.1 and 4.2 have demonstrated that optimal fusion can be expressed in terms of the solution to a WLS problem. However, in both cases, the quantities to be fused, be it $\underline{\hat{x}}_A$ and $\underline{\hat{x}}_B$ in (8) or $\underline{\hat{x}}_k^{\mathrm{p}}$ and $\underline{\hat{z}}_k$ in (10), refer to the same state vector $\underline{\boldsymbol{x}}$. In the remainder of this work, estimates of different state vectors $\underline{\boldsymbol{x}}_1$ and $\underline{\boldsymbol{x}}_2$ are considered, which may partially overlap. So far, the problem of fusing estimates on unequal state vectors has been treated by consensus-based (Stanković et al., 2009) or empirical methods (Sijs et al., 2013). In the following, a minimum MSE fusion method is provided based on an WLS formulation and also, a fusion strategy under unknown correlations is presented.

5. GENERALIZED FUSION STRATEGIES

The considerations of the previous section brings us close to a solution for the general problem. The fusion rule of Section 4.1 represents an important special case of the considered problem, i.e., the estimates to be fused correspond to the same parts of the state vector and there are no exclusive parts. For this special case, Shin et al.

(2006) have shown how to fuse multiple estimates at once, for instance, in order to combine the estimates in Figure 1. The opposite case consists of estimate that are computed for exclusive parts of the state vector, which is studied in the first subsection. After that, the optimal fusion strategy for overlapping state vectors is discussed. The treatment of unknown cross-correlations lies in the focus of the third part of this section.

5.1 Fusion of Non-overlapping Parts

The simplest case is the problem of fusing state estimates of exclusive parts of the state vector. For the state $\underline{\boldsymbol{x}} = [\underline{\boldsymbol{x}}_1^{\mathrm{T}}, \underline{\boldsymbol{x}}_2^{\mathrm{T}}]^{\mathrm{T}}$, the estimates $\hat{\underline{x}}_A$ and $\hat{\underline{x}}_B$ for $\underline{\boldsymbol{x}}_1$ and $\underline{\boldsymbol{x}}_2$, respectively, are considered. In line with (7), these two estimates are again regarded as measurements of the state vector. In contrast to equation (8), the measurement equation now becomes

$$\begin{bmatrix} \hat{\underline{x}}_A \\ \hat{\underline{x}}_B \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \underline{\boldsymbol{x}}_1 \\ \underline{\boldsymbol{x}}_2 \end{bmatrix} + \tilde{\underline{\boldsymbol{x}}} \,,$$

i.e., the matrix $\mathbf{H}$ becomes the identity. As a consequence, the gain (4) that solves the WLS problem also reduces to the identity.

Of course, this result is rather obvious: An estimate for the joint state vector is the joint estimate. The local estimates have to be merged into $\hat{\underline{x}}^{\mathrm{e}} = [\hat{\underline{x}}_A^{\mathrm{T}}, \hat{\underline{x}}_B^{\mathrm{T}}]^{\mathrm{T}}$, and the MSE matrix is simply the joint error matrix $\mathbf{C}^{\mathrm{e}} = \mathbf{C}^x$, which is (9). The fusion of more than two estimates on exclusive parts is carried out in the same fashion.

5.2 Optimal Fusion

The primary aim of this paper is to provide a fusion method for estimates of arbitrary size and number. For the sake of clarity, we first discuss the simpler case of two estimates $\hat{\underline{x}}_A = [\hat{\underline{x}}_{A1}^{\mathrm{T}}, \hat{\underline{x}}_{A2}^{\mathrm{T}}]^{\mathrm{T}}$ and $\hat{\underline{x}}_B = [\hat{\underline{x}}_{B2}^{\mathrm{T}}, \hat{\underline{x}}_{B3}^{\mathrm{T}}]^{\mathrm{T}}$, where $\hat{\underline{x}}_{A2}$ and $\hat{\underline{x}}_{B2}$ are estimates on the same part $\underline{\boldsymbol{x}}_2$ of the state vector $\underline{\boldsymbol{x}} = [\underline{\boldsymbol{x}}_1^{\mathrm{T}}, \underline{\boldsymbol{x}}_2^{\mathrm{T}}, \underline{\boldsymbol{x}}_3^{\mathrm{T}}]^{\mathrm{T}}$. The partial estimates $\hat{\underline{x}}_{A1}$ and $\hat{\underline{x}}_{B3}$ refer to the exclusive parts $\underline{\boldsymbol{x}}_1$ and $\underline{\boldsymbol{x}}_3$. Again, by regarding each estimate as an observation, the measurement equation

$$\begin{matrix} \hat{\underline{x}}_A \left\{ \right. \\ \hat{\underline{x}}_B \left\{ \right. \end{matrix} \begin{bmatrix} \hat{\underline{x}}_{A1} \\ \hat{\underline{x}}_{A2} \\ \hat{\underline{x}}_{B2} \\ \hat{\underline{x}}_{B3} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}}_{=\mathbf{H}} \begin{bmatrix} \underline{\boldsymbol{x}}_1 \\ \underline{\boldsymbol{x}}_2 \\ \underline{\boldsymbol{x}}_3 \end{bmatrix} + \tilde{\underline{\boldsymbol{x}}} \qquad (11)$$

can be set up, where the measurement noise $\tilde{\underline{\boldsymbol{x}}}$ has the covariance matrix (9). The gain (4) can then be computed by means of the matrices $\mathbf{H}$ and $\mathbf{C}^x$, and finally an estimate $(\hat{\underline{x}}^{\mathrm{e}}, \mathbf{C}^{\mathrm{e}})$ is attained according to (5) and (6).

Remark 1. In general, the state estimates for the parts $\underline{\boldsymbol{x}}_1$, $\underline{\boldsymbol{x}}_2$, and $\underline{\boldsymbol{x}}_3$ are correlated due to the process model, and, hence, the joint covariance matrix $\mathbf{C}^x$ of the state estimates is fully occupied. In particular, also the correlations between the estimation errors of $\hat{\underline{x}}_{A1}$ and $\hat{\underline{x}}_{B3}$ have to be addressed. In the particular case that the process model does not cause dependencies between the estimates of $\underline{\boldsymbol{x}}_1$, $\underline{\boldsymbol{x}}_2$, and $\underline{\boldsymbol{x}}_3$, the joint covariance matrix (9) can be simplified to

$$\begin{bmatrix} \mathbf{C}_A & \mathbf{C}_{AB} \\ \mathbf{C}_{BA} & \mathbf{C}_B \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{A1\,A1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{A2\,A2} & \mathbf{C}_{A2\,B2} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{B2\,A2} & \mathbf{C}_{B2\,B2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{C}_{B3\,B3} \end{bmatrix} ,$$

where only correlations between the overlapping parts remain. Fusion, in this case, becomes a combination of the Bar-Shalom/Campo rule for $\hat{\underline{x}}_{A2}$ and $\hat{\underline{x}}_{B2}$ and the fusion of the non-overlapping estimates $\hat{\underline{x}}_{A1}$ and $\hat{\underline{x}}_{B3}$.

In order to fuse arbitrarily many estimates $\hat{\underline{x}}_A, \hat{\underline{x}}_B, \hat{\underline{x}}_C, \ldots$ on the state vector $\underline{\boldsymbol{x}} = [\underline{\boldsymbol{x}}_1^{\mathrm{T}}, \underline{\boldsymbol{x}}_2^{\mathrm{T}}, \underline{\boldsymbol{x}}_3^{\mathrm{T}}, \ldots]^{\mathrm{T}}$ into a single, updated estimate, the two-vector case (11) is expanded to the general measurement equation

$$\begin{bmatrix} \hat{\underline{x}}_{A1} \\ \hat{\underline{x}}_{A2} \\ \vdots \\ \hat{\underline{x}}_{B1} \\ \hat{\underline{x}}_{B2} \\ \vdots \end{bmatrix} = \mathbf{H} \begin{bmatrix} \underline{\boldsymbol{x}}_1 \\ \underline{\boldsymbol{x}}_2 \\ \underline{\boldsymbol{x}}_3 \\ \vdots \end{bmatrix} + \tilde{\underline{\boldsymbol{x}}} \,. \qquad (12)$$

More precisely, the vector on the left-hand side collects all (existent) $\hat{\underline{x}}_{Xj}$ with $X = A, B, C, \ldots$ and $j = 1, 2, 3, \ldots$, where the vector $\hat{\underline{x}}_{Xj}$ is that part of the local estimate $\hat{\underline{x}}_X$ that refers to $\underline{\boldsymbol{x}}_j$. The measurement matrix $\mathbf{H}$ is set up from blocks of identity matrices according to

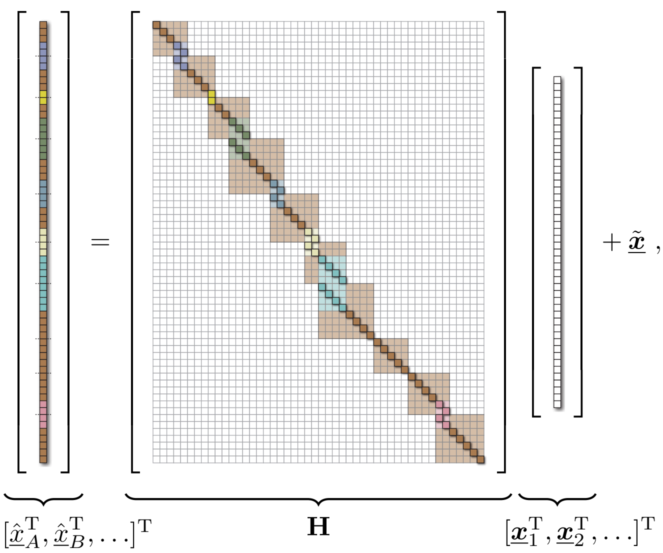
$$\mathbf{H} = \underbrace{\begin{bmatrix} \ddots & & \\ & \cdots\ \mathbf{I}\ \cdots & \\ & & \ddots \end{bmatrix}}_{j} \Big\} X \qquad (13)$$

and maps $\underline{\boldsymbol{x}}_j$ to the component $\hat{\underline{x}}_{Xj}$ of the local estimate $\hat{\underline{x}}_X$. The joint covariance matrix for $\tilde{\underline{\boldsymbol{x}}}$ has the form

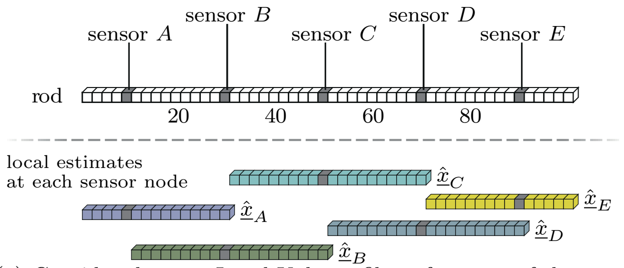
$$\mathbf{C}^x = \begin{bmatrix} \mathbf{C}_A & \mathbf{C}_{AB} & \mathbf{C}_{AC} & \\ \mathbf{C}_{BA} & \mathbf{C}_B & \mathbf{C}_{BC} & \cdots \\ \mathbf{C}_{CA} & \mathbf{C}_{CB} & \mathbf{C}_C & \\ & \vdots & & \ddots \end{bmatrix} . \qquad (14)$$

The formulas (4), (5), and (6) can now be employed to reconstruct an estimate $\hat{\underline{x}}^{\mathrm{e}}$ for the entire state vector $\underline{\boldsymbol{x}}$.

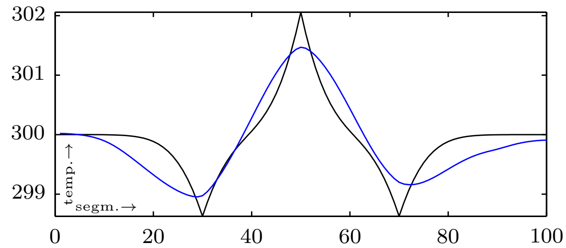
Example 2. For the partition into unequal state representations that has been illustrated in Figure 2, a schematic representation of the corresponding measurement equation (12) is



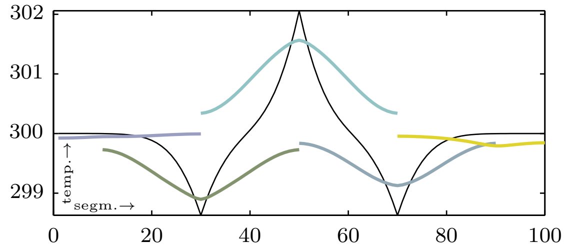
where solid boxes represent entries with value 1 and each other value is 0. More precisely, the shaded boxes indicate the identity matrices and correspond to the structure of



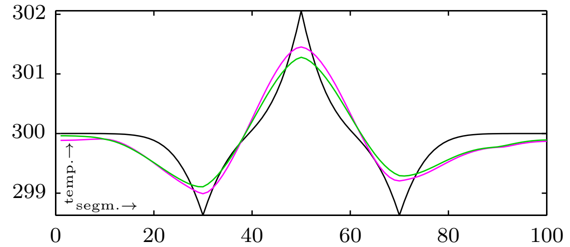
(a) Considered setup. Local Kalman filters for parts of the state.



(b) Centralized Kalman Filter.



(c) Local Kalman filter estimates $\hat{\underline{x}}_A$, $\hat{\underline{x}}_B$, $\hat{\underline{x}}_C$, $\hat{\underline{x}}_D$, and $\hat{\underline{x}}_E$.



(d) Optimal(—) and conservative(—) fusion of estimates in Fig.3(c).

Fig. 3. Temperature profile of rod is to be estimated and is drawn black. A centralized Kalman filter is compared with several local Kalman filters on unequal parts of the state and subsequent fusion.

matrix (13). It should be noticed that $\mathbf{H}$ does not need to be of such a close-to-diagonal structure—for instance, the first estimate could also refer to parts that belong to any other than the second estimate.

5.3 Fusion under Unknown Correlations

The prerequisite that the joint covariance matrix (14) has to be known often renders the optimal fusion strategy difficult to apply. The task of bookkeeping or reconstructing the joint covariance matrix is an often-encountered problem in applying distributed Kalman filters. If the off-diagonal blocks in (14) are nonzero, a local update of an estimate $\hat{\underline{x}}_X$ effects also an update of all cross-covariance matrices $\mathbf{C}_{XY}$, $Y = A, B, \ldots$. This form of interdependencies among the local estimators diminishes important advantages of networked information processing, such as scalability and robustness.

Conservative fusion strategies constitute an alternative to optimal fusion methods that rely on a completely known cross-covariance matrix. In respect of storage and processing complexity, it may be desirable to intentionally discard cross-correlations between local estimates. Julier and Uhlmann (1997) have shown by means of the *covariance intersection* algorithm how to treat unknown cross-correlations when two estimates are to be fused. While this solution is based on the information form of the Kalman filter, Hanebeck et al. (2001) formulate this fusion principle in terms of a bound on the joint covariance matrix. This second formulation—the *covariance bounds* approach—can be employed to compute a conservative bound of the joint covariance matrix (14), which is given by

$$\underbrace{\begin{bmatrix} \frac{1}{w_A}\mathbf{C}_A & \mathbf{0} & \mathbf{0} & \\ \mathbf{0} & \frac{1}{w_B}\mathbf{C}_B & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \frac{1}{w_C}\mathbf{C}_C & \\ & \vdots & & \ddots \end{bmatrix}}_{=:\mathbf{C}^{\mathrm{CB}}} \geq \begin{bmatrix} \mathbf{C}_A & \mathbf{C}_{AB} & \mathbf{C}_{AC} & \\ \mathbf{C}_{BA} & \mathbf{C}_B & \mathbf{C}_{BC} & \cdots \\ \mathbf{C}_{CA} & \mathbf{C}_{CB} & \mathbf{C}_C & \\ & \vdots & & \ddots \end{bmatrix}$$

with positive $w_A, w_B, \ldots$ and $w_A + w_B + \ldots = 1$. The inequality $\mathbf{C}^{\mathrm{CB}} \geq \mathbf{C}^x$ means that the difference $\mathbf{C}^{\mathrm{CB}} - \mathbf{C}^x$ is a positive semidefinite matrix. In place of $\mathbf{C}^x$, the matrix $\mathbf{C}^{\mathrm{CB}}$ can now be utilized to compute a covariance-consistent estimate by means of (4), (5), and (6).

This approach not only bypasses the need for bookkeeping of cross-correlations but also reduces the computational complexity for inverting the joint covariance matrix. However, one has to weigh these advantages against the optimal fusion strategy, which offers a lower estimation uncertainty, i.e., a lower MSE.

6. SIMULATIONS

In order to illustrate and discuss the proposed fusion strategies, the process of heat conduction in a rod is considered. For this purpose, we revisit the example in (Sijs et al., 2013). The rod is divided into 100 segments, for each of which the temperature is to be estimated. The initial temperature is 300 K, and the rod is heated at segment 50 with 15 W and is cooled at segments 30 and 70 with −10 W each. The discretized process model is

$$\boldsymbol{x}_{k+1,n} = 0.17\boldsymbol{x}_{k,n-1} + 0.66\boldsymbol{x}_{k,n} + 0.17\boldsymbol{x}_{k,n+1} + \boldsymbol{w}_{k,n} \ , \quad (15)$$

where the index $k \in \mathbb{N}$ denotes the time step, and $n \in \{1, 2, \ldots, 100\}$ is the n-th segment of the rod. The process noise is normally distributed according to $\boldsymbol{w}_{k,n} \sim \mathcal{N}(0, 30)$ and is used by the estimators to represent the uncertainty about temperature changes, i.e., to represent unknown inputs. The rod is equipped with five sensor nodes, as illustrated in Figure 3(a). They are located at the segments 10, 30, 50, 70 and 90 with the sensor models

$$\boldsymbol{z}_k^A = \boldsymbol{x}_{k,10} + \boldsymbol{v}_k^A \ , \ \boldsymbol{z}_k^B = \boldsymbol{x}_{k,30} + \boldsymbol{v}_k^B \ , \ \boldsymbol{z}_k^C = \boldsymbol{x}_{k,50} + \boldsymbol{v}_k^C \ ,$$
$$\boldsymbol{z}_k^D = \boldsymbol{x}_{k,70} + \boldsymbol{v}_k^D \ , \ \boldsymbol{z}_k^E = \boldsymbol{x}_{k,90} + \boldsymbol{v}_k^E \ .$$

Each sensor noise term is a zero-mean white noise with variance 0.01. The estimation results after 60 time step are shown in Figure 3. The optimal estimate can be seen in Figure 3(b), where each sensor transmits its measurements to the data sink at every time step. A centralized Kalman filter in the data sink computes this estimate. In place of a centralized processing, Figure 3(c) shows the results of local Kalman filters at the five sensor nodes, i.e., each local filter uses only the measurements provided by the corresponding sensor (see Figure 3(a)). At each node, only an estimate for a small part of the state is computed, i.e.,

$$\hat{\underline{x}}_A \text{ of } [\boldsymbol{x}_{k,1}, \ldots, \boldsymbol{x}_{k,30}]^{\mathrm{T}} , \quad \hat{\underline{x}}_B \text{ of } [\boldsymbol{x}_{k,10}, \ldots, \boldsymbol{x}_{k,50}]^{\mathrm{T}} ,$$
$$\hat{\underline{x}}_C \text{ of } [\boldsymbol{x}_{k,30}, \ldots, \boldsymbol{x}_{k,70}]^{\mathrm{T}} , \quad \hat{\underline{x}}_D \text{ of } [\boldsymbol{x}_{k,50}, \ldots, \boldsymbol{x}_{k,90}]^{\mathrm{T}} ,$$
$$\hat{\underline{x}}_E \text{ of } [\boldsymbol{x}_{k,70}, \ldots, \boldsymbol{x}_{k,100}]^{\mathrm{T}} .$$

Each local filter employs a local process model, i.e., (15) is only applied to the corresponding part of the state vector. In order to take the influences from adjacent segments into account, the process noise for the boundary segments of the partial state vectors is increased to 50, e. g., $\boldsymbol{w}_{k,30} \sim \mathcal{N}(0, 50)$ and $\boldsymbol{w}_{k,70} \sim \mathcal{N}(0, 50)$ for the local Kalman filter at sensor node C.

At time step $k = 60$, the local estimates from Figure 3(c) are fused at the data sink. Both the optimal and the conservative fusion result are depicted in Figure 3(d). The magenta plot is closer to the temperature profile, but requires bookkeeping of the entire joint covariance matrix. Since optimal MMSE fusion generally does not yield the MMSE estimate (see Chang et al. (1997)), the magenta result is different from Figure 3(b). The green conservative fusion result has been computed with the aid of a bounding covariance matrix, as explained in Section 5.3. However, due to the high process noise, the difference between the green and magenta estimate is not significant.

7. CONCLUSIONS

For the treatment of unequal state representations, fusion strategies have been presented that are based on a generalized WLS problem. This WLS formulation has been employed to express standard estimation problems, i.e., the Kalman measurement update and track-to-track fusion, and has then been further developed in order to fuse local estimates of arbitrarily sized parts of the state vector. If cross-correlations among all local estimates are known, an optimal fusion result can be derived. If information on cross-correlations is missing, a conservative fusion strategy can be pursued, which employs an upper bound on the unknown joint covariance matrix. The proposed concept is intended for a networked system where an estimate of the entire state vector is to be reconstructed from estimates of unequal state representations. In this regard, we have assumed that the fusion process is carried out by a central node in the network. Consequently, fully decentralized fusion architectures are an important topic for future work. Here, each node is intended to operate independently, and state estimates can be exchanged between nodes. Fusion then takes place locally, and a decentralized formulation of the proposed fusion strategies is to be found.

REFERENCES

Bar-Shalom, Y. and Campo, L. (1986). The Effect of the Common Process Noise on the Two-Sensor Fused-Track Covariance. *IEEE Transactions on Aerospace and Electronic Systems*, 22(6), 803–805.

Bar-Shalom, Y., Li, X.R., and Kirubarajan, T. (2001). *Estimation with Applications to Tracking and Navigation: Theory, Algorithms, and Software.* John Wiley & Sons.

Carli, R., Chiuso, A., Schenato, L., and Zampieri, S. (2008). Distributed Kalman Filtering Based on Consensus Strategies. *IEEE Journal on Selected Areas in Communications*, 26(4), 622–633.

Chang, K.C., Saha, R.K., and Bar-Shalom, Y. (1997). On Optimal Track-to-Track Fusion. *IEEE Transactions on Aerospace and Electronic Systems*, 33(4), 1271–1276.

Chong, C.Y., Mori, S., Chang, K.C., and Barker, W.H. (2000). Architectures and Algorithms for Track Association and Fusion. *IEEE Aerospace and Electronic Systems Magazine*, 15(1), 5–13.

Govaers, F. and Koch, W. (2010). Distributed Kalman Filter Fusion at Arbitrary Instants of Time. In *Proceedings of the 13th International Conference on Information Fusion (Fusion 2010)*. Edinburgh, United Kingdom.

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.

Hanebeck, U.D., Briechle, K., and Horn, J. (2001). 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)*, 85–90. Baden–Baden, Germany.

Ikeda, M., Šiljak, D.D., and White, D.E. (1981). Decentralized Control With Overlapping Information Sets. *Journal of Optimization Theory and Applications*, 34(2), 279–310.

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.

Kalman, R.E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME - Journal of Basic Engineering*, 35–45.

Li, X.R., Zhu, Y., Wang, J., and Han, C. (2003). Optimal Linear Estimation Fusion—Part I: Unified Fusion Rule. *IEEE Transactions on Information Theory*, 49(9), 2192–2208.

Liggins II, M.E., Hall, D.L., and Llinas, J. (eds.) (2009). *Handbook of Multisensor Data Fusion: Theory and Practice.* The Electrical Engineering and Applied Signal Processing Series. CRC Press, 2nd ed. edition.

Mutambara, A.G.O. (1998). *Decentralized Estimation and Control for Multisensor Systems.* CRC Press, Inc., Boca Raton, Florida, USA.

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

Olfati-Saber, R. (2007). Distributed Kalman Filtering for Sensor Networks. In *Proceedings of the 46th IEEE Conference on Decision and Control (CDC 2007)*, 5492–5498. New Orleans, Louisiana, USA.

Shin, V., Lee, Y., and Choi, T.S. (2006). Generalized Millman's Formula and Its Application for Estimation Problems. *Signal Processing*, 86(2), 257–266.

Sijs, J., Hanebeck, U.D., and Noack, B. (2013). An Empirical Method to Fuse Partially Overlapping State Vectors for Distributed State Estimation. In *Proceedings of the 2013 European Control Conference (ECC 2013)*. Zürich, Switzerland.

Stanković, S.S., Stanković, M.S., and Stipanović, D.M. (2009). Consensus Based Overlapping Decentralized Estimator. *IEEE Transactions on Automatic Control*, 54(2), 410–415.