The Hypothesizing Distributed Kalman Filter

Marc Reinhardt, Benjamin Noack, and Uwe D. Hanebeck

***Abstract*— This paper deals with distributed information processing in sensor networks. We propose the Hypothesizing Distributed Kalman Filter that incorporates an assumption of the global measurement model into the distributed estimation process. The procedure is based on the Distributed Kalman Filter and inherits its optimality when the assumption about the global measurement uncertainty is met. Recursive formulas for local processing as well as for fusion are derived. We show that the proposed algorithm yields the same results, no matter whether the measurements are processed locally or globally, even when the process noise is not negligible. For further processing of the estimates, a consistent bound for the error covariance matrix is derived. All derivations and explanations are illustrated by means of a new classification scheme for estimation processes.**

I. Introduction

In times of a strongly interconnected world, there is an increasing need for massive information processing that can hardly be performed at one central node. So the data from millions of individuals are collected, processed, and stored on distributed systems. Then, only a small subset of the data has to be combined at a central node for evaluation.

With this distribution of information processing, additional challenges arise in several research areas from data mining to new hardware developments. In this paper, we model information by probability density functions and provide a solution for distributed information processing when a common phenomenon is estimated.

More formally, we assume to have a network of nodes that irregularly obtain measurements of the common phenomenon at discrete time steps. As we estimate the state over multiple time steps, the estimates are predicted to the next time step by a state transition model that typically involves common noise and leads to dependencies between the estimates and is referred to as *common process noise* in literature. A second challenge comes from past data exchanges between the nodes. The exchanged information must be maintained to prevent the nodes from double counting. This even holds for transitive exchanges, as a node D that obtains estimates from nodes B and C must be aware of information from a source A that has been exchanged with B and C. This effect is referred to as *common prior information* in literature [1].

When we assume that all measurements can be processed at a single node, the Kalman Filter (KF) can be shown to be optimal in the sense of minimizing the mean-squared-error (MSE) matrix when the models are linear. The KF solution, however, is not directly applicable to distributed processing due to the common process noise. So the idea is to find an algorithm that provides the same results as the KF but can be applied to distributed estimation.

So far, a lot of ideas have been proposed concerning distributed and decentralized processing of information. Starting with the Bar-Shalom/Campo formulas [2], [3] that describe the MSE matrix minimizing way of fusing two estimates with known correlations, continuing with Millman's formulas that extend this idea to the N estimates case [4], [5], to the Channel Filter [6] that provides an idea of maintaining common prior information, a bunch of tools is available for distributed information processing. Therefore, algorithms have been proposed that are directly based on those ideas or incorporate some of them. For example, consensus approaches [7], [8] try to reach kind of a common average estimate. An alternative approach is to preprocess the measurements locally and to calculate an actual estimate at the fusion center only [9].

In particular, those algorithms that guarantee MSE matrix bounds are of interest. The Federated Kalman Filter [10], [11] utilizes a standard KF but overestimates the process noise in order to conservatively approximate the MSE matrix globally. When the dependency structure is not maintained, algorithms that operate under unknown correlations such as Covariance Intersection [12], [13] or generalizations of it [14] can be applied.

Indeed, up to the derivation of the Distributed Kalman Filter (DKF) [15], [16], there has been no recursive algorithm that provided the same result as the centralized KF without having access to all measurements over arbitrarily many time steps without communication. The DKF assumes the nodes to have global knowledge about the models of all other nodes and by means of this models obtains globalized estimates locally that allow a distributed processing of measurements without information loss.

This globalized processing comes at the expense of locally suboptimal estimates that are in general biased. In [17], this bias has been derived explicitly and an algorithm to correct the bias has been proposed. The correction is, in particular, valuable when the global knowledge about the models of the other nodes cannot be guaranteed to be exact or when nodes may fail. Thus, the DKF with the extension from [17] is a robust algorithm that provides optimal results when the assumption about a global measurement model is met.

However, the bias correction requires the utilized models to be available for obtaining an unbiased estimate, which makes it necessary to communicate the models. When the

Marc Reinhardt, Benjamin Noack, and Uwe D. Hanebeck are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics, Karlsruhe Institute of Technology (KIT), Germany. marc.reinhardt@ieee.org, benjamin.noack@ieee.org,uwe.hanebeck@ieee.org

models cannot be stored and exchanged effectively, this can lead to a high traffic between the nodes.

In this paper, we relax the aforementioned requirement and derive the Hypothesizing Distributed Kalman Filter (HKF) that calculates a fused correction matrix from recursively calculated local ones. Based on the structure and the variables of a 3-layer classification, which are introduced in Sec. III, we present a slightly modified version of the DKF in Sec. IV and propose the local processing and the fusion algorithm of the HKF in Sec. V. In Sec. VI, we derive MSE matrix bounds for the fused estimate and finally conclude the paper with a summary and an outlook. As a result, we will have a distributed algorithm that is flexible and robust in its application, has a design parameter in order to improve the results up to the global optimal ones, and can be calculated recursively.

II. Problem Statement

We investigate discrete-time information processing in a network of N nodes that cooperatively estimate the unknown state of a common phenomenon. While the optimal solution is given by the well-known KF when all information is available locally and can be processed in a centralized manner, the derivation of a flexible distributed algorithm that is globally optimal, i.e., the MSE matrix is equivalent to that of the KF, is still challenging.

We assume the (potentially time-varying) state transition model to be known to all nodes and to be linear with an additive, white Gaussian noise term

$$\underline{\boldsymbol{x}}_{k+1} = \mathbf{A}_k \underline{\boldsymbol{x}}_k + \underline{\boldsymbol{w}}_k \text{ with } \underline{\boldsymbol{w}}_k \sim \mathcal{N}\left(\underline{0}, \tilde{\mathbf{C}}_k^w\right) \ . \quad (1)$$

The estimator is only aware of a consistent approximation $\mathbf{C}_k^w \geq \tilde{\mathbf{C}}_k^w$, where $\geq$ describes the Loewner partial order, which means that the difference between the larger and smaller covariance matrix is positive semi-definite. The uncertainty or a lack of knowledge concerning the model can be expressed by an increased process noise.

The measurements are linked to the true state by a potentially nonlinear function that is assumed to be approximated by a linear measurement function with additive white Gaussian noise. We assume the nodes with indices $\tilde{M}_k \subseteq \{1, \ldots, N\}$ to obtain measurements according to the models

$$\underline{\tilde{z}}_k^j = \mathbf{H}_k^j \underline{\boldsymbol{x}}_k + \underline{\boldsymbol{v}}_k^j \text{ with } \underline{\boldsymbol{v}}_k^j \sim \mathcal{N}\left(\underline{0}, \tilde{\mathbf{C}}_k^{v_j}\right), \ j \in \tilde{M}_k \ .$$

Again, the estimator is only aware of an approximation of the measurement noise $\mathbf{C}_k^{v_j} \geq \tilde{\mathbf{C}}_k^{v_j}$. It is worth mentioning here, that we distinguish between $a_{k|l}$ and a_k when it is necessary for clarification, where the subscript $k|l$ denotes the incorporation of measurement information up to time step l into a variable of time step k. We especially utilize this notation to differentiate between variables before filtering that are denoted with $k|k-1$ and those after filtering that are denoted with $k|k$.

Due to limited communication capabilities, a full-rate communication of estimates and measurements cannot be employed. Instead, we assume the distributed nodes to communicate irregularly and with limited bandwidth so that a local (pre)processing of measurements is necessary.

As even in linear information processing it is unknown how to simultaneously minimize the MSE matrix of locally and globally fused estimates when the data are calculated recursively over multiple time steps without running two estimators in parallel, one has to differentiate between an optimization of the estimate for distributed and for decentralized processing. The objective in distributed estimation as it is discussed in this paper is to optimize the MSE matrix at a data sink or fusion center while the MSE matrices of the local estimates are not of interest. This is different from decentralized estimation with the focus on locally optimal estimates.

In distributed estimation, the main challenge can be ascribed to the problem of track-to-track fusion, i.e., the distributed processing of measurements with a subsequent fusion of the estimation tracks under consideration of the common process noise. Although an optimal solution is known and is presented in Sec. IV, this algorithm suffers from tight constraints regarding the global knowledge of local measurement models.

III. 3-Layer Classification for Distributed Estimation

In this section, a generalized scheme for the linear distributed estimation is presented. The presented 3-layer classification allows a uniform handling of most distributed estimation algorithms – notably the DKF [15] and the proposed HKF – and may be seen as an extension to the classical idea of the 2-layer differentiation between the Model Layer and the Estimation Layer. The classification does not only provide an overview and the notation of distributed estimation, but also motivates an estimation process based on intermediate synthetic data. This is different from a direct recursive processing of estimates as it is common in the KF and other popular approaches [3].

The idea is summarized by means of an example in Fig. 1. Basically, we differentiate between the Model Layer, the Information Processing Layer, and the Estimation Layer. In the Model Layer, the actual system is described in terms of probabilistic models that are an abstraction of the reality. This description includes the real state, the measurements, and an exact description of the state transition and measurement models. These distributions and functions are unknown to the estimator. Only the concrete instances in form of measurements are observed. In the remainder of this paper, the values of the Model Layer are marked with a tilde.

Estimators operate on assumed uncertainties, models, etc. grouped in the Information Processing Layer. These variables are especially meaningful when the true uncertainties are overestimated and the models allow a consistent processing of estimates and measurements. It is worth mentioning that variables of the Information Processing Layer do not have to represent the real state. For example, it is conceivable that only the combination of multiple variables provides an

unbiased estimate or – in the extreme case – the variables do not even allow a direct access to the real state at all. As most of the occurring variables belong to this layer, we do not characterize them by a specific symbol.

The third layer is the Estimation Layer that contains the actual estimates and is the interface of an estimator. Here, the estimates as well as the MSE matrices of the different nodes are collected. These estimates are typically a direct derivation from the variables of the Information Processing Layer and are marked with a hat in the remainder of this paper.

As mentioned above, estimators operate (typically recursively) on data of the Information Processing Layer. A well-known example is the Information Filter that calculates the information mean and the information matrix based on the measurement and state transition models from the Information Processing Layer. The variables of the Estimation Layer are derived from the variables of the Information Processing Layer by inverting the information matrix, respectively multiplying the information mean with the covariance matrix. In order to prove the linear optimality of the Information Filter, which also holds for the KF, it is furthermore required that the measurement and state transitions models reflect the real situation and the noise terms are not a bound but exact.

The 3-layer classification is even more valuable when the DKF or the HKF is investigated. Then, the synthetic data of the Information Processing Layer represents biased information that must be transformed to obtain correct estimates.

IV. Distributed Kalman Filter

The DKF algorithm [15], [16], [18] is composed of modified KF prediction and filter formulas. Although this algorithm has been originally motivated by a decorrelation of local estimates through globalization, we present it here in slightly different form[1] and formulate the operations on the estimates as simple linear transformations in the context of a 3-layer classification.

Let the true state be distributed according to $\underline{\boldsymbol{x}}_k$. We assume the nodes $i \in \{1, \ldots, N\}$ to be initialized with uncorrelated estimates $\left(\hat{\underline{x}}_k^i, \hat{\mathbf{C}}_k^{x_i}\right)$ with

$$\mathrm{E}\left\{\left(\underline{\boldsymbol{x}}_k - \hat{\underline{x}}_k^i\right)^2\right\} =: \hat{\mathbf{P}}_k^{x_i} \leq \hat{\mathbf{C}}_k^{x_i} \ , \tag{2}$$

where $(\cdot)^2$ is a short version of $(\cdot)(\cdot)^{\mathrm{T}}$. In order to apply the DKF, these estimates are transformed to variables of the Information Processing Layer $\left(\underline{x}_k^i, \mathbf{C}_k^x\right)$ with

$$\begin{aligned} \underline{x}_k^i &= \mathbf{C}_k^x \left(\hat{\mathbf{C}}_k^{x_i}\right)^{-1} \hat{\underline{x}}_k^i \text{ and} \\ \mathbf{C}_k^x &= \left(\sum_{i=1}^{N} \left(\hat{\mathbf{C}}_k^{x_i}\right)^{-1}\right)^{-1} \ , \end{aligned} \tag{3}$$

where each $\underline{x}_k^i$ accounts for approximately $\frac{1}{N}$ of the estimate. The synthetic data (3) are predicted by applying the prediction

[1]The globalized covariance matrix is no longer multiplied by the number of nodes.

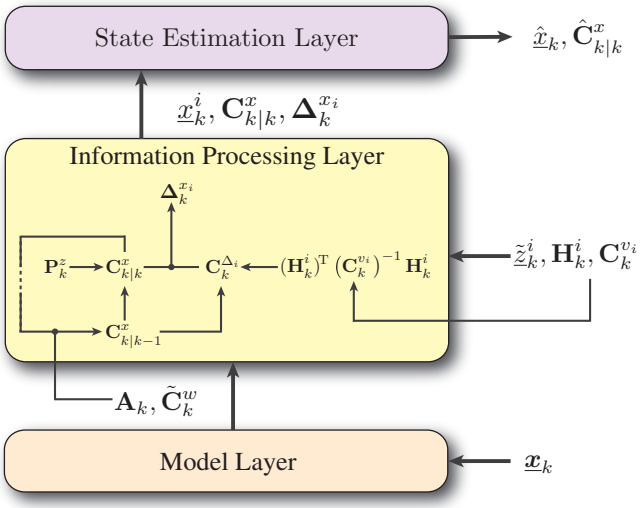

Fig. 1. A classification and overview of the local procedure of the HKF with the dependency structure of some occurring variables.

formulas

$$\begin{aligned} \underline{x}_{k+1}^i &= \mathbf{A}_k \underline{x}_k^i \text{ and} \\ \mathbf{C}_{k+1}^x &= \mathbf{A}_k \mathbf{C}_k^x \mathbf{A}_k^{\mathrm{T}} + \mathbf{C}_k^w \ . \end{aligned} \tag{4}$$

When all measurement models are known to all nodes, we can derive

$$\mathbf{C}_{k|k}^x = \left(\left(\mathbf{C}_{k|k-1}^x\right)^{-1} + \left(\mathbf{P}_k^z\right)^{-1}\right)^{-1} \tag{5}$$

with

$$\left(\mathbf{P}_k^z\right)^{-1} = \sum_{j \in \tilde{M}_k} \left(\mathbf{H}_k^j\right)^{\mathrm{T}} \left(\mathbf{C}_k^{v_j}\right)^{-1} \mathbf{H}_k^j \ . \tag{6}$$

This allows us to calculate

$$\underline{x}_{k|k}^i = \mathbf{K}_k \underline{x}_{k|k-1}^i + \mathbf{L}_k^i \tilde{\underline{z}}_k^i \tag{7}$$

with gains

$$\mathbf{K}_k = \mathbf{C}_{k|k}^x \left(\mathbf{C}_{k|k-1}^x\right)^{-1} \text{ and } \mathbf{L}_k^i = \mathbf{C}_{k|k}^x \left(\mathbf{H}_k^i\right)^{\mathrm{T}} \left(\mathbf{C}_k^{v_i}\right)^{-1} \ , \tag{8}$$

where $\mathbf{L}_k^i = \mathbf{0}\ \forall i \notin \tilde{M}_k$. As in general $\mathbf{K}_k \mathbf{H}_k^i \neq \mathbf{I} - \mathbf{L}_k^i$, this combination of $\underline{x}_k^i$ and $\tilde{\underline{z}}_k^i$ is biased.

As the synthetic variables $\underline{x}_k^i$ from the Information Processing Layer are biased and an associated MSE matrix is not given, the application of the DKF is only meaningful when the variables are combined according to[2]

$$\hat{\underline{x}}_k = \sum_{i=1}^{N} \underline{x}_k^i \ . \tag{9}$$

In this case, the fused estimate equals the result of the centralized KF, i.e., is globally optimal, and the associated MSE matrix is given by $\hat{\mathbf{C}}_k^x = \mathbf{C}_k^x$ [16].

By investigating this algorithm by means of a 3-layer classification it becomes clear that a distinction between the Information Processing Layer and the Estimation Layer is necessary. While the synthetic variables $(\underline{x}_k, \mathbf{C}_k^x)$ have no direct interpretation, only their combination yields an

[2]The fusion rule is slightly different from the one proposed in [16] as the explicit multiplicative variable that describes the number of nodes has been transformed into an implicit one.

unbiased estimate. Although it is possible to "pump" $\mathbf{C}_k^x$ up by the factor N and interpret the variables $\underline{x}_k^i$ as globalized uncorrelated estimates, this is a description of variables in a transformed state space that has not been used in literature for the derivation of further conclusions in the real state space.

The DKF requires distributed exact knowledge about the measurement models at all nodes in order to calculate the global measurement uncertainty $\left(\mathbf{P}_k^z\right)^{-1}$ from (6). Also, an unbiased estimate is only obtained by the fusion formula (9) when all variables $\underline{x}_k^i$ are available at the fusion center. This complicates the application in large scale sensor networks, where the requirements concerning the knowledge about remote measurement models and the availability of estimates are almost impossible to fulfill.

V. The Hypothesizing Distributed Kalman Filter

In this section, we propose the HKF that inherits the global optimality from the DKF when the assumption about the global measurement model is met and provides nearly optimal estimates otherwise. In the first place, we present the local procedure that has been proposed in a similar form as an extension to the DKF in [17]. The idea behind this approach is to derive and correct the bias that is induced by an erroneous assumption about the measurement model or by missing variables $\underline{x}_k^i$. In the second place, the fusion of estimates is discussed and a flexible fusion method is proposed that allows the combination of estimates based on recursively obtained variables.

A. Local Processing of the HKF

We derive a multiplicative correction matrix $\mathbf{\Delta}_k^{x_i}$ with $\mathrm{E}\left\{\left(\mathbf{\Delta}_k^{x_i}\right)^{-1}\underline{x}_k^i\right\} = \mathrm{E}\{\underline{\boldsymbol{x}}_k\}$. This allows us to obtain an estimate

$$\underline{\hat{x}}_k^i = \left(\mathbf{\Delta}_k^{x_i}\right)^{-1}\underline{x}_k^i \tag{10}$$

with a consistent bound $\hat{\mathbf{C}}_{k|k}^{x_i}$ that can be further utilized for fusion or control applications.

When the synthetic variable $\underline{x}_k^i$ has been initialized according to (3), it holds

$$\mathbf{\Delta}_k^{x_i} = \mathbf{C}_k^x\left(\hat{\mathbf{C}}_k^{x_i}\right)^{-1} .$$

Otherwise, we recursively obtain a correction matrix for each node by the following procedure.

The correction matrix for the calculation of an unbiased estimate of the next time step is obtained by

$$\mathbf{\Delta}_k^{x_i} = \mathbf{A}_k\mathbf{\Delta}_{k-1}^{x_i}\left(\mathbf{A}_k\right)^{-1} . \tag{11}$$

It can be easily verified that (10) is still an unbiased estimate and

$$\hat{\mathbf{C}}_k^{x_i} = \mathbf{A}_k\hat{\mathbf{C}}_{k-1}^{x_i}\left(\mathbf{A}_k\right)^{\mathrm{T}} + \mathbf{C}_k^w \tag{12}$$

is the MSE matrix of the bias-corrected estimate $\underline{\hat{x}}_k^i$ [17]. For the derivation of the filter step formulas, we introduce an auxiliary matrix

$$\mathbf{C}_k^{\Delta_i} = \left(\left(\mathbf{C}_{k|k-1}^x\right)^{-1}\mathbf{\Delta}_{k|k-1}^{x_i} + \left(\mathbf{H}_k^i\right)^{\mathrm{T}}\left(\mathbf{C}_k^{v_i}\right)^{-1}\mathbf{H}_k^i\right)^{-1} \tag{13}$$

and set with $\mathbf{C}_k^x$ from (5)

$$\mathbf{\Delta}_k^{x_i} = \mathbf{C}_k^x\left(\mathbf{C}_k^{\Delta_i}\right)^{-1} . \tag{14}$$

By applying the correction matrix to the estimate according to (10), the gains from (8) are modified in such a way that the matrix $\mathbf{C}_{k|k}^x$ is replaced by $\mathbf{C}_k^{\Delta_i}$ and therefore, the originally assumed global measurement uncertainty $\left(\mathbf{P}_k^z\right)^{-1}$ from (6) is replaced by the actually utilized measurement model $\left(\mathbf{H}_k^i\right)^{\mathrm{T}}\left(\mathbf{C}_k^{v_i}\right)^{-1}\mathbf{H}_k^i$. Consequently, even though the quality of the estimate will still depend on the difference between assumed and actual global measurement uncertainty, $\left(\mathbf{P}_k^z\right)^{-1}$ can now be seen as a design parameter that is referred to as $\left(\mathbf{C}_k^z\right)^{-1}$, hereafter. Nevertheless, it is inevitable for the proposed algorithm that this design parameter is the same in all nodes as otherwise the gain $\mathbf{K}_k^1$ from (8) would differ between the nodes.

Again, it is easily proven that (10) with the correction matrix from (14) is unbiased [17]. The MSE matrix $\hat{\mathbf{C}}_k^{x_i}$ is given by

$$\mathbf{C}_k^{\Delta_i}\left(\left(\mathbf{C}_{k|k-1}^x\right)^{-1}\mathbf{\Delta}_{k|k-1}^{x_i}\hat{\mathbf{C}}_{k|k-1}^{x_i}\left(\mathbf{\Delta}_{k|k-1}^{x_i}\right)^{\mathrm{T}}\left(\mathbf{C}_{k|k-1}^x\right)^{-\mathrm{T}} + \left(\mathbf{H}_k^i\right)^{\mathrm{T}}\left(\mathbf{C}_k^{v_i}\right)^{-1}\mathbf{H}_k^i\right)\left(\mathbf{C}_k^{\Delta_i}\right)^{\mathrm{T}} . \tag{15}$$

It is worth to point out that the MSE matrices in (12) and (15) are exact when the process noises and local measurement noises match the real ones, i.e., when the noise covariance matrices in the Information Processing Layer are not a bound but a correct description of the underlying system.

For the calculation of the unbiased estimate and its associated MSE matrix, it is necessary to invert the correction matrix. Although $\mathbf{\Delta}_k^{x_i}$ is in general regular, we have not been able to guarantee this and in addition, the inversion as well as the determination of the MSE matrix in (15) necessitates additional calculations. Therefore, it is reasonable to restrict the calculations to variables of the Information Processing Layer and to avoid the frequent computation of estimates – especially when the measurement rate is high and estimates are needed only rarely. This, however, comes at the expense that a tight MSE matrix bound can only be given when the assumed model matches the real one. Nevertheless, the estimation quality itself remains the same.

In summary, the local version of the HKF is still based on assumptions about the measurement models of all nodes. In contrast to the basic version from [16] it is, however, possible to derive a meaningful estimate in case of a mismatch between the assumed global measurement model and the actual one.

In [17], a version of this algorithm is proposed that allows the fusion of estimates and the subsequent calculation of the global correction matrix. Although it is not necessary to transmit all measurements between the distributed nodes when utilizing this algorithm, the calculation of the correction matrix requires information about the actual utilized models that must be transmitted to the fusion center. In the following, alternative approaches are discussed that allow the fusion of

distributed estimates based on recursively obtained variables only.

B. Fusion Approaches of the HKF

While the optimal fusion rule is known when the model assumptions are met, there are different approaches of fusing the data of the distributed nodes when this is not the case. These approaches will be discussed next, where the focus is laid on an algorithm that requires no information about the cross-covariance matrices between the estimates and provides the global optimal result when the model assumptions are met.

The optimal linear fusion algorithm that minimizes the MSE matrix of the fused estimate is the Bar-Shalom/Campo formula for two or Millman's formula for the general case with arbitrary many nodes. However, the application of these algorithms requires to maintain the complete joint covariance matrix, which in general is not possible in a distributed way. While the local covariance matrices can be maintained as described in V-A, the calculation of the cross-covariance matrices requires knowledge about the models of the other nodes that is typically not available. Therefore, this algorithm can be applied in strongly interlinked networks only where neither the communication nor the memory capacity is considerably limited by temporarily storing the models and transmitting these data to the other nodes.

A straightforward simplification of the optimal algorithm is to approximate or ignore the cross-covariance matrices and focus on the local uncertainties. Although such an approach is practical, it does not provide a globally optimal result, even when the model assumptions are correct. In addition, it is hard to provide tight quality bounds as the real MSE matrix of the fused result depends on the correlations, independently on how the estimates are combined.

We will focus on a different approach that is mainly based on synthetic variables of the Information Processing Layer. We present a simple fusion rule and derive attributes and bounds that allow a further processing of the fusion result similar to that of locally processed estimates.

We investigate the fusion by means of an arbitrary subset of nodes $G_k^f \subseteq \{1, \ldots, N\}$. This subset of nodes can, but is not limited to, be the collection of functioning nodes or a group of nodes in a hierarchical network. The fusion rule

$$\underline{x}_k^f = \sum_{i \in G_k^f} \underline{x}_k^i \tag{16}$$

is defined in the Information Processing Layer. It is the same as in (9) and is only slightly different from the one of the original DKF.

While this rule is sufficient when the model assumptions are correct, it does not have any practical meaning without the possibility of deriving an unbiased estimate from $\underline{x}_k^f$. Therefore, we present a procedure to derive the fused correction matrix from the local ones in the following. Let $\mathbf{\Delta}_k^{x_f}$ be the desired correction matrix. We set

$$\begin{aligned}\mathrm{E}\{\underline{\boldsymbol{x}}_k\} &= \left(\mathbf{\Delta}_k^{x_f}\right)^{-1} \mathrm{E}\left\{\underline{x}_k^f\right\} \overset{(16)}{=} \left(\mathbf{\Delta}_k^{x_f}\right)^{-1} \sum_{i \in G_k^f} \mathrm{E}\{\underline{x}_k^i\} \\ &\overset{(10)}{=} \left(\mathbf{\Delta}_k^{x_f}\right)^{-1} \sum_{i \in G_k^f} \mathbf{\Delta}_k^{x_i} \mathrm{E}\{\hat{\underline{x}}_k^i\}\end{aligned}$$

and obtain

$$\left(\mathbf{\Delta}_k^{x_f}\right)^{-1} \mathbf{\Delta}_k^{x_f} \mathrm{E}\{\underline{\boldsymbol{x}}_k\} = \left(\mathbf{\Delta}_k^{x_f}\right)^{-1} \Big(\sum_{i \in G_k^f} \mathbf{\Delta}_k^{x_i} \Big) \mathrm{E}\{\underline{\boldsymbol{x}}_k\}$$

with $\mathrm{E}\{\underline{\boldsymbol{x}}_k\} = \mathrm{E}\{\hat{\underline{x}}_k^i\}$. Therefore, the fused correction matrix is given by

$$\mathbf{\Delta}_k^{x_f} = \sum_{i \in G_k^f} \mathbf{\Delta}_k^{x_i} \,. \tag{17}$$

By means of (16) and (17), we are able to combine the variables from the Information Processing Layer and can derive an unbiased estimate from this data. Again, we cannot guarantee that the correction matrix $\mathbf{\Delta}_k^{x_f}$ is regular. But when the design parameter $(\mathbf{C}_k^z)^{-1}$ is chosen appropriately, the difference between the assumed and the actual model should be negligible and thus, the fused correction matrix is similar to the identity matrix, which is invertible.

VI. Attributes of the Proposed Algorithm

For further processing of the data, the quality of the estimate, i.e., the MSE matrix is of interest. We will now show that the MSE matrix, is the same as if the measurements of all time steps of all fused estimates would have been processed locally with the procedure from V-A.

A. Independence from Fusion Time

With

$$\begin{aligned}\underline{\boldsymbol{x}}_k - \left(\mathbf{\Delta}_k^{x_f}\right)^{-1} \underline{x}_k^f = \underline{\boldsymbol{x}}_k - \left(\mathbf{\Delta}_k^{x_f}\right)^{-1} \sum_{i \in G_k^f} \underline{x}_k^i = \\ \left(\mathbf{\Delta}_k^{x_f}\right)^{-1} \Big(\mathbf{\Delta}_k^{x_f} \underline{\boldsymbol{x}}_k - \sum_{i \in G_k^f} \underline{x}_k^i \Big) \,,\end{aligned}$$

which equals

$$\left(\mathbf{\Delta}_k^{x_f}\right)^{-1} \Big(\sum_{i \in G_k^f} \mathbf{\Delta}_k^{x_i} \left(\underline{\boldsymbol{x}}_k - \left(\mathbf{\Delta}_k^{x_i}\right)^{-1} \underline{x}_k^i \right) \Big) \,, \tag{18}$$

we derive the MSE matrix of the fused estimate

$$\hat{\mathbf{C}}_k^{x_f} = \mathrm{E}\left\{ \left(\underline{\boldsymbol{x}}_k - \left(\mathbf{\Delta}_k^{x_f}\right)^{-1} \underline{x}_k^f \right)^2 \right\} \tag{19}$$

as

$$\begin{aligned}\hat{\mathbf{C}}_k^{x_f} = \left(\mathbf{\Delta}_k^{x_f}\right)^{-1} \Big(\sum_{i,j \in G_k^f} \mathbf{\Delta}_k^{x_i} \mathrm{E}\Big\{ \left(\underline{\boldsymbol{x}}_k - \left(\mathbf{\Delta}_k^{x_i}\right)^{-1} \underline{x}_k^i \right) \cdot \\ \left(\underline{\boldsymbol{x}}_k - \left(\mathbf{\Delta}_k^{x_j}\right)^{-1} \underline{x}_k^j \right)^{\mathrm{T}} \Big\} \left(\mathbf{\Delta}_k^{x_j}\right)^{\mathrm{T}} \Big) \left(\mathbf{\Delta}_k^{x_f}\right)^{-\mathrm{T}} \,.\end{aligned}$$

This, however, is the bias-corrected sum over all MSE matrices and cross-covariance matrices

$$\left(\mathbf{\Delta}_k^{x_f}\right)^{-1} \Big(\sum_{i,j \in G_k^f} \mathbf{\Delta}_k^{x_i} \hat{\mathbf{C}}_k^{x_{ij}} \left(\mathbf{\Delta}_k^{x_j}\right)^{\mathrm{T}} \Big) \left(\mathbf{\Delta}_k^{x_f}\right)^{-\mathrm{T}} \,. \tag{20}$$

We now show that there is no difference between the processing of all measurements according to the local variant of the HKF and the distributed processing of a subset of measurements each with a subsequent fusion.

Prediction: With

$$\Big(\sum_{i \in G_k^f} \mathbf{A}_k \boldsymbol{\Delta}_k^{x_i} \left(\mathbf{A}_k\right)^{-1} \Big)^{-1} = \mathbf{A}_k \left(\boldsymbol{\Delta}_k^{x_f}\right)^{-1} \left(\mathbf{A}_k\right)^{-1} \ , \tag{21}$$

we simulate the local processing by predicting the variables from (20) according to (11) and (12) and obtain

$$\mathbf{A}_k \big(\boldsymbol{\Delta}_k^{x_f}\big)^{-1} \left(\mathbf{A}_k\right)^{-1} \Bigg(\sum_{i,j \in G_k^f} \mathbf{A}_k \boldsymbol{\Delta}_k^{x_i} \left(\mathbf{A}_k\right)^{-1} \Big(\mathbf{A}_k \hat{\mathbf{C}}_k^{x_{ij}} \left(\mathbf{A}_k\right)^{\mathrm{T}} + \mathbf{C}_k^w \Big) \left(\mathbf{A}_k\right)^{-\mathrm{T}} \left(\boldsymbol{\Delta}_k^{x_j}\right)^{\mathrm{T}} \left(\mathbf{A}_k\right)^{\mathrm{T}} \Bigg) \left(\mathbf{A}_k\right)^{-\mathrm{T}} \big(\boldsymbol{\Delta}_k^{x_f}\big)^{-\mathrm{T}} \left(\mathbf{A}_k\right)^{\mathrm{T}} \ .$$

This is simplified to

$$\mathbf{A}_k \left(\boldsymbol{\Delta}_k^{x_f}\right)^{-1} \Bigg(\sum_{i,j \in G_k^f} \boldsymbol{\Delta}_k^{x_i} \hat{\mathbf{C}}_k^{x_{ij}} \left(\boldsymbol{\Delta}_k^{x_j}\right)^{\mathrm{T}} + \sum_{i,j \in G_k^f} \boldsymbol{\Delta}_k^{x_i} \left(\mathbf{A}_k\right)^{-1} \mathbf{C}_k^w \left(\mathbf{A}_k\right)^{-\mathrm{T}} \left(\boldsymbol{\Delta}_k^{x_j}\right)^{\mathrm{T}} \Bigg) \big(\boldsymbol{\Delta}_k^{x_f}\big)^{-\mathrm{T}} \left(\mathbf{A}_k\right)^{\mathrm{T}} \ . \tag{22}$$

The process noise sum equals

$$\boldsymbol{\Delta}_k^{x_f} \left(\mathbf{A}_k\right)^{-1} \mathbf{C}_k^w \left(\mathbf{A}_k\right)^{-\mathrm{T}} \left(\boldsymbol{\Delta}_k^{x_f}\right)^{\mathrm{T}} \ ,$$

and so, (22) is transformed to

$$\mathbf{A}_k \big(\boldsymbol{\Delta}_k^{x_f}\big)^{-1} \Bigg(\sum_{i,j \in G_k^f} \boldsymbol{\Delta}_k^{x_i} \hat{\mathbf{C}}_k^{x_{ij}} \left(\boldsymbol{\Delta}_k^{x_j}\right)^{\mathrm{T}} \Bigg) \big(\boldsymbol{\Delta}_k^{x_f}\big)^{-\mathrm{T}} \left(\mathbf{A}_k\right)^{\mathrm{T}} + \mathbf{C}_k^w \ ,$$

which is the predicted fused MSE matrix $\mathbf{A}_k \hat{\mathbf{C}}_{k-1}^{x_f} \left(\mathbf{A}_k\right)^{\mathrm{T}} + \mathbf{C}_k^w$.

Filtering: We assume the nodes $\tilde{M}_k^f \subset \tilde{M}_k$ to have performed a filter step and obtain from (14) that

$$\boldsymbol{\Delta}_{k|k}^{x_f} = \sum_{i \in G_k^f} \boldsymbol{\Delta}_{k|k}^{x_i} = \sum_{i \in G_k^f} \mathbf{C}_{k|k}^x \left(\mathbf{C}_{k|k}^{\Delta_i}\right)^{-1} = \mathbf{C}_{k|k}^x \left(\left(\mathbf{C}_{k|k-1}^x\right)^{-1} \boldsymbol{\Delta}_{k|k-1}^{x_f} + \sum_{i \in \tilde{M}_k^f} \left(\mathbf{H}_k^i\right)^{\mathrm{T}} \left(\mathbf{C}_k^{v_i}\right)^{-1} \mathbf{H}_k^i \right)$$

holds, which is simplified to

$$\mathbf{C}_{k|k}^x \left(\left(\mathbf{C}_{k|k-1}^x\right)^{-1} \boldsymbol{\Delta}_{k|k-1}^{x_f} + \left(\mathbf{C}_k^{z_f}\right)^{-1} \right) \ , \tag{23}$$

with

$$\left(\mathbf{C}_k^{z_f}\right)^{-1} = \sum_{i \in \tilde{M}_k^f} \left(\mathbf{H}_k^i\right)^{\mathrm{T}} \left(\mathbf{C}_k^{v_i}\right)^{-1} \mathbf{H}_k^i \ .$$

Furthermore, the inner terms $\boldsymbol{\Delta}_{k|k}^{x_i} \hat{\mathbf{C}}_{k|k}^{x_{ij}} \left(\boldsymbol{\Delta}_{k|k}^{x_j}\right)^{\mathrm{T}}$ of the sum in (20) are transformed to

$$\mathbf{C}_{k|k}^x \left(\mathbf{C}_{k|k}^{\Delta_i}\right)^{-1} \hat{\mathbf{C}}_{k|k}^{x_{ij}} \left(\mathbf{C}_{k|k}^{\Delta_j}\right)^{-\mathrm{T}} \left(\mathbf{C}_{k|k}^x\right)^{\mathrm{T}} \ ,$$

which equals

$$\mathbf{C}_{k|k}^x \Bigg(\Big(\mathbf{C}_{k|k-1}^x\Big)^{-1} \boldsymbol{\Delta}_{k|k-1}^{x_i} \hat{\mathbf{C}}_{k|k-1}^{x_{ij}} \Big(\boldsymbol{\Delta}_{k|k-1}^{x_j}\Big)^{\mathrm{T}} \Big(\mathbf{C}_{k|k-1}^x\Big)^{-\mathrm{T}} + \delta_{ij} \left(\mathbf{H}_k^i\right)^{\mathrm{T}} \left(\mathbf{C}_k^{v_i}\right)^{-1} \mathbf{H}_k^i \Bigg) \left(\mathbf{C}_{k|k}^x\right)^{\mathrm{T}} \ ,$$

where $\delta_{ij} = 1$ for $i = j$ and 0 otherwise. Therefore, the sum $\sum_{i,j \in G_k^f} \boldsymbol{\Delta}_k^{x_i} \hat{\mathbf{C}}_k^{x_{ij}} \left(\boldsymbol{\Delta}_k^{x_j}\right)^{\mathrm{T}}$ is given by

$$\mathbf{C}_{k|k}^x \Bigg(\Big(\mathbf{C}_{k|k-1}^x\Big)^{-1} \Big(\sum_{i,j \in G_k^f} \boldsymbol{\Delta}_{k|k-1}^{x_i} \hat{\mathbf{C}}_{k|k-1}^{x_{ij}} \Big(\boldsymbol{\Delta}_{k|k-1}^{x_j}\Big)^{\mathrm{T}} \Big) \cdot \Big(\mathbf{C}_{k|k-1}^x\Big)^{-\mathrm{T}} + \left(\mathbf{C}_k^{z_f}\right)^{-1} \Bigg) \left(\mathbf{C}_{k|k}^x\right)^{\mathrm{T}} \ ,$$

which is simplified to

$$\mathbf{C}_{k|k}^x \Bigg(\Big(\mathbf{C}_{k|k-1}^x\Big)^{-1} \boldsymbol{\Delta}_{k|k-1}^{x_f} \hat{\mathbf{C}}_{k|k-1}^{x_f} \Big(\boldsymbol{\Delta}_{k|k-1}^{x_f}\Big)^{\mathrm{T}} \Big(\mathbf{C}_{k|k-1}^x\Big)^{-\mathrm{T}} + \left(\mathbf{C}_k^{z_f}\right)^{-1} \Bigg) \left(\mathbf{C}_{k|k}^x\right)^{\mathrm{T}} \ . \tag{24}$$

By help of (23) and (24) we obtain for (20)

$$\mathbf{C}_{k|k}^{\Delta_i} \Bigg(\Big(\mathbf{C}_{k|k-1}^x\Big)^{-1} \boldsymbol{\Delta}_{k|k-1}^{x_f} \hat{\mathbf{C}}_{k|k-1}^{x_f} \Big(\boldsymbol{\Delta}_{k|k-1}^{x_f}\Big)^{\mathrm{T}} \Big(\mathbf{C}_{k|k-1}^x\Big)^{-\mathrm{T}} + \left(\mathbf{C}_k^{z_f}\right)^{-1} \Bigg) \left(\mathbf{C}_{k|k}^{\Delta_i}\right)^{\mathrm{T}} \ ,$$

with

$$\mathbf{C}_{k|k}^{\Delta_f} = \left(\Big(\mathbf{C}_{k|k-1}^x\Big)^{-1} \boldsymbol{\Delta}_{k|k-1}^{x_f} + \left(\mathbf{C}_k^{z_f}\right)^{-1} \right)^{-1} \ ,$$

which equals the real MSE matrix (15) with all measurements processed locally.

In summary, the proposed fusion method provides an estimate that is equivalent to that where all measurements have been processed centrally when the local procedure of the HKF is employed. Consequently, the fusion result is globally optimal when the actually utilized measurement model has met the assumed one at all time steps. In this case, the fused correction matrix $\boldsymbol{\Delta}_{k|k}^{x_f}$ is the identity matrix.

B. MSE Matrix Bound of the Fused Estimate

Even though it has been shown that the fusion result is similar to the locally processed estimate in many respects, it remains to derive an MSE matrix bound for the fused estimate. The bound that has been derived recursively in (12) and (15) is not applicable directly on the MSE matrix of the fused estimate as the MSE matrix $\hat{\mathbf{C}}_k^{x_f}$ of the last time step, respectively before filtering, is unknown.

Instead, the MSE matrix can be calculated based on the local covariance matrices and the cross-covariance matrices as it has been derived in (20). At least the local covariance matrices can be transmitted together with the correction matrices and the estimates. The cross-covariance matrices, however, depend on multiple correction matrices and cannot be obtained at one local node. So, we investigate the

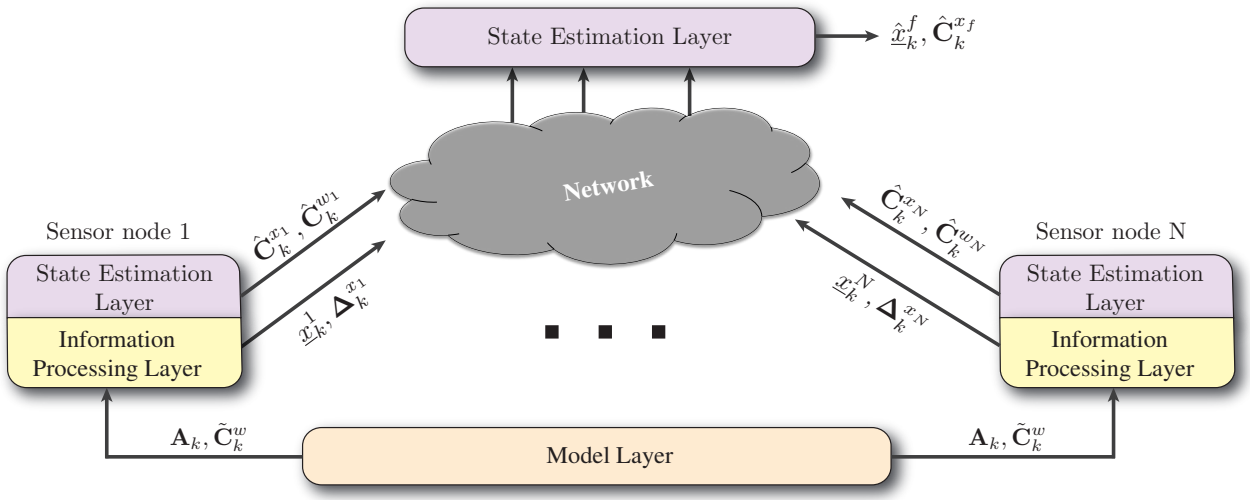

Fig. 2. A scheme describing the calculation and communication of variables for obtaining the fused estimate. The derivation of the MSE matrix bound of the fused estimate requires the transmission of variables of the Estimation Layer from the local nodes while this is not necessary for obtaining the estimate itself.

covariance matrices from (20) for $i \neq j$. In case of a prediction it holds

$$\boldsymbol{\Delta}_k^{x_i}\hat{\mathbf{C}}_k^{x_{ij}}\left(\boldsymbol{\Delta}_k^{x_j}\right)^{\mathrm{T}} = \mathbf{A}_{k-1}\boldsymbol{\Delta}_{k-1}^{x_i}\hat{\mathbf{C}}_{k-1}^{x_{ij}}\left(\boldsymbol{\Delta}_{k-1}^{x_j}\right)^{\mathrm{T}}(\mathbf{A}_{k-1})^{\mathrm{T}} + \boldsymbol{\Delta}_k^{x_i}\mathbf{C}_{k-1}^{w}\left(\boldsymbol{\Delta}_k^{x_j}\right)^{\mathrm{T}} \tag{25}$$

and for a filtering step we obtain by means of simple matrix algebra with a state independent measurement noise

$$\boldsymbol{\Delta}_{k|k}^{x_i}\hat{\mathbf{C}}_{k|k}^{x_{ij}}\left(\boldsymbol{\Delta}_{k|k}^{x_j}\right)^{\mathrm{T}} = \mathbf{C}_{k|k}^{x}\left(\mathbf{C}_{k|k-1}^{x}\right)^{-1}\cdot \boldsymbol{\Delta}_{k|k-1}^{x_i}\hat{\mathbf{C}}_{k|k-1}^{x_{ij}}\left(\boldsymbol{\Delta}_{k|k-1}^{x_j}\right)^{\mathrm{T}}\left(\mathbf{C}_{k|k-1}^{x}\right)^{-\mathrm{T}}\left(\mathbf{C}_{k|k}^{x}\right)^{\mathrm{T}}. \tag{26}$$

When we assume that the estimates were uncorrelated at time step 1, the cross-covariance matrices between the estimates of nodes i and j are therefore given by

$$\boldsymbol{\Delta}_k^{x_i}\hat{\mathbf{C}}_k^{x_{ij}}\left(\boldsymbol{\Delta}_k^{x_j}\right)^{\mathrm{T}} = \sum_{t=1}^{k}\mathbf{G}_t\boldsymbol{\Delta}_t^{x_i}\mathbf{C}_t^{w}\left(\boldsymbol{\Delta}_t^{x_j}\right)^{\mathrm{T}}(\mathbf{G}_t)^{\mathrm{T}}\ , \tag{27}$$

where the different $\mathbf{G}_t$ are products of state transition matrices and matrices from the Information Processing Layer that are known to all nodes. As we later refer to this derivation, it is worth mentioning that $\mathbf{G}_t$ is derived by recursively multiplying $\mathbf{A}_t$ to $\mathbf{G}_t$ according to (25) when a prediction step is performed and by multiplying $\mathbf{G}_t$ with $\mathbf{C}_{t|t}^{x}\left(\mathbf{C}_{t|t-1}^{x}\right)^{-1}$ according to (26) in case of a filter step.

Let G_k^f be defined as above, then the corrected MSE matrix from (20) is given by

$$\sum_{i\in G_k^f}\boldsymbol{\Delta}_k^{x_i}\hat{\mathbf{C}}_k^{x_i}(\boldsymbol{\Delta}_k^{x_i})^{\mathrm{T}} + \sum_{i,j\in G_k^f, i\neq j}\sum_{t=1}^{k}\mathbf{G}_t\boldsymbol{\Delta}_t^{x_i}\mathbf{C}_t^{w}\left(\boldsymbol{\Delta}_t^{x_j}\right)^{\mathrm{T}}(\mathbf{G}_t)^{\mathrm{T}}. \tag{28}$$

The second term is simplified with $\sum_{j\in G_k^f, j\neq i}\boldsymbol{\Delta}_t^{x_j} = \boldsymbol{\Delta}_t^{x_f} - \boldsymbol{\Delta}_t^{x_i}$ to

$$\sum_{t=1}^{k}\sum_{i\in G_k^f}\mathbf{G}_t\boldsymbol{\Delta}_t^{x_i}\mathbf{C}_t^{w}\left(\sum_{j\in G_k^f, j\neq i}\left(\boldsymbol{\Delta}_t^{x_j}\right)^{\mathrm{T}}\right)(\mathbf{G}_t)^{\mathrm{T}}\ ,$$

which equals

$$\sum_{t=1}^{k}\mathbf{G}_t\left(\boldsymbol{\Delta}_t^{x_f}\mathbf{C}_t^{w}\left(\boldsymbol{\Delta}_t^{x_f}\right)^{\mathrm{T}} - \sum_{i\in G_k^f}\boldsymbol{\Delta}_t^{x_i}\mathbf{C}_t^{w}\left(\boldsymbol{\Delta}_t^{x_i}\right)^{\mathrm{T}}\right)(\mathbf{G}_t)^{\mathrm{T}}\ . \tag{29}$$

Here, we present a consistent bound that can be obtained easily and does not require any additional assumptions or communication than the transmission of a local MSE matrix part. Let $\mathbf{R}_1, \mathbf{R}_2$ be two regular matrices and let $\mathbf{R}_3$ be positive definite. Then

$$\begin{aligned}\mathbf{R}_1\mathbf{R}_3\mathbf{R}_1^{\mathrm{T}} + \mathbf{R}_2\mathbf{R}_3\mathbf{R}_2^{\mathrm{T}} - \mathbf{R}_1\mathbf{R}_3\mathbf{R}_2^{\mathrm{T}} - \mathbf{R}_2\mathbf{R}_3\mathbf{R}_1^{\mathrm{T}} \\ = \mathbf{R}_1\mathbf{R}_3\left(\mathbf{R}_1 - \mathbf{R}_2\right)^{\mathrm{T}} + \mathbf{R}_2\mathbf{R}_3\left(\mathbf{R}_2 - \mathbf{R}_1\right)^{\mathrm{T}} \\ = \left(\mathbf{R}_1 - \mathbf{R}_2\right)\mathbf{R}_3\left(\mathbf{R}_1 - \mathbf{R}_2\right)^{\mathrm{T}}\end{aligned} \tag{30}$$

is positive definite. Following this idea, we obtain

$$\begin{aligned}|G_k^f|\sum_{i\in G_k^f}\boldsymbol{\Delta}_t^{x_i}\mathbf{C}_t^{w}\left(\boldsymbol{\Delta}_t^{x_i}\right)^{\mathrm{T}} - \boldsymbol{\Delta}_t^{x_f}\mathbf{C}_t^{w}\left(\boldsymbol{\Delta}_t^{x_f}\right)^{\mathrm{T}} = \\ \sum_{i\in G_k^f}\sum_{j\in G_k^f}\left(\boldsymbol{\Delta}_t^{x_i}\mathbf{C}_t^{w}\left(\boldsymbol{\Delta}_t^{x_i}\right)^{\mathrm{T}} - \boldsymbol{\Delta}_t^{x_i}\mathbf{C}_t^{w}\left(\boldsymbol{\Delta}_t^{x_j}\right)^{\mathrm{T}}\right) \geq \mathbf{0}\ .\end{aligned} \tag{31}$$

Thus, it holds

$$|G_k^f|\sum_{i\in G_k^f}\mathbf{G}_t\boldsymbol{\Delta}_t^{x_i}\mathbf{C}_t^{w}(\boldsymbol{\Delta}_t^{x_i})^{\mathrm{T}}(\mathbf{G}_t)^{\mathrm{T}} \geq \mathbf{G}_t\boldsymbol{\Delta}_t^{x_f}\mathbf{C}_t^{w}\left(\boldsymbol{\Delta}_t^{x_f}\right)^{\mathrm{T}}(\mathbf{G}_t)^{\mathrm{T}}\ ,$$

so that we finally obtain a consistent bound for the unknown part of (29) that is exact when the correction matrices are equally sized and rather conservative when the correction matrices differ significantly. When the nodes calculate and exchange $\hat{\mathbf{C}}_k^{x_i}$ and the synthetic matrices

$$\hat{\mathbf{C}}_k^{w_i} = \sum_{t=1}^{k}\mathbf{G}_t\boldsymbol{\Delta}_t^{x_i}\mathbf{C}_t^{w}\left(\boldsymbol{\Delta}_t^{x_i}\right)^{\mathrm{T}}(\mathbf{G}_t)^{\mathrm{T}}\ , \tag{32}$$

the true MSE matrix from (20) can be bounded by

$$\left(\boldsymbol{\Delta}_k^{x_f}\right)^{-1}\sum_{i\in G_k^f}\left(\boldsymbol{\Delta}_k^{x_i}\hat{\mathbf{C}}_k^{x_i}\left(\boldsymbol{\Delta}_k^{x_i}\right)^{\mathrm{T}} + \left(|G_k^f| - 1\right)\hat{\mathbf{C}}_k^{w_i}\right)\left(\boldsymbol{\Delta}_k^{x_f}\right)^{-\mathrm{T}}. \tag{33}$$

Summarizing the last section, we have proposed a new fusion method for the HKF that solely works on recursively obtained local matrices and provides the user with an estimate that is optimal when the model assumptions are correct and a MSE matrix bound that is tight under appropriate conditions and still consistent otherwise. The procedure is summarized in Fig. 2. While the derivation of a fused estimate (16) and the associated correction matrix (17) is straightforward and easy to implement, the derivation of the MSE matrix bound requires the local calculation and the exchange of the MSE matrix (12), (15) and a local synthetic cross-covariance matrix term (32). The bound itself, however, is easily obtained by (33).

VII. Conclusion

In this paper, we have embedded the DKF and its extension from [17] into a flexible and robust distributed estimation framework that we label the Hypothesizing Distributed Kalman Filter. We have demonstrated that the presented approaches require additional synthetic variables for which we have introduced the Information Processing Layer. In this course, we have proposed a 3-layer classification that simplifies the interpretation and notation of the new algorithm.

The HKF algorithm is based on a slightly modified version of the DKF that no longer requires to estimate the number of nodes in the network. The fusion algorithm itself requires only simple algebraic operations and is easy to implement. Its theoretic results are promising as the quality of the estimates corresponds to that of [17], which are equal to those of a centralized processing scheme when the model assumptions are set correctly. For the post-processing of estimates we also derived a consistent bound for the fused estimate that is based on recursively calculated matrices only. This bound has been shown to be tight under specific circumstances and to be consistent in any case.

In summary, we now have a distributed estimation algorithm that operates recursively and achieves globally optimal results when the design parameter is chosen accordingly. This algorithm allows to be applied to nonlinear systems by means of methods such as the Extended KF and can even be utilized when only a subset of estimates is exchanged. Also, the implementation in its basic version is straightforward and the communication effort is only slightly higher than in algorithms that operate on estimates only.

Future work will focus on finding other bounds for the fused estimate than the one proposed in this paper. Desirable is a tight bound that can be directly obtained from the correction matrix and the covariance matrix of the Information Processing Layer as this would reduce the computational as well as the communication effort. Also, the application of the HKF to decentralized scenarios is a future aim. As things stand at present, this is only possible for strongly interlinked networks where the common prior information can be discarded at every time step or in networks with a lot of computing power where only the locally acquired information are exchanged and each node predicts each available remote estimate until it can be replaced after a further data exchange. In order to provide a more effective procedure, the application of the Channel Filter idea is necessary.

Acknowledgements

This work was partially supported by the German Research Foundation (DFG) within the Research Training Group GRK 1194 “Self-organizing Sensor-Actuator-Networks” and by the Karlsruhe Institute of Technology (KIT).

References

[1] A. Makarenko and H. Durrant-Whyte, “Decentralized Data Fusion and Control in Active Sensor Networks,” in *Proceedings of the 7th International Conference on Information Fusion (FUSION 2004)*, 2004.

[2] Y. Bar-Shalom, “On the Track-to-Track Correlation Problem,” *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 571–572, 1981.

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

[4] V. Shin, Y. Lee, and T.-S. Choi, “Generalized Millman’s Formula and Its Application for Estimation Problems,” *Signal Processing*, vol. 86, no. 2, pp. 257–266, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168405001751

[5] J. Ajgl, M. Šimandl, and J. Dunik, “Millman’s Formula in Data Fusion,” in *Proceedings of the 10th International PhD Workshop on Systems and Control*, 2009, pp. 1–6.

[6] S. Grime and H. F. Durrant-Whyte, “Data Fusion in Decentralized Sensor Networks,” *Control Engineering Practice*, vol. 2, no. 5, pp. 849–863, 1994. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0967066194903492

[7] R. Olfati-Saber, J. A. Fax, and R. M. Murray, “Consensus and Cooperation in Networked Multi-Agent Systems,” *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[8] 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.

[9] Z. Duan and X. R. Li, “Optimal Distributed Estimation Fusion with Transformed Data,” in *Proceedings of the 11th International Conference on Information Fusion (FUSION 2008)*, Cologne, Germany, 2008.

[10] N. A. Carlson, “Federated Filter for Fault-tolerant Integrated Navigation Systems,” in *Proceedings of the Position Location and Navigation Symposium (PLANS’88)*. IEEE, 1988, pp. 110–119.

[11] ——, “Federated Square Root Filter for Decentralized Parallel Processors,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 26, no. 3, pp. 517–525, 1990.

[12] 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.

[13] ——, “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.

[14] U. D. Hanebeck and K. Briechle, “New Results for Stochastic Prediction and Filtering with Unknown Correlations,” in *Proceedings of the 2001 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2001)*. IEEE, 2001, pp. 147–152.

[15] W. Koch, “On Optimal Distributed Kalman Filtering and Retrodiction at Arbitrary Communication Rates for Maneuvering Targets,” in *Proceedings of the 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, 2008, pp. 457–462.

[16] F. Govaers and W. Koch, “On the Globalized Likelihood Function for Exact Track-to-Track Fusion at Arbitrary Instants of Time,” in *Proceedings of the 14th International Conference on Information Fusion (FUSION 2011)*, 2011, pp. 1–5.

[17] M. Reinhardt, B. Noack, and U. D. Hanebeck, “On Optimal Distributed Kalman Filtering in Non-ideal Situations,” in *Proceedings of the 15th International Conference on Information Fusion (FUSION 2012)*, 2012.

[18] W. Koch, “Exact Update Formulae for Distributed Kalman Filtering and Retrodiction at Arbitrary Communication Rates,” in *Proceedings of the 12th International Conference on Information Fusion (FUSION 2009)*, 2009, pp. 2209–2216.