

Advances in Hypothesizing Distributed Kalman Filtering

Marc Reinhardt, Benjamin Noack, and Uwe D. Hanebeck

Intelligent Sensor-Actuator-Systems Laboratory (ISAS)

Institute for Anthropomatics

Karlsruhe Institute of Technology (KIT), Germany

{marc.reinhardt,benjamin.noack,uwe.hanebeck}@ieee.org

Abstract—In this paper, linear distributed estimation is revisited on the basis of the hypothesizing distributed Kalman filter and equations for a flexible application of the algorithm are derived. We propose a new approximation for the mean-squared-error matrix and present techniques for automatically improving the hypothesis about the global measurement model. Utilizing these extensions, the precision of the filter is improved so that it asymptotically yields optimal results for time-invariant models. Pseudo-code for the implementation of the algorithm is provided and the lossless inclusion of out-of-sequence measurements is discussed. An evaluation demonstrates the effect of the new extensions and compares the results to state-of-the-art methods.

Index Terms—Kalman Filtering, Distributed Estimation, Sensor-networks, Track-to-Track Fusion (T2TF).

I. INTRODUCTION

Decentralized processing of data is one of the main challenges in today's strongly interconnected world. The collection of data at distributed sensor nodes about a specific phenomenon is an important aspect. Usually, the value of information decreases over time and when information stems from only a few sources, strong dependencies exist between different data sets. Due to these effects, challenges arise that must be taken into account when data sets are fused in order to harvest as much information as possible.

In this paper, the linear distributed processing of measurements of a common state over arbitrarily many time steps is investigated. In contrast to decentralized estimation, where the estimates are to be optimized at multiple nodes, the objective of distributed estimation is to linearly pre-process measurements at distributed nodes in order to optimize an estimate at one dedicated fusion node. Assuming that the measurements are conditionally independent given the state, an optimal procedure is the Kalman filter (KF) when all *measurements* are processed at one central node.

However, apart from scalability and reliability issues, the KF processing requires all measurements that are acquired at the distributed sensor nodes to be communicated. Therefore, based on seminal works of Wilner [1] and Bar-Shalom [2], the investigation of distributed estimation algorithms have become an important research topic.

One of the main directions in literature is to run KFs on the sensor nodes and to develop methods for combining the obtained estimates, which is referred to as track-to-track fusion (T2TF). The optimal fusion formulas in terms

of the mean-squared-error (MSE) have been proposed for two [3] and arbitrarily many [4] estimates. But, the application of these formulas requires precise information about cross-covariance matrices of the estimates. As the calculation of these cross-covariance matrices is computationally expensive and depends on model information of all nodes, alternative fusion formulas have been derived. For example, when all dependency information is discarded, fusion methods under unknown correlations [5], [6] can be applied.

Approaches that try to reach a consensus between the distributed nodes are investigated in [7]. A similar direction is taken by diffusion algorithms [8], [9] that combine spatially adjacent information and therefore, propagate information through the sensor-network over time.

Indeed, the estimation performance of T2TF approaches is worse than that of the central KF even if the minimum MSE (MMSE) *fusion* formulas are applied [10] as the processing of measurements with local KFs differs from the central processing. Hence, another direction in linear distributed estimation is the decomposition of the central KF in order to yield MMSE estimates. To the authors' knowledge, the first successful attempts in this direction have been made in [11] and [12]. Recently, this topic has regained attention in the context of Gaussian distributions with the derivation of the distributed Kalman filter [13], [14] that decorrelates the distributed estimates in order to allow an independent processing at the sensor nodes.

While the decomposition of the central KF works well for time-invariant measurement models, it becomes challenging otherwise as each node in the sensor-network must be aware of all other measurement models to perform the calculations. Hence, when information about the models is not made available to all sensors or when nodes are error-prone, only T2TF approaches can be applied that are accompanied with a loss of precision.

By proposing the hypothesizing distributed Kalman filter (HKF) [15], [16], we have presented an alternative approach that unites the flexibility of T2TF algorithms with the estimation performance of central KF decompositions. The fundamental idea is to utilize the central KF with a hypothesis about the measurement models instead of the actual models. As the hypothesis, in general, does not meet the actually utilized

models, we have derived correction matrices that compensate an erroneous hypothesis.

In this paper, we improve the basic version of the HKF [16] by two essential extensions. First, we derive an approximation of the MSE matrix in Sec. II and second, we propose a technique to improve the hypothesis about the global measurement model in Sec. III, based on data that is available at the fusion node anyway. In Sec. IV, we discuss the application of the HKF and present pseudo-code for sensor and fusion node algorithms. Finally, an evaluation that is inspired by the one in [15] is given in Sec. V.

A. Problem Formulation

We consider a sensor-network consisting of multiple nodes collected in a set \mathcal{G} . The sensors process measurements at discrete time steps and transmit recursively obtained data after arbitrarily many time steps to a dedicated fusion node. In particular, no assumptions about the network topology and communication frequency are made. Hence, at each time step, the data of a subset of sensors $\mathcal{G}_f \subseteq \mathcal{G}$ is available at the fusion node. We aim at estimating the state of a phenomenon that evolves according to the linear model

$$\mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \mathbf{w}_k, \quad (1)$$

with additive noise $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_k^w)$ and regular state transition matrix \mathbf{A}_k . The measurements obtained from the sensors are linearly related to the state

$$\tilde{z}_k^i = \mathbf{H}_k^i \mathbf{x}_k + \mathbf{v}_k^i, \quad i \in \mathcal{G} \quad (2)$$

with $\mathbf{v}_k^i \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_k^{v_i})$. All noise terms are assumed to be independent from each other. It is worth mentioning that the considerations in the remainder of this paper are not limited to Gaussian distributions. However, for other noise characteristics, we still aim at providing the best linear estimator in terms of the MSE.

Each node in the network has access to the system model (1) and to its local sensor model (2). The measurement models of other sensors are concealed as they are considered to be potentially time- and state-dependent. Thus, the global measurement model, which we define in anticipation of the next section as

$$(\mathbf{C}_k^z)^{-1} = \sum_{i \in \mathcal{G}_f} (\mathbf{H}_k^i)^\top (\mathbf{C}_k^{v_i})^{-1} \mathbf{H}_k^i, \quad (3)$$

is unknown to the sensors. The objective is to minimize the MSE of the combined estimate at the fusion node.

B. Motivation

As motivation, we investigate a scalar-valued system with $\mathbf{A}_0 = 1$, $\mathbf{C}_0^w = 1$, $\mathbf{H}_1^i = 1$, and $\mathbf{C}_1^{v_i} = 1$ for two sensors $i \in \{1, 2\}$. The initial estimates of the sensors are denoted by \hat{x}_0^i with MSE matrix $\hat{\mathbf{C}}_0^{x_i} = 1$. When local KFs are applied at the sensors, the estimates after one prediction and one filtering step are given (after some algebraic simplifications) by $\hat{x}_1^i = \frac{1}{3}\hat{x}_0^i + \frac{2}{3}\tilde{z}_1^i$. When all data, i.e., both initial estimates and both measurements, are processed with a central KF, the estimate is given by $\hat{x}_1^c = \frac{1}{8}\hat{x}_0^1 + \frac{1}{8}\hat{x}_0^2 + \frac{3}{8}\tilde{z}_1^1 + \frac{3}{8}\tilde{z}_1^2$. Note, that

the precision of the estimates depends on the ratios between initial estimates and measurements. As the KF is an MMSE estimator, the ratios are optimal in both cases, i.e., for the local optimization, the best ratio between initial estimates and measurements is $\frac{1}{3}$ to $\frac{2}{3}$, and for the global optimization $\frac{1}{8}$ to $\frac{3}{8}$ respectively. However, when only recursive information is sent from the sensors to the fusion node, the fused estimate must be a linear combination of \hat{x}_1^1 and \hat{x}_1^2 . Here, the main problem of distributed estimation arises: as the inner ratio between initial estimates and measurements in \hat{x}_1^i is one to two, the fused estimate is different from the globally optimal one with an inner ratio of one to three. In fact, analyzing the MMSE combination [3] of the local estimates demonstrates that the MSE is 3.4% higher than that of \hat{x}_1^c . For more than two time steps and sensors, the performance degradation becomes even more significant, as it will be seen in Sec. V.

Therefore, the challenge of linear distributed estimation is to locally find gains that equal those of the central KF. As these gains depend on state and measurement models only, they can be obtained locally when the global measurement model (3) is known to all nodes. Indeed, this kind of global knowledge is typically not available. Nevertheless, distributed algorithms can *hypothesize* about the global measurement model, e.g., each sensor could anticipate that, subsequently, its local estimate is combined with another one of the same type. Actually, given this hypothesis, the HKF provides the globally optimal estimate for the simple scalar scenario.

II. THE HYPOTHESIZING DISTRIBUTED KALMAN FILTER

As motivated in Sec. I-B, the challenge of distributed linear algorithms is to find matrix gains by which initial estimates and measurements of different time steps are combined. When the global measurement model (3) is unknown to the sensors, the matrix gains are, in general, restricted to approximations of the optimal ones. The idea of the HKF is to optimize the matrix gains according to a hypothesis that approximates the global measurement model, and to provide a fusion algorithm in order to subsequently compensate differences between the hypothesis and the actual global measurement model.

More precisely, each sensor runs a modified version of the KF that is optimized based on the hypothesis about the global measurement model (HGMM), which is given by the matrix $(\mathbf{P}_k^z)^{-1}$, instead of the local measurement model. In contrast to estimates, auxiliary variables \underline{x}_k^i are processed, which we refer to as pseudo-estimates. While a more detailed discussion about how to choose the HGMM follows in Sec. III, it is worth mentioning that the HGMM should meet the global measurement model (3). However, in general, the HGMM differs from the actual global measurement model. Thus, recursively obtained correction matrices $\Delta_k^{x_i}$ are calculated at the sensors that specify the difference between local measurement model and the HGMM. By combining correction matrices and pseudo-estimates, unbiased estimates are derived.

At the fusion node, multiple pseudo-estimates that have been optimized according to the same HGMM are combined

and a fused correction matrix is derived that allows the reconstruction of an unbiased combined estimate. Note that the fused correction matrix compensates the difference between the sum of local measurement models (3) and the HGMM. Taking into account that the fused estimate is optimal when this difference is zero [16], the HGMM must be optimized subject to the slowly changing measurement model of the entire sensor-network and not according to local measurement models.

In the following, we present formulas for the recursive calculation of local variables, and equations for their combination. Apart from a common value for the HGMM, the calculations are performed on local variables only. The key relation between pseudo-estimates and correction matrices that remains valid during all operations and, in particular, during the fusion, is that the transformed pseudo-estimate is unbiased, i.e.,

$$\mathbb{E}\{(\Delta_k^{x_i})^{-1}\underline{x}_k^i\} = \mathbb{E}\{\underline{x}\}, i \in \mathcal{G}. \quad (4)$$

When **initial** estimates $\hat{\underline{x}}_0^i$ with MSE matrices $\hat{\mathbf{C}}_0^{x_i}$ are available at the sensors, we set

$$\underline{x}_0^i = \Delta_0^{x_i} \hat{\underline{x}}_0^i, \Delta_0^{x_i} = \mathbf{C}_0^x (\hat{\mathbf{C}}_0^{x_i})^{-1} \text{ with } (\mathbf{C}_0^x)^{-1} = \sum_{i \in \mathcal{G}_f} (\hat{\mathbf{C}}_0^{x_i})^{-1}. \quad (5)$$

In case no initial estimates are given, we initialize the local variables with the first measurements \underline{z}_0^i and the HGMM, i.e., $(\mathbf{P}_0^z)^{-1}$, according to

$$\underline{x}_0^i = \mathbf{C}_0^x (\mathbf{H}_0^i)^\top (\mathbf{C}_0^{v_i})^{-1} \underline{z}_0^i, \Delta_0^{x_i} = \mathbf{L}_0^i \mathbf{H}_0^i \text{ with } (\mathbf{C}_0^x)^{-1} = (\mathbf{P}_0^z)^{-1}. \quad (6)$$

Note that by means of (6), a new pseudo-estimate can be obtained at any time as long as \mathbf{C}_k^x is up-to-date.

The **prediction** step of the local variables is realized by

$$\underline{x}_{k+1|k}^i = \mathbf{A}_k \underline{x}_k^i, \Delta_{k+1|k}^{x_i} = \mathbf{A}_k \Delta_k^{x_i} (\mathbf{A}_k)^{-1} \text{ and } \quad (7)$$

$$\mathbf{C}_{k+1|k}^x = \mathbf{A}_k \mathbf{C}_k^x (\mathbf{A}_k)^\top + \mathbf{C}_k^w, \quad (8)$$

where the subscript $k|k-1$ denotes predicted estimates. Note that for the prediction of the correction matrix, we require the state transition matrix \mathbf{A}_k to be regular.

When a measurement \underline{z}_k^i is obtained according to the linear model (2), it is included into the local pseudo-estimate in the **filtering** step by

$$\underline{x}_k^i = \mathbf{K}_k \underline{x}_{k|k-1}^i + \mathbf{L}_k^i \underline{z}_k^i, \Delta_k^{x_i} = \mathbf{K}_k \Delta_{k|k-1}^{x_i} + \mathbf{L}_k^i \mathbf{H}_k^i \quad (9)$$

$$\text{and } (\mathbf{C}_k^x)^{-1} = (\mathbf{C}_{k|k-1}^x)^{-1} + (\mathbf{P}_k^z)^{-1}, \quad (10)$$

where

$$\mathbf{K}_k = \mathbf{C}_k^x (\mathbf{C}_{k|k-1}^x)^{-1} \text{ and } \mathbf{L}_k^i = \mathbf{C}_k^x (\mathbf{H}_k^i)^\top (\mathbf{C}_k^{v_i})^{-1} \quad (11)$$

are modified gains of the KF that incorporate the HGMM instead of the local measurement model. It is worth pointing out that the HKF has been proposed for alternating prediction and filtering steps. When multiple measurements $1, \dots, m$ are obtained at one sensor at one time step, the pseudo-estimates

are given by

$$\underline{x}_k^i = \mathbf{K}_k \underline{x}_{k|k-1}^i + \sum_{j=1}^m \mathbf{L}_k^{i_j} \underline{z}_k^{i_j} \quad (12)$$

and the correction matrices by

$$\Delta_k^{x_i} = \mathbf{K}_k \Delta_{k|k-1}^{x_i} + \sum_{j=1}^m \mathbf{L}_k^{i_j} \mathbf{H}_k^{i_j}. \quad (13)$$

In this case, the global measurement model (3) needs to be adapted as well. When, on the other side, no measurement is obtained at a specific time step, the sum terms in (12) and (13) are omitted.

Based on the proposed formulas, the sensors are able to process measurements with the HKF. For the combination of multiple pseudo-estimates, \underline{x}_k^i and $\Delta_k^{x_i}$ are transmitted to the fusion node, and **fused** according to the simple sum formulas

$$\underline{x}_k^f = \sum_{i \in \mathcal{G}_f} \underline{x}_k^i \text{ and } \Delta_k^{x^f} = \sum_{i \in \mathcal{G}_f} \Delta_k^{x_i}, \quad (14)$$

where \mathcal{G}_f denotes the set of sensors that have transmitted data to the fusion node.

It is worth mentioning that the gains for the predicted estimates \mathbf{K}_k are the same at all nodes as they do not depend on the actually utilized measurement gain \mathbf{L}_k^i but only on the HGMM. This specific structure together with the fusion formulas (14) allows to subsequently combine measurements of the same time step from different sensors although only recursive variables are maintained. More precisely, let us without loss of generality assume that the local pseudo-estimates have been initialized from measurements with (6). Then, the fused pseudo-estimate is obtained by

$$\underline{x}_k^f = \sum_{i \in \mathcal{G}_f} \underline{x}_k^i = \sum_{i \in \mathcal{G}_f} \sum_{t=0}^k \underbrace{\mathbf{K}_k \mathbf{A}_{k-1} \cdots \mathbf{K}_{t+1} \mathbf{A}_t}_{\mathbf{G}_{t:k}} \mathbf{L}_t^i \underline{z}_t^i, \quad (15)$$

which equals

$$\underline{x}_k^c = \sum_{t=0}^k \mathbf{G}_{t:k} \sum_{i \in \mathcal{G}_f} \mathbf{L}_t^i \underline{z}_t^i, \quad (16)$$

while the second term is the pseudo-estimate of a node that has processed all measurements centrally with the HKF (note the ordering of the sums). A recursive representation of the correction matrix shows that the combination of correction matrices follows the same scheme. Hence, we derived a central Lemma that has been proven in detail in [16].

Lemma II.1 *The estimate gained by the HKF only depends on the information included and not on the time step when pseudo-estimates are combined or where measurements are processed, i.e.,*

$$p(\hat{\underline{x}}_k^{HKF} | \mathcal{Z}_0, \dots, \mathcal{Z}_k) \quad (17)$$

is uniquely determined, where \mathcal{Z}_t are sets containing all available initial estimates and measurements of time step t .

Based on the presented formulas we are able to estimate an unknown state. When the HGMM meets the global measurement model, the correction matrix is the identity matrix and the estimate equals the one obtained by the KF. However, in general, this is not the case and therefore, we are interested in quality attributes such as the MSE.

The MSE matrix of the HKF estimates is given by

$$\hat{\mathbf{C}}_k^{x_i} := \mathbb{E} \left\{ \left((\Delta_k^{x_i})^{-1} \underline{x}_k^i - \underline{x}_k \right)^2 \right\}, \quad (18)$$

with $(\cdot)^2 = (\cdot)(\cdot)^\top$. As for unbiased estimates (18) is determined by the model and covariance matrices of noise terms only, we present a recursive closed-form representation in the following. For this purpose, we introduce the pseudo-covariances $\hat{\mathbf{B}}_k^{x_i}$ and $\hat{\mathbf{B}}_k^{w_i}$ that account for local uncertainties, and common uncertainties respectively. The pseudo-covariances are maintained locally according to the formulas in Appendix A. Eventually, the MSE matrix (18) for HKF estimates is given by

$$\hat{\mathbf{C}}_k^{x_i} = (\Delta_k^{x_i})^{-1} (\hat{\mathbf{B}}_k^{x_i} + \hat{\mathbf{B}}_k^{w_i}) (\Delta_k^{x_i})^{-\top}, i \in \mathcal{G} \cup f. \quad (19)$$

When multiple pseudo-estimates are exchanged and the utilized models of the remote sensor nodes are made available to the fusion node, the exact MSE matrix is likewise obtained by the formulas from Appendix A. However, as this would require a lot of data to be transmitted, we want to determine $\hat{\mathbf{C}}_k^{x_f}$ from recursively obtained local variables only. In [16], it has been shown that $\hat{\mathbf{B}}_k^{x_f}$ is given as the sum of the local MSE matrices $\hat{\mathbf{B}}_k^{x_i}$ according to

$$\hat{\mathbf{B}}_k^{x_f} = \sum_{i \in \mathcal{G}_f} \hat{\mathbf{B}}_k^{x_i}. \quad (20)$$

The process noise term is obtained by

$$\hat{\mathbf{B}}_k^{w_f} = \sum_{t=0}^{k-1} \mathbf{G}_{t:k} \Delta_k^{x_f} (\mathbf{A}_k)^{-1} \mathbf{C}_k^{w_f} (\mathbf{A}_k)^{-\top} (\Delta_k^{x_f})^\top (\mathbf{G}_{t:k})^\top.$$

Remembering (14), $\Delta_t^{x_f} = \sum_{i \in \mathcal{G}_f} \Delta_t^{x_i}$, $\forall t \in \{0, \dots, k\}$ holds, and hence, $\hat{\mathbf{B}}_k^{w_f}$ is given by

$$\sum_{i,j \in \mathcal{G}_f} \sum_{t=0}^{k-1} \mathbf{G}_{t:k} \Delta_k^{x_i} (\mathbf{A}_k)^{-1} \mathbf{C}_k^{w_f} (\mathbf{A}_k)^{-\top} (\Delta_k^{x_j})^\top (\mathbf{G}_{t:k})^\top,$$

which in turn, is consistently¹ bounded [16] by

$$\hat{\mathbf{B}}_k^{w_f} \leq |\mathcal{G}_f| \sum_{i \in \mathcal{G}_f} \hat{\mathbf{B}}_k^{w_i}. \quad (21)$$

Even though this bound is meaningful for theoretical derivations, it lacks on practical relevance when the quality of the local pseudo-estimates differs considerably as in this case the process noise term of the MSE matrix is overestimated up to the factor $|\mathcal{G}_f|$. In the following, we propose an approximation of the MSE matrix that is not necessarily consistent, but aims at giving a precise value for the real MSE matrix.

It is well known that for time-invariant models, the KF matrices level off, i.e., the MSE matrix of the estimate as well as the Kalman gains converge to constant values [17]. As \mathbf{C}_k^x is obtained by means of KF equations and thus, the local gains and correction matrices are derivatives of constant matrices, it is meaningful to approximate them by time-independent matrices.

Therefore, with $\mathbf{K} = \mathbf{K}_k$, $\mathbf{A} = \mathbf{A}_k$, and $(\mathbf{K} \cdot \mathbf{A})^p = \mathbf{G}^p$, $\forall p \in \mathbb{N}_0$, the common process noise term $\hat{\mathbf{B}}_k^{w_f}$ is approximated by

$$\hat{\mathbf{B}}_k^{w_f} \approx \sum_{t=0}^{k-1} \mathbf{G}^{k-t} \underbrace{\Delta_k^{x_f} (\mathbf{A})^{-1} \mathbf{C}_k^{w_f} (\mathbf{A})^{-\top} (\Delta_k^{x_f})^\top}_{\mathbf{N}} (\mathbf{G}^{k-t})^\top. \quad (22)$$

From a computational point of view, the costs for evaluating (22) increase unbounded for $k \rightarrow \infty$. Thus, we split (22) into the sum terms

$$\sum_{t=0}^{l-1} \mathbf{G}^{k-t} \mathbf{N} (\mathbf{G}^{k-t})^\top + \sum_{t=l}^{k-1} \mathbf{G}^{k-t} \mathbf{N} (\mathbf{G}^{k-t})^\top. \quad (23)$$

With $\mathbf{N}' = \mathbf{G}^{k-l} \mathbf{N} (\mathbf{G}^{k-l})^\top$, we obtain for the first term

$$\begin{aligned} \left\| \sum_{t=0}^{l-1} \mathbf{G}^{t+1} \mathbf{N}' (\mathbf{G}^{t+1})^\top \right\| &\leq \sum_{t=0}^{l-1} \left\| \mathbf{G}^{t+1} \right\| \left\| \mathbf{N}' \right\| \left\| \mathbf{G}^{t+1} \right\| \\ &\leq \left\| \mathbf{N}' \right\| \left\| \mathbf{G} \right\|^2 \sum_{t=0}^{l-1} \left\| \mathbf{G} \right\|^{2t} = \left\| \mathbf{N}' \right\| \left\| \mathbf{G} \right\|^{2l} \frac{1 - \left\| \mathbf{G} \right\|^{2l}}{1 - \left\| \mathbf{G} \right\|^2}, \end{aligned} \quad (24)$$

for any matrix norm. Thus, a meaningful strategy to evaluate (22) for large k is to calculate the sum terms of the last $k-l$ time steps and approximate the remaining terms by means of (24). In order to guarantee a fixed precision, l should be decreased as long as the approximation (24) accounts for more than a fixed value of the overall MSE matrix from (22). Note that when l is decreased iteratively, \mathbf{N}' diminishes while most of the calculations, e.g., the calculation of $\left\| \mathbf{G} \right\|$, have to be performed only once.

Although, in general, neither the MSE matrix from (22) nor the evaluation of the term are exact, the provided MSE matrix is a good approximation of the real data. The reasons are twofold. First, the KF converges to the steady state quickly [17], which justifies the utilization of constant values in (22)—especially when they are based on multiple time steps such as the correction matrix. Second, the estimation quality depends foremost on the well-approximated last filtering steps, while the poorly approximated earlier filtering steps account for a fraction of the MSE matrix only. This is also measured by matrix norms in (24), where the evaluation of the MSE matrix stops for all tested scenarios after a few time steps even if the approximation error is forced to be less than one percent.

III. ADAPTIVE IMPROVEMENT OF HGMM

In the following, we propose a technique for the adaptive optimization of the HGMM. As the estimation quality depends on the difference between HGMM and actual global

¹The difference between bound and actual MSE matrix is positive semi-definite

measurement model (3), minimizing the difference leads to a lower MSE, which asymptotically reaches the globally optimal error for time-invariant models. In particular, the application of the adaptive optimization renders the approximation or pre-calculation of initial values for the HGMM obsolete.

A naïve adaption of the HGMM would be to calculate it based on the measurement models of the last time step or based on an average from an appropriate past period. However, this would require the matrices of the measurement models from all nodes to be transmitted to the fusion node. We propose an alternative technique that is based on the fused correction matrix and does not demand any additional transmissions except a broadcast message that contains the new HGMM.

With $\mathbf{G}_{t:k}$ from (15), the fused correction matrix is given by

$$\Delta_k^{x_f} = \sum_{t=0}^k \mathbf{G}_{t:k} \left(\sum_{i \in \mathcal{G}_f} \mathbf{L}_t^i \mathbf{H}_t^i \right) (\mathbf{A}_t)^{-1} \cdots (\mathbf{A}_{k-1})^{-1}. \quad (25)$$

As $(\mathbf{P}^z)^{-1}$, i.e., the HGMM, should meet the global measurement model $(\mathbf{C}_t^z)^{-1} = \sum_{i \in \mathcal{G}_f} (\mathbf{H}_t^i)^\top (\mathbf{C}_t^{v_i})^{-1} \mathbf{H}_t^i$, the optimal procedure is to model $(\mathbf{C}_t^z)^{-1}$ as an unknown state with an appropriate evolution model, estimate $(\mathbf{C}_t^z)^{-1}$ from past time steps, and find an HGMM that minimizes the difference to future states of $(\mathbf{C}^z)^{-1}$. As the evolution model of $(\mathbf{C}_t^z)^{-1}$ is in general nonlinear, state dependent, and unknown, we assume $(\mathbf{C}_t^z)^{-1}$ to be constant. This is a reasonable approximation as even major changes in local measurement models only have a minor effect on the global measurement model, which is the sum over all local models. Given this assumption, we obtain the linear equation

$$\Delta_k^{x_f} = \sum_{t=0}^k \mathbf{G}_{t:k} \mathbf{C}_t^x (\mathbf{C}^z)^{-1} (\mathbf{A}_t)^{-1} \cdots (\mathbf{A}_{k-1})^{-1}.$$

For such equations, a closed-form solution is available [18] and thus, a transformed global measurement model $\text{vec}(\mathbf{C}_t^x (\mathbf{C}^z)^{-1})$ is given by

$$\left(\sum_{t=0}^k \left((\mathbf{A}_{k-1})^{-\top} \cdots (\mathbf{A}_t)^{-\top} \right) \otimes \mathbf{G}_{t:k} \right)^{-1} \text{vec}(\Delta_k^{x_f}) \quad (26)$$

where \otimes denotes the Kronecker product and $\text{vec}(\cdot)$ denotes an operator that transforms a matrix into a row vector by concatenating all columns. Indeed, the computational load and the memory consumption for this method are high when k is large.

However, as motivated in the derivation of the MSE matrix approximation in Sec. II, the gains and uncertainty terms quickly reach a steady state. Again, we follow this argumentation and consider \mathbf{A}_k and \mathbf{K}_k to be constant. Under this assumption, we simplify the matrix sum from (26) and obtain with $(A^{k+1} \otimes G^{k+1}) = (A \otimes G)(A^k \otimes G^k)$ [18] that

$$\left(\sum_{t=0}^k (\mathbf{A}^{-\top})^{k-t} \otimes \mathbf{G}^{k-t} \right)^{-1} = \left(\sum_{t=0}^k (\mathbf{A}^{-\top} \otimes \mathbf{G})^t \right)^{-1}.$$

As the terms with $t \gg 0$ are negligible, we approximate the sum for $k \rightarrow \infty$ by the corresponding Neumann series that is given by

$$\left(\sum_{t=0}^{\infty} (\mathbf{A}^{-\top} \otimes \mathbf{G})^t \right)^{-1} = \left((\mathbf{I} - (\mathbf{A}^{-\top} \otimes \mathbf{G}))^{-1} \right)^{-1},$$

and obtain

$$\text{vec}(\mathbf{C}^x (\mathbf{C}^z)^{-1}) = \left(\mathbf{I} - (\mathbf{A}^{-\top} \otimes \mathbf{G}) \right) \text{vec}(\Delta_k^{x_f}), \quad (27)$$

from which we can extract $(\mathbf{C}^z)^{-1}$ in a computationally efficient way. Of course, the consideration of additional time steps in the Neumann series induces an approximation error. However, this error is negligible and is bounded by means of similar bounds as in Sec. II.

Note that for time-invariant system and measurement models, the assumptions about steady state matrices are asymptotically satisfied. Therefore, asymptotically, the global measurement model is met by the HGMM and the fused estimate equals the one of the central KF, which is the MMSE estimate.

IV. APPLYING THE HKF

After having presented and derived attributes of the HKF, we focus on the application in the following. In particular, we present pseudo-code for the application of the HKF, discuss the communication of variables, and suggest application strategies.

Data: $\mathbf{A}_k, \mathbf{C}_k^w, (\mathbf{P}_k^z)^{-1}, k_{begin}$

Result: $\underline{x}_k^i, \Delta_k^{x_i}, \hat{\mathbf{B}}_k^{x_i}, k_{cur}$

/ initialization */*

$k_{cur} \leftarrow k_{begin};$

Initialize $\underline{x}_k^i, \Delta_k^{x_i}, \mathbf{C}_k^x$ with (5) or (6);

Initialize $\hat{\mathbf{B}}_k^{x_i}, \hat{\mathbf{B}}_k^{w_i}$ with (31) or (32);

while *Sensor is active do*

/ prediction */*

$k_{cur} \leftarrow k_{cur} + 1;$

Predict $\underline{x}_k^i, \Delta_k^{x_i}, \mathbf{C}_k^x$ by means of $\mathbf{A}_k, \mathbf{C}_k^w$ with (7);

Predict $\hat{\mathbf{B}}_k^{x_i}, \hat{\mathbf{B}}_k^{w_i}$ by means of \mathbf{A}_k with (33);

/ filtering */*

Update $\underline{x}_k^i, \Delta_k^{x_i}, \mathbf{C}_k^x$ by means of $\mathbf{H}_k^i, \mathbf{C}_k^{v_i}, (\mathbf{P}^z)^{-1}$ with (9);

Update $\hat{\mathbf{B}}_k^{x_i}, \hat{\mathbf{B}}_k^{w_i}$ by means of $\mathbf{H}_k^i, \mathbf{C}_k^{v_i}, (\mathbf{P}^z)^{-1}$ with (35);

/ communication */*

if *IsConnected then*

Send $\underline{x}_k^i, \Delta_k^{x_i}, k_{cur}$, checksum(\mathbf{C}_k^x) to fusion node;

Send $\hat{\mathbf{B}}_k^{x_i}, \hat{\mathbf{B}}_k^{w_i}$ to fusion node;

Receive updates for $\mathbf{A}_k, \mathbf{C}_k^w, (\mathbf{P}_k^z)^{-1};$

end

end

Algorithm 1: Sensor node processing of the HKF

The operations to be performed at the sensors are summarized in Alg. 1. The variables $\mathbf{A}_k, \mathbf{C}_k^w$, and $(\mathbf{P}_k^z)^{-1}$ are potentially time-variant and thus, an effective implementation

is to model each of them as a buffer object containing matrices that are valid for consecutive time periods. The measurement model consisting of \mathbf{H}_k^i and $\mathbf{C}_k^{v_i}$ is considered to be available at the sensors. The gray variables in Alg. 1 are optional. The MSE matrix $\hat{\mathbf{B}}_k^{x_i}$ is only necessary when an MSE matrix is to be obtained at the fusion node. For the derivation of a consistent bound, $\hat{\mathbf{B}}_k^{w_i}$ must be maintained as well.

Furthermore, by means of a checksum of \mathbf{C}_k^x , the fusion node is able to verify that the HGMM and the utilized prediction models have been the same at all sensors. When a sensor fails to provide the correct checksum, a remote reset should be triggered from the fusion node.

The main task of the fusion node is to collect data from the sensors and extract an unbiased estimate. When the fusion node does not receive up-to-date estimates of all nodes, it is moreover meaningful to predict the missing pseudo-estimates from past time steps. The overall processing is summarized in Alg. 2. It is worth pointing out that two types of MSE matrices are calculated. The consistent bound that is obtained by means of (21) is only a tight approximation when all pseudo-estimates contain a similar portion of the overall information. In particular, when pseudo-estimates from different time steps are fused, this is often not the case and thus, the application of the—in terms of computation and communication—more efficient approximation based on (22) is reasonable.

Although the HKF in the proposed form is directly applicable², we think it is beneficial to discuss some properties and strategies regarding the application. In the first instance, it is worth mentioning that the distributed calculation of the HKF is basically a sum decomposition. Therefore, the effort in networks with a hierarchical multi-hop communication can be reduced by combining pseudo-estimates to fused pseudo-estimates at intermediate tree nodes and exchange only the fused data. Apart from the consistent MSE matrix bound (21), the result at the fusion node is exactly the same as when only sensor estimates are transmitted.

As indicated with the “IsConnected” function in Alg. 1, no communication constraints for applying the HKF exist. In particular, sensors can filter measurements at high frequency without precision loss even if communication links are established at low frequency. Instead of fixed communication intervals, it is also conceivable that estimates are exchanged when new measurements are obtained or the local MSE is lower than a designated bound.

Indeed, a communication strategy that provides an approximately constant global measurement model at the fusion node is preferable as the HGMM can be adapted to a fixed value then. When no feedback system is implemented, the value must be chosen offline in order to match the expected measurement quality (3) in the network as well as possible. However, in general, an adaption of the HGMM as it is described in Sec. III is more flexible and yields more precise estimates. Note that either way the HGMM must be the same at all sensors as

Data: $\mathbf{A}_k, \mathbf{C}_k^w, (\mathbf{P}_k^z)^{-1}, k_{begin}$

Result: $\hat{x}_k^f, \mathbf{C}_k^{x_f}, k_{cur}$

/ initialization */*

$k_{cur} \leftarrow k_{begin};$

$X \leftarrow \emptyset;$

while *Estimation process is running* **do**

$k_{cur} \leftarrow k_{cur} + 1;$

/ communication and preparation */*

while $x = \{\hat{x}_k^i, \Delta_k^{x_i}, k_{cur}^i, checksum(\mathbf{C}_k^x)\}$ is received **do**

if *checksum(\mathbf{C}_k^x) is incorrect* **then**

 Reset sensor and continue;

end

if $y \in X$ from the same sensor **then**

 If y sent earlier, remove y , else **continue**;

end

$x = x \cup \{\hat{\mathbf{B}}_k^{x_i}, \hat{\mathbf{B}}_k^{w_i}\};$

 Add x to X ;

end

 Predict all $x \in X$ by means of $\mathbf{A}_k, \mathbf{C}_k^w, (\mathbf{P}_k^z)^{-1}$ to k_{cur} ;

/ fusion */*

 Fuse $x \in X$ with (14) and derive estimate by (4);

 Derive MSE matrix by (19), (20), (21) and/or (22);

 Improve $(\mathbf{P}_k^z)^{-1}$ with (27) and communicate it to sensors;

end

Algorithm 2: Fusion node processing of the HKF

otherwise the pseudo-estimates cannot be combined. Hence, when a feedback system is employed, each application of an improvement of the HGMM at the local sensors is to be terminated at a future time step in order to guarantee that all sensors have obtained the new HGMM.

Finally, it is worth pointing out that sensors can be run in a triggered mode, i.e., they are only activated when a measurement is obtained, as the HKF is capable of processing the measurements by initializing pseudo-estimates with (6) without precision loss. In particular, an exact processing of out-of-sequence measurements is realized by this procedure when the initialized pseudo-estimates are also predicted to the current time step.

V. EVALUATION

In this section, the HKF is compared to state-of-the-art methods and the new adaptive improvement of the HGMM as well as the MSE matrix approximation are evaluated in a scenario that is inspired by the evaluation in [16].

We consider a network consisting of 64 sensors. The sensors are arranged in grid form in an area of 100×100 . A target moves with initial state $[50; 50; 1; 2]$ according to the disturbed

²A reference implementation of the HKF with the proposed extensions is given at <http://www.cloudrunner.eu/algorithm/109/>.

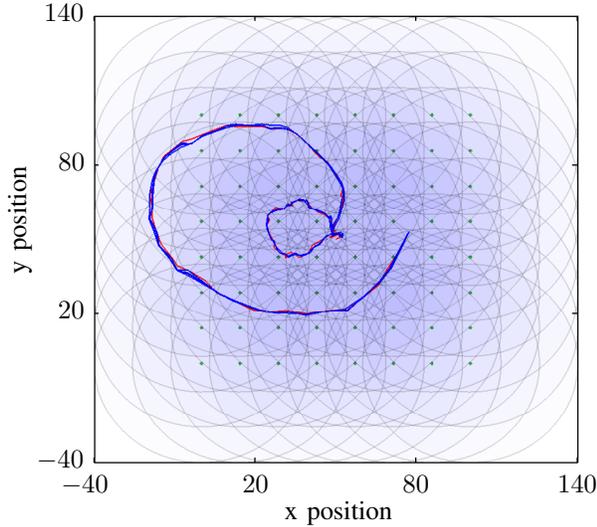


Figure 1. The sensor-network setup with the target path.

rotational motion model

$$\underline{\mathbf{x}}_{k+1} = \begin{pmatrix} 1 & 0 & \frac{\sin(\omega)}{\omega} & -\frac{1-\cos(\omega)}{\omega} \\ 0 & 1 & \frac{1-\cos(\omega)}{\omega} & \frac{\sin(\omega)}{\omega} \\ 0 & 0 & 1.02 \cdot \cos(\omega) & 1.02 \cdot \sin(\omega) \\ 0 & 0 & 1.02 \cdot \sin(\omega) & 1.02 \cdot \cos(\omega) \end{pmatrix} \underline{\mathbf{x}}_k + \underline{\mathbf{w}}_k \quad (28)$$

with $\underline{\mathbf{w}}_k \sim \mathcal{N}(\underline{\mathbf{0}}, \text{diag}(\mathbf{I}_2, 0.01 \cdot \mathbf{I}_2))$ and $\omega = 0.13$. Let \underline{p}^i denote the position of the i -th sensor and let \underline{p}_k^x denote the position of the target. We set the maximum sensor range to $d_{max} = 40$. The nodes obtain distance dependent measurements according to

$$\tilde{\mathbf{z}}_k^i = \underline{\mathbf{x}}_k + \underline{\mathbf{v}}_k^i \text{ with } \underline{\mathbf{v}}_k^i \sim \mathcal{N}(\underline{\mathbf{0}}, d(\underline{p}^i, \underline{p}_k^x) \cdot \mathbf{I}) , \quad (29)$$

with

$$d(\underline{p}^i, \underline{p}_k^x) = \begin{cases} 0 & , \text{ if } \|\underline{p}^i - \underline{p}_k^x\| > d_{max} \\ 10 \cdot \sqrt{\frac{1 + \|\underline{p}^i - \underline{p}_k^x\|}{1 + d_{max}}} & , \text{ else.} \end{cases} \quad (30)$$

The local estimates are initialized with an uncertainty of $1000 \cdot \mathbf{I}_4$ in order to focus on the measurement processing. The target is tracked over 100 time steps by five estimators. As a baseline, the central KF (CKF) processes all measurements at one node. The reference for T2TF algorithms is an estimator that utilizes KFs locally and fuses the estimates by means of MMSE combination formulas [4]. Note that this estimator is not realizable without knowledge about all measurement models at the fusion node as the cross-covariance matrices must be determined exactly. Nevertheless, the MSE of this approach defines a lower bound for all other T2TF algorithms. As a distributed estimation algorithm, the federated Kalman filter (FKF) [19] is employed. Finally, the HKF is employed with (HKF+) and without (HKF-) the adaptive improvement of the HGMM. The initial HGMM is set to the sum of the local measurement models with $d(\underline{p}^i, \underline{p}_k^x) = 1$. The HGMM is improved every tenth time step.

In Fig. 1, the target path and the considered sensor-network is given. As can be seen, the sensors cover the central area better

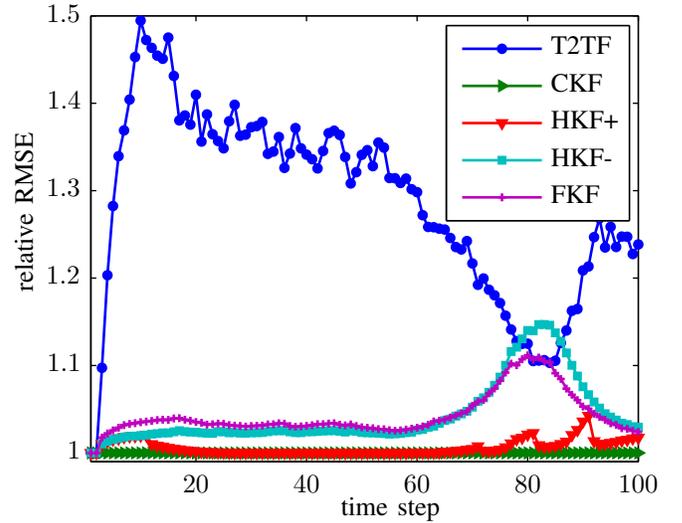


Figure 2. The root MSE of the estimators relative to the MMSE estimator for different time steps.

than the border areas. Thus, when the object path includes central and secluded positions as well, the global measurement model varies so that the performance of the HKF- suffers.

The root MSE (RMSE) for the different estimators relative to the root MMSE is compared in Fig. 2 for different time steps. Note that we compare the root MSE that are derived in closed form from the estimator characteristics, and not the errors of one Monte Carlo run.

As the target position is initially set to the center of the area, multiple sensors measure the state with low uncertainty and thus, the distributed estimators perform better than T2TF approaches. In time step 11, the adapted HGMM is utilized the first time. Therefore, the RMSE of the HKF+ with adaptive HGMM is almost optimal in subsequent time steps. In time steps 70 - 90, the target approaches the borders. As the sensor coverage in those areas is low, the HGMM fits the actually utilized models poorly and the number of sensing nodes in the FKF is dramatically overestimated. Hence, the relative RMSEs of the corresponding estimators increase while that of the T2TF decreases. The relative RMSE of the HKF+ becomes more sensitive to differences between the HGMM and the global measurement model, which causes a slight increase in the relative RMSE.

While T2TF algorithms perform well when only a small number of sensors is involved, the performance is significantly worse than distributed algorithms such as the HKF or the FKF when the target is measured by many nodes. In the given example, the RMSE is up to 50 percent higher than that of the other approaches. However, tests have even shown that the relative RMSE increases linearly with the number of nodes in the sensor-network.

The performance of the HKF- and that of the FKF are almost equal as the sensor nodes are assumed to be structural identical and thus, both approaches work by means of scalar changes of the relation between predicted state and measurements. In theory, a more sophisticated HGMM would improve the results of the HKF-, but also would have required a manual

tuning of the parameter. Instead, it is reasonable to employ the self-controlled adaptive improvement of the HGMM when no parameter should be chosen in advance. In fact, the results are almost optimal even if the global measurement model changes quickly.

A detailed evaluation of the MSE approximation from Sec.II is omitted due to space limitations. When the MSE matrices are compared with the approximations of the MSE, the deviation for the different time steps can be determined. As the maximum deviation between MSE matrix and approximation for HKF- over all time steps is 2,45% and the average deviation is 0.7%, the approximation of the MSE matches the real data quite well. For the HKF+ the average deviation still is only 3,36%. However, abrupt changes of the HGMM, which occur at time steps 71, 81, and 91 affect the calculation significantly so that the maximum deviation between MSE matrix and approximation is 21,8%. So, when a precise approximation of the MSE is necessary, (22) must be evaluated under consideration of the old HGMM.

In summary, distributed algorithms overcome T2TF approaches when multiple sensors are involved in the estimation process. When the measurement models are identical, the FKF and the HKF- perform almost equally for similar assumptions. In case the estimation error is to be minimized, the measurement models change significantly over the estimation period, or no tuning parameter should be chosen in advance, the adaptive HGMM technique provides a valuable extension to the HKF.

VI. CONCLUSIONS

We have extended the hypothesizing distributed Kalman filter (HKF) by an approximate calculation of MSE matrices and an adaptive technique to improve the hypothesis about the global measurement model. In particular, the latter one simplifies the application of the HKF significantly as no design parameters has to be chosen anymore. The results of the proposed extension are promising. In fact, for time-invariant models, the filter asymptotically approaches the globally optimal estimates.

Future work will focus on transferring the HKF to distributed *nonlinear* systems. In the linear context, it remains to show that the MSE monotonically improves when the hypothesis approaches the actual global measurement model. And last but not least, the HKF should be applied to real world systems.

APPENDIX

PSEUDO-COVARIANCE FORMULAS

Initialization from initial estimates:

$$\hat{\mathbf{B}}_0^{x_i} = \Delta_0^{x_i} \hat{\mathbf{C}}_k^{x_i} (\Delta_0^{x_i})^\top, \hat{\mathbf{B}}_0^{w_i} = \mathbf{0} \quad (31)$$

Initialization from measurements:

$$\hat{\mathbf{B}}_0^{x_i} = \mathbf{I}_0^i \mathbf{C}_0^{v_i} (\mathbf{I}_0^i)^\top, \hat{\mathbf{B}}_0^{w_i} = \mathbf{0} \quad (32)$$

Prediction:

$$\hat{\mathbf{B}}_{k+1|k}^{x_i} = \mathbf{A}_k \hat{\mathbf{B}}_k^{x_i} (\mathbf{A}_k)^\top, \quad (33)$$

$$\hat{\mathbf{B}}_{k+1|k}^{w_i} = \mathbf{A}_k \hat{\mathbf{B}}_k^{w_i} (\mathbf{A}_k)^\top + \Delta_{k+1|k}^{x_i} \mathbf{C}_k^w (\Delta_{k+1|k}^{x_i})^\top \quad (34)$$

Filtering:

$$\hat{\mathbf{B}}_k^{x_i} = \mathbf{K}_k \hat{\mathbf{B}}_{k|k-1}^{x_i} (\mathbf{K}_k)^\top + \mathbf{L}_k^i \mathbf{C}_k^{v_i} (\mathbf{L}_k^i)^\top, \quad (35)$$

$$\hat{\mathbf{B}}_k^{w_i} = \mathbf{K}_k \hat{\mathbf{B}}_{k|k-1}^{w_i} (\mathbf{K}_k)^\top \quad (36)$$

ACKNOWLEDGMENTS

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

REFERENCES

- [1] D. Willner, C. Chang, and K. Dunn, “Kalman Filter Algorithms for a Multi-sensor System,” *Proceedings of the 1976 IEEE Conference on Decision and Control*, pp. 1–5, 1976.
- [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.
- [5] 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.
- [6] J. Sijs, M. Lazar, and P. P. J. v. d. Bosch, “State Fusion with Unknown Correlation: Ellipsoidal Intersection,” in *Proceedings of the American Control Conference 2010*, 2010, pp. 3992–3997.
- [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] F. S. Cattivelli, S. Member, and A. H. Sayed, “Diffusion Strategies for Distributed Kalman Filtering and Smoothing,” *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2069–2084, 2010.
- [9] J. Hu, L. Xie, and C. Zhang, “Diffusion Kalman Filtering Based on Covariance Intersection,” *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 891–902, Feb. 2012.
- [10] S. Mori, K.-C. Chang, and C.-Y. Chong, “Essence of Distributed Target Tracking,” in *Distributed Data Fusion for Network-Centric Operations*, 2012, pp. 125–160.
- [11] H. R. Hashemipour, S. Roy, and A. J. Laub, “Decentralized structures for parallel Kalman filtering,” *IEEE Transactions on Automatic Control*, vol. 33, no. 1, pp. 88–94, 1988.
- [12] B. Rao and H. Durrant-Whyte, “Fully decentralised algorithm for multisensor Kalman filtering,” *IEE Proceedings D Control Theory and Applications*, vol. 138, no. 5, pp. 413–420, 1991.
- [13] 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.
- [14] F. Govaers and W. Koch, “An Exact Solution to Track-to-Track-Fusion at Arbitrary Communication Rates,” *Transactions on Aerospace and Electronic Systems*, vol. 48, no. 3, 2012.
- [15] 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.
- [16] —, “The Hypothesizing Distributed Kalman Filter,” in *Proceedings of the 2012 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2012)*, 2012.
- [17] G. Welch and G. Bishop, “An Introduction to the Kalman Filter,” *University of North Carolina at Chapel Hill, Chapel Hill, NC*, vol. 7, no. 1, 1995.
- [18] K. B. Petersen, “The Matrix Cookbook,” 2004.
- [19] N. A. Carlson, “Federated Square Root Filter for Decentralized Parallel Processors,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 26, no. 3, pp. 517–525, 1990.