Distributed Kalman Filtering in the Presence of Packet Delays and Losses

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

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

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

***Abstract*—Distributed Kalman filtering aims at optimizing an estimate at a fusion center based on information that is gathered in a sensor network. Recently, an exact solution based on local estimation tracks has been proposed and an extension to cope with packet losses has been derived. In this contribution, we generalize both algorithms to packet delays. The key idea is to introduce augmented measurement vectors in the sensors that permit the optimization of local filter gains according to time-dependent measurement capabilities at the fusion center. In the most general form, the algorithm provides optimized estimates in sensor networks with packets delays and losses. The precision depends on the actual arrival patterns, and the results correspond to those of the centralized Kalman filter when specific assumptions about the measurement capability are satisfied.**

Keywords—Kalman Filtering, Target Tracking, Distributed Estimation.

I. Introduction

In an increasingly interconnected world, the distributed collection and processing of information finds its way in more and more domains, ranging from military applications such as surveillance systems and target tracking to the decomposition of query operations in data centers.

Information is often noisy and must be processed by filter algorithms. Although density estimators such as particle filters [1], have been derived and applied to specific, often low-dimensional systems, the basis for current algorithms is still provided by simple linear approaches and, in particular, the Kalman filter (KF) [2]. Due to its flexibility, computational efficiency, and ease of application, the KF and its modifications are used in small mobile devices [3] as well as in huge data centers (e.g., for climate prediction [4]).

Distributing the calculation of a centralized KF to several (sensor) nodes carries additional challenges due to limited communication and knowledge. Not only is it necessary to process locally sensed measurements at the sensors without data from other sensors, but quite often also without access to the models and processing matrices at remote nodes. However, as measurements are recursively combined into tracks, local estimates contain all information provided by the corresponding sensing device up to the current time step.

A different approach is to transmit measurements instead of estimates. The naïve solution is to separately send all measurements (and the corresponding measurement models) from each sensor to the fusion center, which then calculates centralized KF gains for the processing of the received measurements. Apparently, this processing is more susceptible to packet losses than decomposition techniques as information is not re-sent in every time step. Still, a lot of research has been carried out in this regard for communication networks with stochastic properties [5]. For example, optimal data independent gains have been derived [6].

In the area of track-to-track fusion, techniques to quantify correlations [7] and to optimally fuse two [8] or several estimates [9], [10] have been derived when arbitrary estimators, e.g., KFs, are applied at the sensors. Unfortunately, the distributed calculation of correlations is challenging, taken by itself [11], and thus, approximate fusion methods [12] and fusion algorithms under unknown correlations have been proposed [13].

A holistic view on sensor networks with a fusion center has been taken in [14], [15] by exclusively considering the precision of the estimate at the fusion center while neglecting the quality of local estimates at the sensors. This has been realized by a decomposition of the centralized KF for different noise term characteristics.

Recently, the distributed KF (DKF) has been proposed as an optimal, i.e., centralized KF equivalent, estimator for sensor networks in the context of Gaussian distributions [16], [17]. The key idea is to optimize the local processing at the sensors according the centralized KF and to recover the optimal estimate from locally obtained vectors at the fusion node. As the DKF is only applicable when the communication is deterministic and all sensors have full access to the measurement models of remote nodes, the hypothesizing KF (HKF) has been derived to bypass these limitations [18].

In this paper, we aim at reconstructing the centralized KF results based on recursively obtained variables at the sensors. We generalize the DKF and the HKF by deriving formulas for the optimal inclusion of packets that are subject to communication delays. The proposed technique is based on a mixed processing at the sensors, consisting of a recursive part and an augmentation part that allows for filtering with optimal gains in all time steps. While the computational effort for the proposed approach increases proportionally to the maximum packet delay in the network, the communication effort is the same as for the basic versions of distributed and hypothesizing KF.

II. Problem Formulation

We consider a sensor network $\mathcal{S}$ consisting of a dedicated fusion center and sensors $s \in \mathcal{S}$ that observe a common phenomenon at discrete time steps. The packet transmissions from sensors to the fusion center are realized in a communication-delayed network, e.g., by means of multihop communication. The sensors in $\mathcal{S}$ are clustered according to communication delays $\tau \in \{0, \ldots, \overline{\tau}\}$ into sets $\mathcal{S}_0, \ldots, \mathcal{S}_{\overline{\tau}}$ with $\mathcal{S} = \mathcal{S}_0 \cup \cdots \cup \mathcal{S}_{\overline{\tau}}$. For example, in the multihop sensor network depicted in Fig. 1, the sensor specific communication delays are given by the number of hops to the fusion center.

Let the system and measurement models for $s \in \mathcal{S}$

$$\underline{\mathbf{x}}_{k+1} = \mathbf{A}\underline{\mathbf{x}}_k + \underline{\mathbf{w}}_k \text{ and} \tag{1}$$
$$\underline{z}_k^s = \mathbf{H}_k^s \underline{\mathbf{x}}_k + \underline{\mathbf{v}}_k^s \ , \tag{2}$$

be linear with zero-mean additive noise terms $\underline{\mathbf{w}}_k$ and $\underline{\mathbf{v}}_l^s$, respectively. The noise terms are assumed to be independent of each other with covariances $\mathrm{E}\{\underline{\mathbf{w}}_k(\underline{\mathbf{w}}_k)^\top\} = \mathbf{Q}_k$ and $\mathrm{E}\{\underline{\mathbf{v}}_k^s(\underline{\mathbf{v}}_k^s)^\top\} = \mathbf{R}_k^s$.

In the considered system, information is obtained by means of sensor measurements. Hence, we collect all information that is available to the fusion center at time step k in $\mathcal{I}^k = \{\underline{z}_t^s \mid s \in \mathcal{S}, t \leq k - \tau_s\}$. Let $\mathcal{T}_{k-t} = \mathcal{S}_0 \cup \cdots \cup \mathcal{S}_{k-t} \subseteq \mathcal{S}$ denote the set of sensors that have made the measurements of time step t available by time step k, i.e., the sensors $s \in \mathcal{S}$ with $\tau_s \leq k - t$.

The objective is to minimize the MSE of the estimate at the fusion node in each time step based on recursively calculated data at the sensors. The main challenge is to consider communication delays that impede the application of the DKF [17] and the HKF [18].

III. DKF for Time-delayed Information

In this section, we assume the delays to be deterministic, neglect packet losses, and require global model knowledge, i.e., all sensors have full knowledge about measurement models and uncertainties of all other sensors in the network. All of these assumptions are relaxed in Sec. IV.

In the first instance, we derive the centralized KF with delayed measurement processing and then, propose an approach to decompose the calculation. The centralized KF calculates the optimal estimate $\mathrm{E}\{\underline{\mathbf{x}}_k \mid \mathcal{I}^k\}$ in alternating prediction and filtering steps. The prediction of estimate $\underline{\hat{x}}_t$ and covariance $\mathbf{P}_t$ over one time step is given by

$$\underline{\hat{x}}_{t+1} = \mathbf{A}\underline{\hat{x}}_t \text{ and} \tag{3}$$
$$\mathbf{P}_{t+1} = \mathbf{A}\mathbf{P}_t\mathbf{A} + \mathbf{Q}_t \ . \tag{4}$$

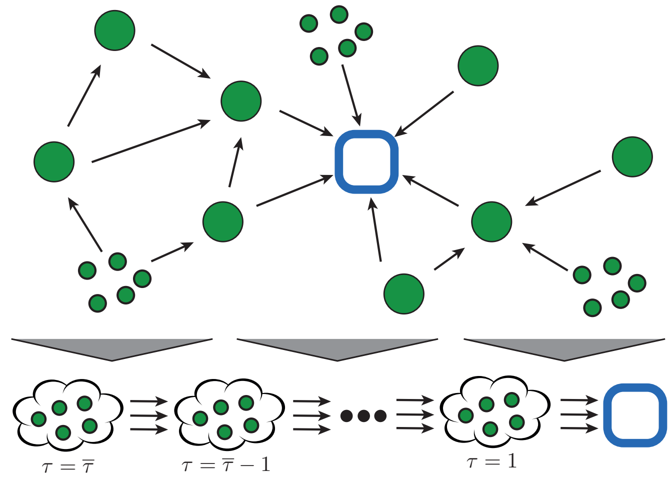

Figure 1. An arbitrary sensor network with sensor nodes (circles, green), fusion center (rounded rectangle, blue), and communications paths (arrows). At the bottom, a schematic clustering according to communication delays from sensors to the fusion center is given.

In the information form of the KF [19], measurements are filtered by

$$\underline{\hat{x}}_{t|t} = \mathbf{L}_t \underline{\hat{x}}_t + \sum_{s \in \mathcal{T}_{k-t}} \mathbf{K}_t^s \underline{z}_t^s \text{ with} \tag{5}$$
$$\left(\mathbf{P}_{t|t}\right)^{-1} = \left(\mathbf{P}_t\right)^{-1} + \left(\mathbf{P}_{t|k}^z\right)^{-1} \ , \tag{6}$$

where

$$\mathbf{L}_t = \mathbf{P}_{t|t}\left(\mathbf{P}_t\right)^{-1} \ , \tag{7}$$
$$\mathbf{K}_t^s = \mathbf{P}_{t|t}(\mathbf{H}_t^s)^\top \left(\mathbf{R}_t^s\right)^{-1} \ , \text{ and} \tag{8}$$
$$\left(\mathbf{P}_{t|k}^z\right)^{-1} = \sum_{s \in \mathcal{T}_{k-t}} (\mathbf{H}_t^s)^\top \left(\mathbf{R}_t^s\right)^{-1} \mathbf{H}_t^s \ . \tag{9}$$

We have chosen to present the KF equations in the information form as this allows an easy integration of multiple measurements at one time step. Note that the gains $\mathbf{L}_t$ and $\mathbf{K}_t^s$ implicitly depend on the *measurement capabilities* $\left(\mathbf{P}_{t|k}^z\right)^{-1}$ in (9), $t \leq k$, and therefore on which measurements are available to the fusion center at time step k.

The following observation follows directly from the communication structure

Observation 1 *Let $t < k - \overline{\tau}$, then $\mathcal{T}_{k-t} = \mathcal{T}_{k-t-1}$ and thus, the gains $\mathbf{L}_t$, $\mathbf{K}_t^s$, and the estimates $\underline{\hat{x}}_t$ with $t \leq k - \overline{\tau}$ do not change in time steps $k' \geq k$.*

Following Observation 1, we propose a mixed algorithm consisting of a recursive and a state augmentation part. Information up to time step $k - \overline{\tau}$ is stored in the estimate $(\underline{\hat{x}}_{k-\overline{\tau}}, \mathbf{P}_{k-\overline{\tau}})$. The gains (7) and (8) for $t > k - \overline{\tau}$ ensue from $\mathbf{P}_{k-\overline{\tau}}$ with $\left(\mathbf{P}_{k-\overline{\tau}+1|k}^z\right)^{-1}, \ldots, \left(\mathbf{P}_{k|k}^z\right)^{-1}$. In order to highlight temporary quantities, which are discarded at every time step, we use a tilde and a reference time step k, i.e., $\tilde{\mathbf{P}}_t^k$, $\tilde{\mathbf{L}}_t^k$, and $\tilde{\mathbf{K}}_t^s$ for $t > k - \overline{\tau}$. Taken together, the covariances depicted in Tab. I must be calculated at every time step.

k	$\mathbf{P}^{k}_{k-\tau}$	$\tilde{\mathbf{P}}^{k}_{k-\tau+1}$	$\tilde{\mathbf{P}}^{k}_{k-\tau+2}$	$\dots$	/	/
$k{+}1$	/	$\mathbf{P}^{k+1}_{k-\tau+1}$	$\tilde{\mathbf{P}}^{k+1}_{k-\tau+2}$	$\dots$	$\tilde{\mathbf{P}}^{k+1}_{k+1}$	/
$k{+}2$	/	/	$\mathbf{P}^{k+2}_{k-\tau+2}$	$\dots$	$\tilde{\mathbf{P}}^{k+2}_{k+1}$	$\tilde{\mathbf{P}}^{k+2}_{k+2}$

Table I. COVARIANCES THAT ARE NEEDED FOR THE GAIN CALCULATION AT TIME STEPS k TO $k+2$.

We can "roll out" the recursive calculations in (3) and (5) for $t > k-\overline{\tau}$ and obtain a linear combination of measurements and the estimate $\underline{\hat{x}}_{k-\overline{\tau}}$

$$\mathrm{E}\{\underline{\mathbf{x}}_k \mid \mathcal{I}^k\} = \sum_{t=k-\overline{\tau}+1}^{k} \sum_{s\in\mathcal{T}_{k-t}} \tilde{\mathbf{N}}^s_{t|k}\underline{z}^s_t + \tilde{\mathbf{N}}^s_{k-\overline{\tau}|k}\underline{\hat{x}}_{k-\overline{\tau}} \quad (10)$$

with

$$\tilde{\mathbf{N}}^s_{t|k} = \begin{cases} \mathbf{A}\tilde{\mathbf{L}}_{k|k}\cdots\mathbf{A}\tilde{\mathbf{L}}_{t+1|k}\mathbf{A}\tilde{\mathbf{K}}^s_{t|k} & , t > k-\overline{\tau} \\ \mathbf{A}\tilde{\mathbf{L}}_{k|k}\cdots\mathbf{A}\tilde{\mathbf{L}}_{t+1|k}\mathbf{A} & , t = k-\overline{\tau} \end{cases} . \quad (11)$$

Unfortunately, due to the delayed inclusion of measurements, the temporary covariances, in general, do not transfer to the covariances in subsequent time steps. For example, $\tilde{\mathbf{P}}^k_{k-\tau+1}$ differs from $\mathbf{P}^{k+1}_{k-\tau+1}$ as the latter one is obtained by means of the measurement capability $\left(\mathbf{P}^z_{k-\tau+1|k+1}\right)^{-1} \geq \left(\mathbf{P}^z_{k-\tau+1|k}\right)^{-1}$. However, when the communication network and the system models are time-invariant, the covariance $\mathbf{P}_{k-\overline{\tau}}$ converges to a steady-state matrix, the measurement capabilities $\left(\mathbf{P}^z_{t|k}\right)^{-1}$ exclusively depend on the difference between t and k, and thus, the gains (7), (8), and (11) are constants with respect to $k-t$.

It is worth mentioning that even if the gains $\tilde{\mathbf{K}}^s_{t|k}$ converge to steady-state values, it holds in general $\tilde{\mathbf{K}}^s_{t_1|k} \neq \tilde{\mathbf{K}}^s_{t_2|k}$ and $\tilde{\mathbf{K}}^s_{t_1|k+1}\left(\mathbf{A}\tilde{\mathbf{K}}^s_{t_1|k}\right)^{-1} \neq \tilde{\mathbf{K}}^s_{t_2|k+1}\left(\mathbf{A}\tilde{\mathbf{K}}^s_{t_2|k}\right)^{-1}$ for $t_1 \neq t_2$, $t_1, t_2 \in \{t+1,\ldots,k\}$, which, in turn, prevents a recursive processing of measurements.

Decomposition of centralized KF Equations

In the following, the derived formulas are decomposed in order to allow a distributed processing of the estimate. We use the same technique as in [17] and define the local processing based on centralized KF equations. Remember that the maximum packet delay in the sensor network is denoted by $\overline{\tau}$. We introduce a local quantity $\underline{\hat{x}}^s_{k-\overline{\tau}}$ at each sensor that recursively comprises measurements up to time step $k-\overline{\tau}$. For $t \leq k-\overline{\tau}$, the gains $\mathbf{L}_t$ and $\mathbf{K}^s_t$ are obtained as in the centralized KF and the quantity $\underline{\hat{x}}^s_{k-\overline{\tau}}$ is recursively calculated according to

$$\underline{\hat{x}}^s_0 = \frac{1}{|\mathcal{S}|}\underline{\hat{x}}_0 \ , \quad (12)$$

$$\underline{\hat{x}}^s_{t+1} = \mathbf{A}\underline{\hat{x}}^s_t \ \text{, and} \quad (13)$$

$$\underline{\hat{x}}^s_{t|t} = \mathbf{L}_t\underline{\hat{x}}^s_t + \mathbf{K}^s_t\underline{z}^s_t \ . \quad (14)$$

In general, $\mathbf{L}_t \neq \mathbf{I} - \mathbf{K}^s_t\mathbf{H}^s_t$ and thus, $\underline{\hat{x}}^s$ is a biased estimate. However, comparing it to $\underline{\hat{x}}_{k-\overline{\tau}}$ in (3) and (5) reveals the property

$$\underline{\hat{x}}_{k-\overline{\tau}} = \sum_{s\in\mathcal{S}} \underline{\hat{x}}^s_{k-\overline{\tau}} \ , \quad (15)$$

i.e., we are able to reconstruct the centralized KF estimate of time step $k-\overline{\tau}$ based on variables that have been gained recursively at the sensors. An initialization based on local estimates or measurements instead of a decomposition of the centralized estimate is feasible as long as (15) is ensured (c.f. [20]).

For time steps $t > k-\overline{\tau}$, temporary gains (7) and (8) are utilized that are optimized according to the measurement capability at the fusion node. As packets from sensor s are affected by a communication delay of τ_s, information that will be available at the fusion center at time step k must be transmitted by time step $k-\tau_s$ already. Hence, we define temporary estimates $\underline{\tilde{y}}^s_k$ reflecting the information of time step $k-\tau_s$ that will be available to the fusion center in time step k as

$$\underline{\tilde{y}}^s_k = \tilde{\mathbf{N}}^s_{\overline{\tau}-1:\tau_s|k}\ \underline{z}^s_{\overline{\tau}-1:\tau_s|k} + \mathbf{N}^s_{k-\overline{\tau}|k}\underline{\hat{x}}^s_{k-\overline{\tau}} \quad (16)$$

with

$$\begin{aligned} \tilde{\mathbf{N}}^s_{\overline{\tau}-1:\tau_s|k} &= \left(\tilde{\mathbf{N}}^s_{k-\overline{\tau}+1|k} \quad \ldots \quad \tilde{\mathbf{N}}^s_{k-\tau_s|k}\right) \text{ and} \\ \underline{z}^s_{\overline{\tau}-1:\tau_s|k} &= \left((\underline{z}^s_{k-\overline{\tau}+1})^\top \quad \ldots \quad (\underline{z}^s_{k-\tau_s})^\top\right)^\top . \end{aligned} \quad (17)$$

Note that $\underline{\tilde{y}}^s_k$ is calculated at node s in time step $k-\tau_s$ by means of gains $\tilde{\mathbf{N}}^s_{\overline{\tau}-1:\tau_s|k}$ that depend on covariances and the communication of time steps $t > k-\tau_s$. Again, the proposed method is a decomposition of the centralized KF equations. We obtain

$$\underline{\hat{x}}_k = \mathrm{E}\{\underline{\mathbf{x}}_k \mid \mathcal{I}^k\} = \sum_{s\in\mathcal{S}} \underline{\tilde{y}}^s_k \quad (18)$$

as the combination rule at the fusion center.

The proposed technique for the distributed calculation of the centralized KF estimate in the presence of deterministic communication delays is summarized in Alg. 1. For $\overline{\tau} = 0$, the equations simplify to the recursive processing in (13) and (14) with fusion method (15) and hence, are the same as the processing of the DKF [17]. For $\overline{\tau} > 0$, a state augmentation technique (17) is utilized to facilitate a remote processing of measurements that is adapted to the measurement capability at the fusion node.

Alternatively to the proposed processing, it is conceivable to calculate $\underline{\tilde{y}}^s_{k-\tau_s}$, i.e., estimates of time step $k-\tau_s$ instead of k at the sensors. The forward projection from time step $k-\tau_s$ to k is then done at the fusion center by multiplying $\underline{\tilde{y}}^s_{k-\tau_s}$ with $\mathbf{A}\tilde{\mathbf{L}}_{k|k}\cdots\mathbf{A}\tilde{\mathbf{L}}_{k-\tau_s+1|k}$, $t > k-\overline{\tau}$ (c.f. (11)). The estimate yielded by this technique still corresponds to the centralized KF result, but instead of repeating the computations at all sensors, the forward projection gains are obtained only once at the fusion node.

For both variants, the computational effort depends on the communication delay between sensor and fusion node. More precisely, at node s, the number of gains $\tilde{\mathbf{N}}^s_{t|k}$ calculated per

Algorithm 1 Sensor Processing

```
Initialize x̂_0^s with (12)
for k = 0 ; ... ; k = k + 1 do
    Filter P^s_{k-τ̄}, x̂^s_{k-τ̄} with (6) and (14)
    Add new measurement to z^s_{τ̄:τ_s|k} and remove z^s_{k-τ̄} from list (if k - τ̄ ≥ 0)
    Set P̃^s_{k-τ̄} = P^s_{k-τ̄}
    for τ = τ̄ - 1 ; τ ≥ 0 ; τ = τ - 1 do
        Process P̃^s_{k-τ+1} with (4) and (6)
        Calculate L̃_{k-τ} with (7)
        Calculate K̃^s_{k-τ} with (8) (if τ ≥ τ_s)
    end for
    Calculate Ñ^s_{τ̄-1:τ_s|k} from L̃_t, K̃^s_t with (17)
    Calculate ỹ^s_k with (16) and transmit it to the fusion center
    Predict P^s_{k-τ̄}, x̂^s_{k-τ̄} with (3), (4), and (13)
end for
```

time step and the number of measurements stored is $\overline{\tau}-\tau_s$ each. Indeed, independently from the computations, only the vector $\tilde{\underline{y}}_k^s$ with $\dim\{\tilde{\underline{y}}_k^s\} = \dim\{\underline{\mathbf{x}}\}$ from (16) must be transmitted to the fusion center.

To the authors' knowledge, the only technique available in literature to reconstruct the centralized KF result for the considered system in the presence of packet delays is to utilize a naïve measurement transmission technique and to employ a KF at the fusion node. The proposed algorithm is superior to this approach in the following scenarios:

- The computational power of the fusion center does not scale with the number of sensors in the network. For the proposed technique, the only operation necessary at the fusion node is to sum up the received vectors. As this is an aggregation operation, it can be done in-network when the communication structure is modeled as a tree.
- The measurement space is larger than the state space. As only the vector $\tilde{\underline{y}}_k^s$ (dimension equals that of the state) is transmitted to the fusion center, the data amount that must be transmitted is reduced at the costs of pre-processing at the sensors.
- An estimate at the fusion center is not needed in every sensing cycle. As all past measurements of a sensor are comprised in $\tilde{\underline{y}}_k^s$, it is not necessary to transmit measured values separately, but instead establish a communication rarely, e.g., each 100th time step only, and communicate all information in one vector then.

As measurements from several time steps are bundled in one vector, the last-mentioned aspect is, in particular, relevant for energy constrained sensor networks that suffer from expensive communication and for sensor networks with a high measurement sampling rate.

For the considered system with global model knowledge and deterministic communication, the proposed algorithm is optimal. However, it has been argued that the assumptions about packet losses and global model knowledge are limiting and quite often not realistic [21].

Distributed KF	→	Hypothesizing KF
↓		↓
Distributed KF with time delays	→	Hypothesizing KF with time delays

Table II. Classification of proposed algorithms. Arrows indicate generalizations.

IV. HKF for Time-delayed Information

The objective pursued in this section is to relax the assumptions from Sec. III while maintaining a high precision of the estimate at the fusion center. For this purpose, the HKF [18] is generalized. Instead of relying on guaranteed packet arrival patterns, the main idea of the HKF is to substitute the measurement capability $\left(\mathbf{P}^z_{t|k}\right)^{-1}$ (9) and provide techniques to compensate for the bias of the estimate that is induced by a difference between substitution and underlying quantity. In the following, we expand the HKF to the considered scenario with packet delays. We obtain a simple relation between the DKF, the HKF, and the proposed techniques that is depicted in Tab. II.

The main difference between the HKF versus the DKF is that the gains (7) and (8) are not obtained with the true measurement capabilities $(\mathbf{P}^z_{t|k})^{-1}$ that depend on the measurement models of the sensors and the packet delays and losses in the sensor network, but based on hypotheses $(\mathbf{C}^z_{k-t})^{-1} \approx (\mathbf{P}^z_{t|k})^{-1}$ that aim at approximating the underlying true model as best as possible. Apart from replacing $(\mathbf{P}^z_{t|k})^{-1}$ by $(\mathbf{C}^z_{k-t})^{-1}$, the processing of estimates and covariances is the same as in Sec. III.

However, the substitution of the true measurement capability has mainly two consequences. First, the covariance $\mathbf{P}_k$ does not represent the true MSE matrix of the estimate, but is an auxiliary variable to obtain optimized filter gains. Second, the fused estimate (18) is biased when the hypotheses do not match the true measurement capabilities. For this reason, debiasing matrices $\mathbf{\Delta}_k^s$ are maintained in addition to local estimates that allow the reconstruction of unbiased estimates according to

$$\mathrm{E}\{\left(\mathbf{\Delta}_k^s\right)^{-1} \tilde{\underline{y}}_k^s\} = \mathrm{E}\{\underline{\mathbf{x}}_k\} \ . \tag{19}$$

In [18], the processing of multiplicative debiasing matrices that satisfy (19) has been derived for initialization, prediction, filtering, and fusion. In the following, we adapt these methods to networks with a delayed communication.

Assume that the processing of Sec. III is applied with hypotheses $\left(\mathbf{C}^z_\tau\right)^{-1}$, $\tau \in \{0, \dots, \overline{\tau}\}$, as substitutions for the measurement capabilities. Then, in order to guarantee property (19), the inverse of the debiasing matrix must reverse the initialization operation. The initialization from (12) is for example reversed with

$$\mathbf{\Delta}_0^s = \frac{1}{|\mathcal{S}|}\mathbf{I} \ . \tag{20}$$

Property (19) specifies prediction and filtering operations as well. Taking into account that $\tilde{\underline{y}}_k^s$ is obtained with gains $\mathbf{L}_t^s$ and $\mathbf{K}_t^s$ from (7) and (8) (with $(\mathbf{C}^z_{k-t})^{-1}$ instead of $(\mathbf{P}^z_{t|k})^{-1}$), we apply the derivation of the debiasing matrix from [18] and obtain

$$\mathbf{\Delta}_{t+1}^s = \mathbf{A}\mathbf{\Delta}_t^s(\mathbf{A})^{-1} \text{ and} \tag{21}$$

$$\mathbf{\Delta}_{t|t}^s = \mathbf{L}_t^s\mathbf{\Delta}_t^s + \mathbf{K}_t^s\mathbf{H}_t^s \ . \tag{22}$$

As the gains for $t > k - \overline{\tau}$ change with temporary covariance matrices (c.f. Sec. III), the debiasing matrices consist of recursively obtained and of augmented parts as well. We obtain the temporary debiasing matrix for $\underline{\tilde{y}}_k^s$ in matrix-vector representation as

$$\mathbf{\Delta}_k^s = \mathbf{N}_{k-\overline{\tau}|k}^s \mathbf{\Delta}_{k-\overline{\tau}}^s (\mathbf{A})^{-\overline{\tau}} + \sum_{t=k-\overline{\tau}+1}^{k-\tau_s} \tilde{\mathbf{N}}_{t|k}^s (\mathbf{A})^{-(k-t)} \ . \tag{23}$$

A summary of the debiasing processing is separately given in Alg. 2. Indeed, the processing of debiasing matrices is best integrated in the processing cycle of Alg. 1.

Algorithm 2 Processing of Debias Matrix

Initialize $\mathbf{\Delta}_0^s$ with (20)
for $k = 0$; ... ; $k = k + 1$ **do**
  Filter $\mathbf{\Delta}_{k-\overline{\tau}}^s$ with (22)
  Calculate $\mathbf{\Delta}_k^s$ with (23) and transmit it to fusion center
  Predict $\mathbf{\Delta}_{k-\overline{\tau}}^s$ with (21)
end for

In the fusion center, the sensor specific vectors $\underline{\tilde{y}}_k^s$ are combined. Note that due to packet losses it is possible that only a subset of sensors $\mathcal{S}' \subseteq \mathcal{S}$ has provided local quantities.

Yet, the representation of the debiasing matrix in the information space allows the fusion

$$\underline{\tilde{y}}_k^{\mathcal{S}'} = \sum_{s \in \mathcal{S}'} \underline{\tilde{y}}_k^s \text{ and} \tag{24}$$

$$\mathbf{\Delta}_t^{\mathcal{S}'} = \sum_{s \in \mathcal{S}'} \mathbf{\Delta}_t^s \ , \tag{25}$$

with $\underline{\hat{x}}_k = (\mathbf{\Delta}_k^{\mathcal{S}'})^{-1} \underline{\tilde{y}}_k^{\mathcal{S}'}$. It directly follows from

$$\begin{aligned}(\mathbf{\Delta}_k^{\mathcal{S}'})^{-1} \sum_{s \in \mathcal{S}'} \underline{\tilde{y}}_k^s - \underline{\mathbf{x}}_k &= (\mathbf{\Delta}_k^{\mathcal{S}'})^{-1} \Big(\sum_{s \in \mathcal{S}'} \underline{\tilde{y}}_k^s - \sum_{s \in \mathcal{S}'} \mathbf{\Delta}_k^s \underline{\mathbf{x}}_k \Big) \\ &= (\mathbf{\Delta}_k^{\mathcal{S}'})^{-1} \sum_{s \in \mathcal{S}'} \mathbf{\Delta}_k^s \Big((\mathbf{\Delta}_k^s)^{-1} \underline{\tilde{y}}_k^s - \underline{\mathbf{x}}_k \Big) \ , \end{aligned} \tag{26}$$

that the difference between fused estimate (24) and the underlying true state $\underline{\mathbf{x}}_k$ diminishes in expectation when (19) holds for $s \in \mathcal{S}'$. Therefore, we have $\mathrm{E}\{\underline{\hat{x}}_k\} = \mathrm{E}\{\underline{\mathbf{x}}_k\}$.

Principally, there exist two approaches for handling missing packets in the considered framework. The naïve technique is to ignore past transmissions and to derive an estimate according to (24) and (25). Indeed, in this case, measurement information from sensors $s \in \mathcal{S} \backslash \mathcal{S}'$ is not included in the fused estimate although previous transmissions might have been successful.

However, as shown in (26), the only requirement to obtain an unbiased estimate at the fusion center is to guarantee local unbiasedness (19). Therefore, outdated estimates $\underline{\tilde{y}}_t^s$ can be projected forward at the fusion center by means of the simple formulas

$$\underline{\tilde{y}}_{t+1}^s = \mathbf{A}\underline{\tilde{y}}_t^s \text{ and } \mathbf{\Delta}_{t+1}^s = \mathbf{A}\mathbf{\Delta}_t^s(\mathbf{A})^{-1} \ ,$$

and can then be included in the fusion process (24) and (25).

As discussed in Sec. III, for time-invariant models, the covariances and gains converge to steady-state values. Obviously, the same holds for $\mathbf{\Delta}_k^s$ as it is a function of covariances and gains so that the computational effort is significantly reduced as well.

For the proposed extension of the HKF, we can go even one step further. In fact, the proposed framework provides flexible techniques to optimize filter gains at the local sensors based on expected sensor network capabilities in communication environments that are subject to packet delays and losses. As long as the local debiasing matrix satisfy (19), unbiased estimates that are optimized according to centralized KF equations are obtained with help of the fusion algorithm (24) and (25).

In particular, it is feasible to derive constant gains prior to application and to use them from the beginning, which limits the necessary computations in each time step to the matrix operations in (13) and (14). Although such processing is suboptimal during the initial period of an estimation process (with time-varying covariances), long-term applications that tend to settle to a steady-state are well covered.

Moreover, the results of [18], [20] transfer to the generalization derived in this paper. Among other things, the derivation of a MSE matrix bound [18] and the initialization from measurements [20] can be applied directly. In particular, it is worth mentioning that the estimate at the fusion center corresponds to the optimal estimate $\mathrm{E}\{\underline{\mathbf{x}}_k \mid \mathcal{I}^k\}$ if and only if the estimated measurement capabilities $(\mathbf{C}_\tau^z)^{-1}$ have met the underlying true values $(\mathbf{P}_{t|k}^y)^{-1}$, $\tau = k - t$. Otherwise, the precision depends nonlinearly on the difference between assumption and underlying truth of the measurement capabilities.

The application of the extension of the HKF is advantageous in the same scenarios that have been discussed in Sec. III. Additionally, estimators with recursive processing at the sensors comprise data from past transmissions in subsequent packets and thus information cannot be lost but only arrive delayed at the fusion center. As the algorithms from Sec. III and Sec. IV employ this local processing and the proposed extension of the HKF is capable of handling packet losses, it has a conceptual benefit versus measurement processing techniques.

V. Evaluation

As mentioned above, the extension of the DKF provides optimal results and thus, does not have to be evaluated. However, the algorithm from Sec. IV can be applied in a variety of scenarios and provides suboptimal results when the assumptions about the measurement capability do not meet the underlying true models. An examination of the performance of the HKF that is not subject to packet delays is given in [18], [21].

We focus on the impact of packet delays on the estimation performance in the following. Consider a white-noise jerk

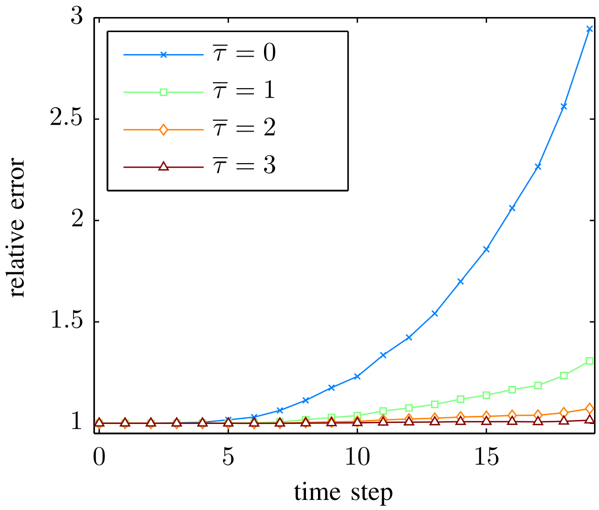

Figure 2. The relative MSE of estimators with time horizons $\overline{\tau} = 0$ to $\overline{\tau} = 3$ relative to the centralized Kalman filter results in 500 Monte Carlo runs.

model where the acceleration is given by a continuous-time white noise term [22] with time-discrete equivalent

$$\underline{\mathbf{x}}_{k+1} = \begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \underline{\mathbf{x}}_k + \underline{\mathbf{w}}_t \ ,$$

where

$$\underline{\mathbf{w}}_t \sim \mathcal{N}(\underline{0}, p \cdot \mathbf{Q}_t) \text{ with } \mathbf{Q}_t = \begin{pmatrix} t^5/20 & t^4/8 & t^3/6 \\ t^4/8 & t^3/3 & t^2/2 \\ t^3/6 & t^2/2 & t \end{pmatrix} \ .$$

We set the time interval to $t = 0.1$ and the power spectral density to $p = 10$. Let the network consist of 10 identical sensors that observe the acceleration according to

$$\underline{z}_k^s = (0 \quad 0 \quad 1)\, \underline{\mathbf{x}}_k + \underline{\mathbf{v}}_k^s \text{ with } \underline{\mathbf{v}}_k^s \sim \mathcal{N}(0, 10) \ .$$

The communication delays of the sensors vary between 0 and 4 time steps with two sensors each belonging to one delay cluster. Therefore, the measurement capabilities are given by $\left(\mathbf{P}^z_{k-\tau|k}\right)^{-1} = 2 \cdot (\tau + 1)(0\ 0\ 1)^\top \frac{1}{10}(0\ 0\ 1)$, $\tau \in \{0, \ldots, 4\}$, $k - \tau \geq 0$.

In order to determine the impact of the proposed packet delay processing on the estimation performance, we consider different delay horizons in the optimization of the algorithm from Sec. IV. More precisely, we optimize the local estimates with the correct measurement capabilities $\left(\mathbf{C}^z_\tau\right)^{-1} = \left(\mathbf{P}^z_{k-\tau|k}\right)^{-1}$, $\tau \in \{0, \ldots, \overline{\tau}\}$, but vary the considered time horizon from $\overline{\tau} = 0$ (standard HKF without delay processing) to $\overline{\tau} = 4$ (exact) so that time horizons $\overline{\tau} \in \{0, \ldots, 3\}$ yield suboptimal results.

The MSEs relative to the MSE of the optimal centralized KF of 500 Monte Carlo runs are given in Fig. 2. For $\overline{\tau} = 4$, the algorithm from Sec. IV equals the centralized KF and thus, provides optimal results. As depicted in Fig. 2, the performance of the other estimators is suboptimal and deteriorates with the considered optimization horizon. For $\overline{\tau} = 0$, i.e., the regular HKF [18], the MSE at time step 19 is almost three times as big as for the optimal approach. The more time steps are considered in the optimization, the smaller is the deterioration. For example, for $\overline{\tau} = 3$, the MSE at time step 19 is only 1.5% higher than for $\overline{\tau} = 4$.

Note that independently of the considered time horizon, the same information is processed and communicated to the fusion center. The only difference lies in the optimization of the filter gains that is done according to different measurement capabilities. Obviously, even for this simple scenario, the optimization has a distinct impact on the estimation performance. In particular, the application of local KFs with a subsequent fusion resembles the algorithm without packet delay processing optimized according to local measurement models $(\mathbf{H}^s_t)^\top (\mathbf{R}^s_t)^{-1} \mathbf{H}^s_t$, $s \in \mathcal{S}$. As local measurement models are poor approximations of the sensor network capability (9), the local KF processing yields even worse results than $\overline{\tau} = 0$.

While it remains to evaluate the approaches extensively, we demonstrated by means of a simple scenario that the extension for delayed packet processing has a significant impact on the estimation performance.

VI. Conclusion

We proposed extensions to distributed estimation algorithms for the optimal processing of information in sensor networks in the presence of packet delays. In the first instance, we generalized the distributed Kalman filter and showed that the centralized Kalman filter estimate is obtained at the fusion center based on recursively obtained quantities from the sensors. As the distributed Kalman filter relies on global model knowledge and deterministic communication, we also derived an extension of the hypothesizing Kalman filter that is capable of providing unbiased estimates with up to optimal precision for arbitrary sensor networks.

Hence, in the most general form, the proposed algorithm is able to cope with packet delays and losses without knowledge about models of remote sensors. In particular, in scenarios with a reliable description of the (stochastic) communication attributes, the application of the proposed schemes significantly improves the estimation performance compared to local Kalman filters or standard track-to-track fusion algorithms.

Acknowledgment

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

References

[1] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, and P.-J. Nordlund, "Particle Filters for Positioning, Navigation, and Tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 425–437, 2002.

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

[3] A. Bry, A. Bachrach, and N. Roy, "State Estimation for Aggressive Flight in GPS-denied Environments using onboard Sensing," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 2012*. IEEE, 2012, pp. 1–8.

[4] G. Evensen, "The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation," *Ocean dynamics*, vol. 53, no. 4, pp. 343–367, 2003.

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

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

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

[8] 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.

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

[10] 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.

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

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

[13] 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.

[14] H. R. Hashemipour, S. Roy, and A. J. Laub, "Decentralized structures for parallel Kalman filtering," *IEEE Transactions on Automatic Control*, vol. 33, no. I, pp. 88–94, 1988.

[15] 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.

[16] 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.

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

[18] M. Reinhardt, B. Noack, and U. D. Hanebeck, "The Hypothesizing Distributed Kalman Filter," in *Proceedings of the 2012 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2012)*, Hamburg, Germany, Sep. 2012.

[19] A. Mutambara, *Decentralized Estimation and Control for Multisensor Systems*. CRC press, 1998.

[20] M. Reinhardt, B. Noack, and U. D. Hanebeck, "Advances in Hypothesizing Distributed Kalman Filtering," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, Istanbul, Turkey, Jul. 2013.

[21] ——, "On Optimal Distributed Kalman Filtering in Non-ideal Situations," in *Proceedings of the 15th International Conference on Information Fusion (Fusion 2012)*, Singapore, Jul. 2012.

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