2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

A Distributed Kalman Filter with Global Covariance

J. Sijs, *Student Member, IEEE,* M. Lazar, *Member, IEEE*

***Abstract*— Most distributed Kalman filtering (DKF) algorithms for sensor networks calculate a local estimate of the global state-vector in each node. An important challenge within distributed estimation is that all sensors in the network contribute to the local estimate in each node. In this paper, a novel DKF algorithm is proposed with the goal of attaining the above property, which is denoted as global covariance. In the considered DKF set-up each node performs two steps iteratively, i.e., it runs a standard Kalman filter using local measurements and then fuses the resulting estimates with the ones received from its neighboring nodes. The distinguishing aspect of this set-up is a novel state-fusion method, i.e., *ellipsoidal intersection (EI)*. The main contribution consists of a proof that the proposed DKF algorithm, in combination with EI for state-fusion, enjoys the desired property under similar conditions that should hold for observability of standard Kalman filters. The advantages of developed DKF with respect to alternative DKF algorithms are illustrated for a benchmark example of cooperative adaptive cruise control.**

***Index Terms*— Distributed estimation; Kalman filter; Fusion; Asymptotic analysis.**

I. Introduction

The standard state-estimator for a linear process with Gaussian noise distributions is the Kalman filter, which was formally presented in [1]. The estimator calculates the global state-vector based on all measurements of the process. Nowadays, measurements are often acquired by means of a sensor network, especially in large-area processes, e.g., [2]. Employing the centralized Kalman filter in a sensor network requires global communication and central data-processing. Since this is known to be infeasible for large-scale sensor networks, a continuously increasing interest has been shown in distributed Kalman filtering (DKF) set-ups, e.g., [3]–[5]. In a typical DKF each node performs a standard Kalman filter on the local measurement and exchanges data only with neighboring nodes. Moreover, due to the unpredictable system changes within deployed large-scale sensor networks, keeping track of the shared data between different nodes is intractable. Hence, none of the nodes can afford direct knowledge of global information and thus the performance of DKFs should be assessed differently compared to centralized solutions. A first performance objective of DKF is that each individual measurement in the network contributes to any local estimate, even though direct communication is limited to neighboring nodes. As such, the objective of a DKF is a *global covariance*, i.e., each measurement that is available in the sensor network sets asymptotic bounds on some parts of the local state-covariance in any other node. A DKF with this property will be referred to as *global covariance DKF*.

In what follows, it is shown that an approach to attain this property is employing a state-fusion method in each node on the resulting estimates of local Kalman filters. This fusion method merges the local estimate of a node with the estimates received from neighboring nodes. Notice that in contrast to this approach, it can be proven that the *global covariance* property is not attainable if measurements are exchanged instead of estimates, as proposed in [3], [5], [6] (unless each node communicates with all the other nodes in the network). However, simply performing an arbitrary state-fusion method, such as the one in, e.g., [7], [8], does not yield *global covariance* either, since certain required properties of the state-fusion method are not satisfied.

As such, the main contribution of this paper is a proof that a DKF where each node performs a local Kalman filter followed by the state-fusion method ellipsoidal intersection (EI) of [9], is a *global covariance DKF*. This proof consists of deriving asymptotic bounds of the full state-covariance for each node and showing that these bounds are a function of the global sensor data available within the sensor network. Moreover, it is proven that to attain asymptotic bounds of the full state-covariance in *all* nodes, a sufficient condition is that there exists at least one node in the network for which the state is locally observable. A benchmark "platoon of vehicles" example is employed to illustrate the benefits of the developed global covariance DKF in comparison with other existing DKF algorithms.

II. Preliminaries

$\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{Z}$ and $\mathbb{Z}_+$ define the set of real numbers, non-negative real numbers, integer numbers and non-negative integer numbers, respectively. For any $\mathcal{C} \subset \mathbb{R}$, let $\mathbb{Z}_{\mathcal{C}} := \mathbb{Z} \cap \mathcal{C}$. The notation 0 is used to denote either zero, the null-vector or the null-matrix of appropriate dimensions. The transpose, inverse and determinant of a matrix $A \in \mathbb{R}^{n\times n}$ are denoted as $A^\top$, A^{-1} and $|A|$, respectively. Further, $[A]_{qr} \in \mathbb{R}$ denotes the element in the q-th row and r-th column of A. Given that $A, B \in \mathbb{R}^{n\times n}$ are positive definite, denoted with $A \succ 0$ and $B \succ 0$, then $A \succ B$ denotes $A - B \succ 0$. $A \succeq 0$ denotes that A is positive semi-definite. For any $A \succ 0$, $A^{\frac{1}{2}}$ denotes its Cholesky decomposition and $A^{-\frac{1}{2}}$ denotes $(A^{\frac{1}{2}})^{-1}$. Furthermore, the following matrix properties for any $A, B \succ 0$ will be frequently used:

- If $A \succeq B$, then $CAC^\top \succeq CBC^\top$ for any C of appropriate dimensions (Proposition 8.1.2 [10]);
- If $A \succeq B$, then $A^{-1} \preceq B^{-1}$ (Proposition 8.5.5 [10]).

J. Sijs is with TNO Science and Industry, P.O. Box 155, 2600 AD Delft, The Netherlands, E-mail: joris.sijs@tno.nl.

M. Lazar is with Eindhoven University of Technology, Eindhoven, The Netherlands, E-mail: m.lazar@tue.nl.

Given $A \in \mathbb{R}^{n\times n}$, let $\nu_q(A) \in \mathbb{R}^n$ and $\lambda_q(A) \in \mathbb{R}$ denote its q-th eigenvector and eigenvalue, respectively. If $\nu_q(A)$ and $\lambda_q(A)$ contain only real values, for all $q \in \mathbb{Z}_{[1,n]}$, then the Jordan decomposition of A yields: $A = SDS^{-1}$. Hence, $S := (\nu_1(A) \ \dots \ \nu_n(A))$, i.e., $[S]_{qr} = [\nu_r(A)]_q$ for all $q,r \in \mathbb{Z}_{[1,n]}$, and $D := \text{diag}\,(\lambda_1(A),\dots,\lambda_n(A))$, i.e., $[D]_{qr} = \lambda_q(A)$ if $q = r$ and $[D]_{qr} = 0$ otherwise, for all $q,r \in \mathbb{Z}_{[1,n]}$. The Gaussian function (Gaussian in short) of vectors $x,\mu \in \mathbb{R}^n$ and matrix $\Sigma \in \mathbb{R}^{n\times n}$ is denoted as $G(x,\mu,\Sigma)$.

To model the sensor network, consider an undirected graph $\mathcal{G} = (V,E)$, where $V = \{v_1,\cdots,v_N\}$ is the set of nodes (vertices), for some $N \in \mathbb{Z}_+$, $E \subseteq (V \times V)$ is the set of edges and (v_i,v_j) is the edge from v_i to v_j. If $(v_i,v_j) \in E$, then $(v_j,v_i) \in E$. Also, $\mathcal{N} := \mathbb{Z}_{[1,N]}$ denotes the set of node-indices.

Definition II.1 Let $r,s \in \mathcal{N}$ and a finite, undirected graph $\mathcal{G} = (V,E)$ be given. Then a graph path that starts at $v_r \in V$ and ends at $v_s \in V$ is a sequence of vertices $\tau_{r,s} = \{\upsilon^{(1)},\cdots,\upsilon^{(l)}\} \subseteq V$, where $(\upsilon^{(j)},\upsilon^{(j+1)}) \in E$, for all $j \in \mathbb{Z}_{[1,l-1]}$, and $\upsilon^{(1)} = v_r$, $\upsilon^{(l)} = v_s$. Furthermore, the length of the path is $L(\tau_{r,s}) := l$ and $L(\tau_{r,r}) := 0$.

Definition II.2 Let $r,s \in \mathcal{N}$ and a finite, undirected graph $\mathcal{G} = (V,E)$ be given. Then the graph distance between $v_r, v_s \in V$, denoted with $d(v_r,v_s)$, is the length of the shortest path between them, i.e., $d(v_r,v_s) := \min_{\tau_{r,s}\in\mathcal{T}_{r,s}} L(\tau_{r,s})$, where $\mathcal{T}_{r,s}$ is the set of all graph paths from v_r to v_s. In case $v_r = v_s$ we define $d(v_r,v_r) := 0$.

For any $i \in \mathcal{N}$, let $\mathcal{N}_{i(q)} := \{j \in \mathcal{N} | d(v_i,v_j) = q\}$ denote the set of nodes corresponding to the q-th order neighbors of node v_i. Further, let $\mathcal{N}_{i(0,1)} := \mathcal{N}_{i(1)} \cup \{i\}$ and let $\sharp\mathcal{N}_{i(q)}$ denote the cardinality of $\mathcal{N}_{i(q)}$. Notice that $\cup_{q\in\mathbb{Z}_+}\mathcal{N}_{i(q)} = \mathcal{N}$.

III. Problem formulation

Let an autonomous process and a sensor network of N nodes be given. The state-vector of the process, denoted as $x \in \mathbb{R}^n$, is affected by process-noise, denoted as $w \in \mathbb{R}^n$. The measurements that are taken at a node i, denoted as $y_i \in \mathbb{R}^{m_i}$, are affected by measurement-noise, denoted as $d_i \in \mathbb{R}^{m_i}$, where $m_i \in \mathbb{Z}_{\geq 1}$ is the output dimension in each node i. In case $k \in \mathbb{Z}_+$ denote the sampling-instants of the system, then the discrete-time process-model at a node i, yields

$$\begin{aligned} x[k+1] &= Ax[k] + w[k], \\ y_i[k] &= C_i x[k] + d_i[k]. \end{aligned} \tag{1}$$

Both the process-noise and the measurement-noise are assumed to have a zero-mean Gaussian PDF for all k, i.e.,

$$p(w[k]) := G(w[k],0,Q) \text{ and } p(d_i[k]) := G(d_i[k],0,R_i).$$

The goal of the sensor network is to estimate x in each node by means of a DKF. In line with the current literature, the estimation algorithm at each node i performs a "local Kalman filter" (LKF) given y_i. The resulting PDF of the LKF at node i at sample instant k is denoted as $p_i(x[k]) := G(x[k],\hat{x}_i[k],P_i[k])$, for some $\hat{x}_i[k] \in \mathbb{R}^n$ and $P_i[k] \in \mathbb{R}^{n\times n}$. To make y_i improve the accuracy of $p_j(x)$, for all $j \in \mathcal{N}$, i.e., to enable *global covariance*, nodes exchange the result of their LKF with each other. As such, node i receives $p_j(x[k])$ for all $j \in \mathcal{N}_{i(1)}$. The received PDFs are then merged with $p_i(x[k])$ in a "local state-fusion algorithm" (LSF). The resulting PDF of this fusion-step is denoted with $p_{i_f}(x[k]) := G(x[k],\hat{x}_{i_f}[k],P_{i_f}[k])$, for some $\hat{x}_{i_f}[k] \in \mathbb{R}^n$ and $P_{i_f}[k] \in \mathbb{R}^{n\times n}$.

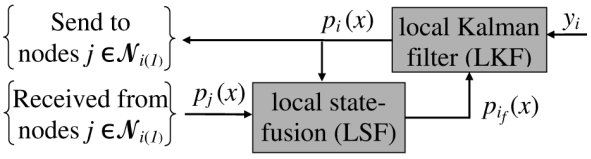

Fig. 1. Schematic set-up of the estimation algorithm of node i.

To complete the estimation algorithm of a node, two aspects are still to be addressed. One is the state-fusion method, which for clarity is simply regarded as the problem of fusing two arbitrary estimates of the same state-vector x, i.e., $p_i(x)$ and $p_j(x)$. Since keeping track of shared estimates between different nodes is intractable, the employed fusion method cannot require any knowledge on correlations, or mutual information, of $p_i(x)$ and $p_j(x)$. Only then the LSF of each node i is able to deal with correlations of the local estimate $p_i(x[k])$ and any other estimate $p_j(x[k])$ in the network, i.e., also for any node $j \notin \mathcal{N}_{i(1)}$. The second aspect is a complete description of the estimation algorithm, which is schematically depicted in Figure 1, and its properties. The properties of interest for any DKF algorithm are, firstly, *global covariance*, as explained in the Introduction and, secondly, an asymptotic bound guarantee for each state-covariance matrix P_{i_f}. The next section motivates our choice of the state-fusion method, after which the properties of the proposed DKF algorithm are presented.

IV. State fusion: Ellipsoidal intersection

This section recalls the recently developed state-fusion method EI, as presented in [9]. The method fuses $p_i(x) := G(x,\hat{x}_i,P_i)$ with $p_j(x) := G(x,\hat{x}_j,P_j)$ into the new estimate $p_{i_f}(x) := G(x,\hat{x}_{i_f},P_{i_f})$, for some $\hat{x}_i,\hat{x}_j,\hat{x}_{i_f} \in \mathbb{R}^n$ and $P_i,P_j,P_{i_f} \in \mathbb{R}^{n\times n}$. The distinguishing feature of this method, compared to alternative fusion methods, e.g., [7], [8], [11], is that mutual information is parameterized a priori to deriving a fusion formula via estimation theory. As such, it was proven in [9] that EI satisfies the following criterion, which is necessary to attain a *global covariance DKF*.

Criterion IV.1 Let $p_i(x)$ and $p_j(x)$ be given. Then independent on correlations it should hold that $P_{i_f} \preceq P_i$ and $P_{i_f} \preceq P_j$.

Criterion IV.1 is a formal characterization of the fact that $p_{i_f}(x)$ is a "more accurate" estimate of x, compared to $p_i(x)$ and $p_j(x)$, independent of mutual information. Mutual information refers to, for example, measurements and process-model parameters that were used by both $p_i(x)$ and $p_j(x)$. Hence, in the first step of EI a new estimate is introduced that is only based on the mutual information of $p_i(x)$ and $p_j(x)$.

This estimate is denoted with $p_\gamma(x) := G(x,\gamma,\Gamma)$, for some "mutual mean" $\gamma \in \mathbb{R}^n$ and "mutual covariance" $\Gamma \in \mathbb{R}^{n\times n}$. In case γ and Γ are known, then the expressions of the fused mean $\hat{x}_{i_f}$ and fused covariance P_{i_f} follow from estimation theory, i.e.,

$$\begin{aligned} P_{i_f} &= \left(P_i^{-1} + P_j^{-1} - \Gamma^{-1}\right)^{-1}, \\ \hat{x}_{i_f} &= P_{i_f}\left(P_i^{-1}\hat{x}_i + P_j^{-1}\hat{x}_j - \Gamma^{-1}\gamma\right). \end{aligned} \tag{2}$$

The values for γ and Γ are determined in the second step of EI. To ensure that the unknown correlation is treated correctly, γ and Γ are derived by assuming a maximum effect of the mutual information on $p_i(x)$ and $p_j(x)$. To that extent, the matrices S_i, D_i, S_j and D_j are introduced, such that the following Jordan decompositions hold, i.e.,

$$P_i = S_i D_i S_i^{-1}, \quad D_i^{-\frac{1}{2}} S_i^{-1} P_j S_i D_i^{-\frac{1}{2}} = S_j D_j S_j^{-1}.$$

Also, let $H := P_i^{-1} + P_j^{-1} - 2\Gamma$ and let $\lambda_{0_+}(H) \in \mathbb{R}_+$ denote the smallest, non-zero eigenvalue of H. Then the mutual covariance and mutual mean, for some $\varepsilon \in \mathbb{R}_+$, yields

$$\Gamma = S_i D_i^{\frac{1}{2}} S_j D_\Gamma S_j^{-1} D_i^{\frac{1}{2}} S_i^{-1}, \tag{3}$$

$$\begin{aligned} \gamma = &\left(P_i^{-1} + P_j^{-1} - 2\Gamma^{-1} + 2\varepsilon I\right)^{-1} \times \\ &\left(\left(P_j^{-1} - \Gamma^{-1} + \varepsilon I\right)\hat{x}_i + \left(P_i^{-1} - \Gamma^{-1} + \varepsilon I\right)\hat{x}_j\right), \end{aligned} \tag{4}$$

$$\text{where } [D_\Gamma]_{qr} = \begin{cases} \max([D_j]_{qr}, 1) & \text{if} \quad q = r, \\ 0 & \text{if} \quad q \neq r, \end{cases} \tag{5}$$

$$\text{and} \quad \varepsilon = \begin{cases} 0 & \text{if} \quad |H| \neq 0, \\ \varepsilon \ll \lambda_{0_+}(H) & \text{if} \quad |H| = 0. \end{cases} \tag{6}$$

The interested reader is referred to [9] for a more detailed description of EI. Next, the algorithm of the proposed DKF, as it is schematically depicted in Figure 1, is presented.

V. A distributed Kalman filter

The algorithm that is performed by each node i consists of a LKF followed by a LSF. The part that corresponds to the LKF is derived from a Kalman filter in the information form, see, e.g., [3]. State-fusion of one estimate with multiple other estimates is commonly conducted recursively. This means that the LSF algorithm fuses $p_i(x[k])$ with the first received $p_j(x[k])$, after which their resulting fused estimate is further merged with the PDF that is received next, and so on. Let the initial local estimate at sample-instant k be defined as $p_{i(0)}(x) := p_i(x[k])$. Then this recursive behavior implies that $p_{i(l)}(x)$, for all $l \in \mathbb{Z}_{[1,L]}$ and $L := \sharp\mathcal{N}_{i(1)}$, is defined as the fused estimate of $p_{i(l-1)}(x)$ and the l-th received estimate $p_j(x[k])$, which will be denoted as $p_{j(l)}(x)$. The final estimate after fusing $p_i(x[k])$ with all received PDFs is therefore $p_{i_f}(x[k]) := p_{i(L)}(x)$. Hence, the following algorithm is performed by a node i at each sample-instant k:

Algorithm V.1 Distributed ellipsoidal intersection (DEI)

Step 1: local Kalman filter (LKF)

$M_i[k] = AP_{i_f}[k-1]A^\top + Q;$

$P_i[k] = \left(M_i^{-1}[k] + C_i^\top R_i^{-1} C_i\right)^{-1};$

$\hat{x}_i[k] = P_i[k]\left(M_i^{-1}[k]A\hat{x}_{i_f}[k-1] + C_i^\top R_i^{-1} y_i[k]\right);$

Step 2: local state fusion (LSF)

$\hat{x}_{i(0)} = \hat{x}_i[k], \quad P_{i(0)} = P_i[k];$

for $l = 1, \ldots, L$, do:

$\hat{x}_{j(l)} = \hat{x}_j[k], \quad P_{j(l)} = P_j[k], \quad j \in \mathcal{N}_{i(1)};$

$\Gamma_{(l)} = \text{MutualCovariance}(P_{i(l-1)}, P_{j(l)})$, i.e., (3);

$\gamma_{(l)} = \text{MutualMean}(P_{i(l-1)}, P_{j(l)}, \Gamma_{(l)}, \hat{x}_{i(l-1)}, \hat{x}_{j(l)})$, i.e., (4);

$P_{i(l)} = \left(P_{i(l-1)}^{-1} + P_{j(l)}^{-1} - \Gamma_{(l)}^{-1}\right)^{-1};$

$\hat{x}_{i(l)} = P_{i(l)}\left(P_{i(l-1)}^{-1}\hat{x}_{i(l-1)} + P_{j(l)}^{-1}\hat{x}_{j(l)} - \Gamma_{(l)}^{-1}\gamma_{(l)}\right);$

end

$\hat{x}_{i_f}[k] = \hat{x}_{i(L)}, \quad P_{i_f}[k] = P_{i(L)}.$

□

The proposed DKF of Algorithm V.1 is denoted as DEI to distinguish it from other DKF algorithms. Let us proceed this analysis by proving that the DEI is a global covariance DKF and that under certain conditions P_{i_f}, for all $i \in \mathcal{N}$, is asymptotically bounded. In terms of the bounds, more precisely, it is shown that there exists a covariance-matrix $\Theta_i \in \mathbb{R}^{n\times n}$, such that $\lim_{k\to\infty} P_{i_f}[k] \preceq \Theta_i$. To that extent, assume that each node i performs a stand-alone LKF in parallel to the DEI, i.e., only step 1 of Algorithm V.1. Hence, if μ_i and Σ_i denote the state-estimates of the stand-alone LKF as performed by node i, then they are updated as follows,

$$\begin{aligned} \Lambda_i[k] &= A\Sigma_i[k-1]A^\top + Q, \\ \Sigma_i[k] &= \left(\Lambda_i^{-1}[k] + C_i^\top R_i^{-1} C_i\right)^{-1}, \\ \mu_i[k] &= \Sigma_i[k]\left(\Lambda_i^{-1}[k]A\mu_i[k-1] + C_i^\top R_i^{-1} y_i[k]\right). \end{aligned} \tag{7}$$

The asymptotic property of $\Sigma_i[k]$ was analyzed in [12]. Therein, it was established that the following result, where $\Sigma_i[\infty] := \lim_{k\to\infty}\Sigma_i[k]$ and $\Psi_i := A\Sigma_i[\infty]A^\top + Q$, holds.

Proposition V.2 *If (A, C_i) is an observable pair and $\lambda_q\left(A - A\Psi_i C_i^\top (C_i\Psi_i C_i^\top + R_i)^{-1} C_i\right) \leq 1$, for all $q \in \mathbb{Z}_{[1,n]}$, then $\Sigma_i[\infty]$ exists and satisfies $\Sigma_i^{-1}[\infty] = \Psi_i^{-1} + C_i^\top R_i^{-1} C_i$.*

Before P_{i_f} of the DEI is related to Σ_i of the stand-alone LKF, let us start by deriving certain bounds on both $P_i[k]$ and $P_{i_f}[k]$. The first set of bounds depends on state-covariances of node i and its direct neighboring nodes, i.e., all nodes within $\mathcal{N}_{i(0,1)}$.

Lemma V.3 *Let each node i perform the DEI. Then $P_{i_f}[k] \preceq P_j[k]$ and $P_i[k+1] \preceq AP_j[k]A^\top + Q$, for all $j \in \mathcal{N}_{i(0,1)}$, $k \in \mathbb{Z}_+$.*

Lemma V.3 is proven in Appendix A. Notice that this lemma also indicates that the received state-covariances of a node i are bounded by their neighboring state-covariances as well. If, for example, node $j \in \mathcal{N}_{i(2)}$ is a direct neighbor of a certain node $h \in \mathcal{N}_{i(1)}$, i.e., $j \in \mathcal{N}_{h(1)}$, then Lemma V.3 can be used to establish that $P_h[k] \preceq AP_j[k-1]A^\top + Q$. Moreover, since node h is a direct neighbor of node i, the same lemma also gives that $P_{i_f}[k] \preceq P_h[k]$. Hence, although node $j \in \mathcal{N}_{i(2)}$ is not a direct neighbor of node i, there exists a bound on $P_{i_f}[k]$ depending on $P_j[k-1]$. In the next lemma, this principle is extended to prove that $P_{i_f}[k]$ is bounded by a prediction of $P_j[k-m]$ to the k-th sample-instant for any node $j \in \mathcal{N}_{i(m+1)}$. Notice that the value of m is such that the graph-distance between nodes i and j is equal to $m+1$.

Lemma V.4 *Let each node i perform the DEI. Then $P_{i_f}[k] \preceq A^m P_j[k-m](A^m)^\top + \sum_{c=0}^{m-1} A^c Q (A^c)^\top$ holds for all $k \in \mathbb{Z}_{\geq m}$, $j \in \mathcal{N}_{i(m+1)}$ and all $m \in \mathbb{Z}_{\geq 1}$.*

Lemma V.4 is proven in Appendix B. Notice that, since P_j depends on R_j, this lemma guarantees that y_j, for all $j \in \mathcal{N}$, contribute to improving $p_i(x[k])$ at each node i. *Hence, the proposed DKF algorithm enjoys the global covariance property*. The derivation of the asymptotic bounds on P_{i_f} continues by relating P_{i_f} to Σ_i of the stand-alone LKF.

Lemma V.5 *Let each node i perform the DEI and the stand-alone LKF. Then $P_i[1] = \Sigma_i[1]$ and $P_i[k] \preceq \Sigma_i[k]$, for all $k \in \mathbb{Z}_{\geq 2}$.*

Lemma V.5 is proven in Appendix C. Now the main result on the asymptotic analysis of the DEI can be stated, for which $P_{i_f}[\infty] := \lim_{k\to\infty} P_{i_f}[k]$ and $\tilde{\Sigma}_j \in \mathbb{R}^{n\times n}$, for all $j \in \mathcal{N}$ and $m \in \mathbb{Z}_{\geq 1}$ such that $\Sigma_j[\infty]$ exists, is defined as follows

$$\tilde{\Sigma}_j := \begin{cases} \Sigma_j[\infty] & \text{if} \quad j \in \mathcal{N}_{i(0,1)}, \\ A^m \Sigma_j[\infty](A^m)^\top + \sum_{c=0}^{m-1} A^c Q (A^c)^\top & \text{if} \quad j \in \mathcal{N}_{i(m+1)}. \end{cases}$$

Theorem V.6 *Let each node i perform the DEI and the stand-alone LKF and suppose that the hypothesis of Proposition V.2 holds. Then it holds that $P_{i_f}[\infty] \preceq \tilde{\Sigma}_j$ for all $j \in \mathcal{N}$.*

Proof: Using the definition of $\tilde{\Sigma}_j$, let us first prove that $P_{i_f}[\infty] \preceq \Sigma_j[\infty]$ holds for all $j \in \mathcal{N}_{i(0,1)}$. Lemma V.3 and Lemma V.5 give that for all $j \in \mathcal{N}_{i(0,1)}$ and $k \in \mathbb{Z}_+$ it holds that $P_{i_f}[k] \preceq P_j[k]$ and $P_j[k] \preceq \Sigma_j[k]$. Hence, also $P_{i_f}[k] \preceq \Sigma_j[k]$ holds for all $j \in \mathcal{N}_{i(0,1)}$ and $k \in \mathbb{Z}_+$. This inequality can be rewritten as $P_{i_f}[k] - \Sigma_j[k] \preceq 0$, and thus,

$$P_{i_f}[k] - \Sigma_j[k] + \Sigma_j[\infty] - \Sigma_j[\infty] \preceq 0, \quad \forall j \in \mathcal{N}_{i(0,1)},\ k \in \mathbb{Z}_+. \tag{8}$$

Proposition V.2 gives that $\lim_{k\to\infty}(\Sigma_j[k] - \Sigma_j[\infty]) = 0$. Hence, if $k \to \infty$, then (8) yields $\lim_{k\to\infty}(P_{i_f}[k] - \Sigma_j[\infty]) \preceq 0$. As $\Sigma_j[\infty]$ is a constant, this inequality is equivalent to $\lim_{k\to\infty} P_{i_f}[k] - \Sigma_j[\infty] \preceq 0$ and thus,

$$P_{i_f}[\infty] \preceq \Sigma_j[\infty],\ \forall j \in \mathcal{N}_{i(0,1)}. \tag{9}$$

Next, it is shown that for all $j \in \mathcal{N}_{i(m+1)}$ and $m \in \mathbb{Z}_{\geq 1}$, $P_{i_f}[\infty] \preceq A^m \Sigma_j[\infty](A^m)^\top + \sum_{c=0}^{m-1} A^c Q (A^c)^\top$ holds. Let us start from Lemma V.4, which for all $m \in \mathbb{Z}_{\geq 1}$ and $k \in \mathbb{Z}_{\geq m}$ gives the following inequality, i.e.,

$$P_{i_f}[k] \preceq A^m P_j[k-m](A^m)^\top + \sum_{c=0}^{m-1} A^c Q (A^c)^\top, \quad \forall j \in \mathcal{N}_{i(m+1)}.$$

Using Lemma V.5, i.e., $P_j[k-m] \preceq \Sigma_j[k-m]$, in the above inequality gives that for all $m \in \mathbb{Z}_{\geq 1}$ and $k \in \mathbb{Z}_{\geq m}$ it holds that

$$P_{i_f}[k] \preceq A^m \Sigma_j[k-m](A^m)^\top + \sum_{c=0}^{m-1} A^c Q (A^c)^\top, \quad \forall j \in \mathcal{N}_{i(m+1)}.$$

As $m < \infty$, we have that $\lim_{k\to\infty} k - m = \infty$. Therefore, similarly as to (8) and (9) one can derive that

$$P_{i_f}[\infty] \preceq A^m \Sigma_j[\infty](A^m)^\top + \sum_{c=0}^{m-1} A^c Q (A^c)^\top, \quad \forall j \in \mathcal{N}_{i(m+1)},$$

holds for any $m \in \mathbb{Z}_{\geq 1}$, which completes the proof. ■

Notice that in Theorem V.6 it is assumed that the hypothesis of Proposition V.2 is satisfied for all $j \in \mathcal{N}$. However, the same theorem also proves that $P_{i_f}[\infty]$ is already a bounded matrix if there exists only one bounded matrix $\Sigma_j[\infty]$, for any $j \in \mathcal{N}$, as $m < \infty$, which is formally stated next.

Proposition V.7 *Let there exists at least one node $j \in \mathcal{N}$, for which the hypothesis of Proposition V.2 holds. Then $P_{i_f}[\infty] \prec \tilde{\Sigma}_j$ is a bounded matrix for all $i \in \mathcal{N}$.*

Remark V.8 A possible extension to Algorithm V.1 is incorporating $y_j[k]$, for all $j \in \mathcal{N}_{i(1)}$, in the LKF of a node i. This extension can be used if for none of the nodes $i \in \mathcal{N}$ the stand-alone LKF yields a bounded $\Sigma_i[\infty]$. Then one can choose a node $j \in \mathcal{N}$, for which all the nodes $h \in \mathcal{N}_{j(1)}$ sent their local measurements $y_h[k]$ on top of $p_h(x[k])$. Incorporating these measurements in the LKF of node j is equivalent to extending y_j with $y_{j(1)} := (y_{h_1}, \ldots, y_{h_L})^\top$, for all $h_l \in \mathcal{N}_{j(0,1)}$ and $L := \sharp\mathcal{N}_{j(0,1)}$. The neighboring nodes $h \in \mathcal{N}_{j(1)}$ can thus be chosen so that node j satisfies the hypothesis of Proposition V.2, due to which Theorem V.6 gives that P_{i_f}, for all $i \in \mathcal{N}$, is asymptotically bounded. □

Remark V.9 A non-observable pair (A, C_j) implies that $\Sigma_j^{-1}[\infty] \succeq 0$ rather than $\Sigma_j^{-1}[\infty] \succ 0$. Hence, the inverse of the covariance matrix $\tilde{\Sigma}_j$, i.e., $\tilde{\Sigma}_j^{-1}$, does exists for all nodes j *independent* of whether the hypothesis of Proposition V.2 holds or not. For $P_{i_f}[\infty]$ to be bounded it should satisfy $P_{i_f}^{-1}[\infty] \succ 0$. Since the statement of Theorem V.6 can be rewritten into $P_{i_f}^{-1}[\infty] \succeq \tilde{\Sigma}_j^{-1}$, for all $j \in \mathcal{N}$, $P_{i_f}[\infty]$ is bounded if $\left(\sum_{j\in\mathcal{N}} \Sigma_j^{-1}[\infty]\right) \succ 0$. Notice that is similar to the condition for stability of the centralized Kalman filter. □

VI. Illustrative case study

The estimation error of the proposed DEI algorithm is compared to alternative DKF algorithms. The benchmark application is a platoon of four vehicles, in which each vehicle has a cooperative adaptive cruise controller [13]. The main reason for choosing this application is that the controller in

each vehicle requires an estimate of the kinematic state x of the leading vehicle in the platoon, which is defined as its longitudinal position and speed. The goal is to estimate the state-vector in each vehicle, for which $x[0] = \left(\begin{smallmatrix}10\\15\end{smallmatrix}\right)$. In this example, the acceleration of the leading vehicle is set to $a(t) = 0.1 + 3\sin(0.2t)$ and its continuous-time process-model is defined as the double integrator, i.e.,

$$\dot{x}(t) := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} a(t).$$

The sampling-time is 0.1 seconds. As a result, the discrete-time process-model of (1) is defined with $A = \left(\begin{smallmatrix}1 & 0.1\\ 0 & 1\end{smallmatrix}\right)$ and $w(t) = \left(\begin{smallmatrix}0.005\\ 0.1\end{smallmatrix}\right)a(t)$, due to which $Q = 10^{-3}\left(\begin{smallmatrix}0.1 & 2.5\\ 2.5 & 50\end{smallmatrix}\right)$.

Each vehicle can only communicate with its front and rear vehicle. Hence, the graph-model $\mathcal{G}(V,E)$ of this network yields $V := \{v_1, v_2, v_3, v_4\}$ and $E := \{(v_i, v_j) | (i-j)^2 = 1\}$, for all $i, j \in \mathbb{Z}_{[1,4]}$. A vehicle measures its own position and speed and the distance to the vehicle in front. Notice that due to this dispersion of the sensors we have that the process-model of (1) for vehicles 1 and 2 is such that $C_1 = \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)$ and $C_2 = (1\ 0)$. The corresponding covariance of d_1 and d_2 is set to $R_1 = \left(\begin{smallmatrix}0.05 & 0\\ 0 & 0.05\end{smallmatrix}\right)$ and $R_2 = 0.5$, respectively. Vehicles 3 and 4 have no sensors related to vehicle 1. Therefore, the process-model of (1) for vehicles 3 and 4 uses $C_3 = C_4 = (0\ 0)$ and $R_3^{-1} = R_4^{-1} = 0$. Three different DKF algorithms are compared: (i) DEI as defined in Algorithm V.1; (ii) a DKF, denoted as DCI, which is similar to DEI only that state-fusion is done according to covariance intersection of [7] instead of EI; (iii) the DKF as presented in [3], denoted as DIF, which performs a stand-alone LKF at each vehicle/node i on all measurements $y_j \in \mathcal{N}_{i(0,1)}$. Notice that, due to the dispersion of sensors and communication, vehicles 1 and 2 in the DIF both perform a centralized Kalman filter.

Initial values are set to $\hat{x}_i[0] = \left(\begin{smallmatrix}10\\15\end{smallmatrix}\right)$ and $P_i[0] = \left(\begin{smallmatrix}10 & 0\\ 0 & 10\end{smallmatrix}\right)$ for all $i \in \mathbb{Z}_{[1,4]}$. The estimation-error of vehicles 2 and 4 are presented in Figure 2. However, the DIF of vehicle 4 does not receive any measurements of the first vehicle. Hence, x is not observable in vehicle 4 and we will omit the results of the DIF in plots (C) and (D) of Figure 2.

Figure 2 shows that the DEI clearly outperforms the DCI in both vehicles. This is mainly a result of the employed fusion method covariance intersection in the DCI, which does not take the mutual information explicitly into account as EI does. Instead, the fused estimate is parameterized as a consensus of the original estimates while minimizing P_{i_f}. As such, $p_{i_f}(x[k])$ resembles to a consensus of all the different local estimates in the network. Since vehicles 3 and 4 have no measurements depending on x, their local estimation error is high, which then merges into vehicles 1 and 2 as a result of covariance intersection. Further, notice that the DEI and the DIF have comparable estimation results in vehicle 2. This is because the DEI in vehicle 2 extracts the exclusive information of vehicle 1, i.e., y_1 and R_1. These are exactly the same variables that are exchanged in the DIF, due to which the DEI and the DIF of vehicle 2 result in almost the same estimate. However, as the DIF exchanges measurements, the state-vector is only observable in the first three vehicles of a platoon. Hence, the DEI outperforms the DIF as well.

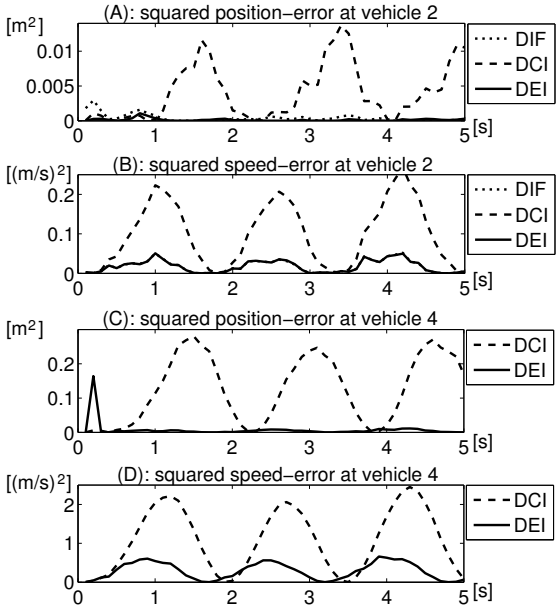

Fig. 2. The squared estimation-error at vehicles 2 and 4.

VII. Conclusions

In this paper, a novel distributed Kalman filtering algorithm was proposed to attain *global covariance*. This means that all the measurements in the sensor network improve the local estimate at any other node in the network. In the considered DKF set-up each node performs two steps iteratively, i.e., it runs a Kalman filter using local measurements and then fuses the resulting estimate with the ones received from the neighboring nodes. The distinguishing aspect of the set-up was a novel state fusion method, i.e., ellipsoidal intersection. The main contribution consisted of a proof that the proposed DKF algorithm, in combination with ellipsoidal intersection for fusion, enjoys the *global covariance* property. The advantages of developed DKF with respect to alternative DKF algorithms were illustrated for a benchmark example. Future work is concerned with establishing convergence of the state-covariances in case the state is not locally observable for all nodes.

VIII. Acknowledgements

Part of this research is supported by the Veni grant "Flexible Functions for Real-time Control", grant number 10230, awarded by STW (Dutch Technology Foundation) and NWO (The Netherlands Organisation for Scientific Research).

References

[1] R. Kalman, "A new approach to linear filtering and prediction problems," *Transaction of the ASME Journal of Basic Engineering*, vol. 82, no. D, pp. 35–42, 1960.

[2] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless Sensor Networks: a survey," *Elsevier, Computer Networks*, vol. 38, pp. 393–422, 2002.

[3] H. Durant-Whyte, B. Rao, and H. Hu, "Towards a fully decentralized architecture for multi-sensor data fusion," in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, Cincinnati, USA, 1990, pp. 1331–1336.

[4] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proc. of the 46th IEEE Conf. on Decision and Control*, New Orleans, USA, 2007.
[5] J. Sijs, M. Lazar, P. Van de Bosch, and Z. Papp, "An overview of non-centralized Kalman filters," in *Proc. of the IEEE Int. Conf. on Control Applications*, San Antonio, USA, 2008, pp. 739–744.
[6] U. Khan and J. Moura, "Distributed Kalman filters in sensor networks: Bipartite Fusion Graphs," in *IEEE 14th Workshop on Statistical Signal Processing*, Madison, USA, 2007, pp. 700–704.
[7] S. J. Julier and J. K. Uhlmann, "A Non-divergent Estimation Algorithm in the Presence of Uknown Correlations," in *Proc. of the American Control Conference*, Piscataway, USA, 1997, pp. 2369–2373.
[8] D. Franken and A. Hupper, "Improved Fast Covariance Intersection for Distributed Data Fusion," in *Proc. of the 8th Int. Conf. on Information Fusion*, Philidalphia, PA, USA, 2005.
[9] J. Sijs, M. Lazar, and P. v.d. Bosch, "State fusion with unknown correlation: Ellipsoidal intersection," in *Proc. of the American Control Conference*, Baltimore, USA, 2010, pp. 3992 – 3997.
[10] D. Bernstein, *Matrix Mathematics*. Princeton University Press, 2005.
[11] Y. Zhuo and J. Li, "Data fusion of unknown correlations using internal ellipsoidal approximations," in *Proc. of the 17th IFAC World Congress*, Seoul, Korea, 2008, pp. 2856–2860.
[12] S. W. Chan, G. C. Goodwin, and K. S. Sin, "Convergence properties of the Riccati difference equation in optimal filtering of nonstabilizable systems," *IEEE Trans. on Automatic Control*, vol. 29, pp. 110–118, 1984.
[13] B. van Arem, C. van Driel, and R. Visser, "The impact of cooperative adaptive cruise control on traffic-flow characteristics," *IEEE Trans. on Intelligent Transportation Systems*, vol. 7, pp. 429–436, 2006.

APPENDIX

A. Proof of Lemma V.3

We will first prove that $P_i[k+1] \preceq AP_j[k]A^\top + Q$, in case $P_{i_f}[k] \preceq P_j[k]$ holds, both for all $j \in \mathcal{N}_{i(0,1)}$ and $k \in \mathbb{Z}_+$. Algorithm V.I gives that $P_i^{-1}[k+1] = (AP_{i_f}[k]A^\top + Q)^{-1} + C_i^\top R_i^{-1} C_i$. As $R_i^{-1} \succ 0$ and thus $C_i^\top R_i^{-1} C_i \succeq 0$, it follows that

$$P_i[k+1] \preceq AP_{i_f}[k]A^\top + Q.$$

Therefore, if $P_{i_f}[k] \preceq P_j[k]$, for all $j \in \mathcal{N}_{i(0,1)}$ and $k \in \mathbb{Z}_+$, then also $P_i[k+1] \preceq AP_j[k]A^\top + Q$ holds for all $j \in \mathcal{N}_{i(0,1)}$ and $k \in \mathbb{Z}_+$.

The second part is to show that $P_{i_f}[k] \preceq P_j[k]$ holds for all $j \in \mathcal{N}_{i(0,1)}$ and $k \in \mathbb{Z}_+$. From Algorithm V.I it follows that this holds if $P_{i(l)} \preceq P_{j(l)}$ and $P_{i(l)} \preceq P_{i(l-1)}$ hold for all $l \in \mathbb{Z}_{[1,L]}$. Notice that in the l^{th} fusion-cycle, i.e., with $P_{i(l-1)}$ and $P_{j(l)}$ as the initial covariances, we have that $\Gamma_{(l)}$ and $P_{i(l)}$ represent the mutual and fused covariance, respectively. Hence, the result of Criterion IV.1, which was proven for EI in [9], can be applied by substituting $P_i = P_{i(l-1)}$, $P_j = P_{j(l)}$, $\Gamma = \Gamma_{(l)}$ and $P_{i_f} = P_{i(l)}$. With these substitutions Criterion IV.1 gives that $P_{i(l)} \preceq P_{j(l)}$ and $P_{i(l)} \preceq P_{i(l-1)}$, which completes the proof. ■

B. Proof of Lemma V.4

Let us recall the inequality that needs to be proven, i.e.,

$$P_{i_f}[k] \preceq A^m P_j[k-m](A^m)^\top + \sum_{c=0}^{m-1} A^c Q (A^c)^\top, \qquad (10)$$

for all $j \in \mathcal{N}_{i(m+1)}$ and $k \in \mathbb{Z}_{\geq m}$. Also, let $\mathcal{N}_{h(1)}$ represent the set of all direct neighboring nodes of a node h. Furthermore, let $\mathcal{M} \subset \mathcal{N}$ denote a subset of nodes. Then it follows that $\{\cup_{h\in\mathcal{M}} \mathcal{N}_{h(0,1)}\}$ denotes the set of all the nodes $h \in \mathcal{M}$ and all the nodes j that are a direct neighbor of a node $h \in \mathcal{M}$.

The proof of (10) proceeds by induction. The first step, for which $m = 1$ in (10), is to prove that $P_{i_f}[k] \preceq AP_j[k-1]A^\top + Q$, for all $j \in \mathcal{N}_{i(2)}$ and $k \in \mathbb{Z}_{\geq 1}$. Let a node $h \in \mathcal{N}_{i(1)}$ be given. Substituting node i with node h in Lemma V.3 gives

$$P_h[k] \preceq AP_j[k-1]A^\top + Q, \quad \forall j \in \mathcal{N}_{h(0,1)}, k \in \mathbb{Z}_{\geq 1}. \qquad (11)$$

Notice that, as $h \in \mathcal{N}_{i(0,1)}$, Lemma V.3 implies that $P_{i_f}[k] \preceq P_h[k]$ holds for all $h \in \mathcal{N}_{i(0,1)}$ and $k \in \mathbb{Z}_+$. Combining this inequality with (11) yields that for all $k \in \mathbb{Z}_{\geq 1}$ it holds that

$$P_{i_f}[k] \preceq AP_j[k-1]A^\top + Q, \quad \forall j \in \{\cup_{h\in\mathcal{N}_{i(1)}} \mathcal{N}_{h(0,1)}\}. \qquad (12)$$

Notice that $j \in \{\cup_{h\in\mathcal{N}_{i(1)}} \mathcal{N}_{h(0,1)}\}$ implies that (12) holds for all the nodes j, such that the graph distance $d(v_i, v_j) \leq 2$. This means that $\{\cup_{h\in\mathcal{N}_{i(1)}} \mathcal{N}_{h(0,1)}\} = \{\cup_{m\leq 2} \mathcal{N}_{i(m)}\}$, and because $\mathcal{N}_{i(2)} \subset \{\cup_{m\leq 2} \mathcal{N}_{i(m)}\}$, also (12) holds for all $j \in \mathcal{N}_{i(2)}$.

The second step is to show that (10) holds for any $m \geq 2$, by assuming that (10) holds for $m-1$. This latter assumption gives that, for all $h \in \mathcal{N}_{i(m)}$ and $k \in \mathbb{Z}_{\geq m-1}$, it holds that

$$P_{i_f}[k] \preceq A^{m-1} P_h[k-m+1](A^{m-1})^\top + \sum_{c=0}^{m-2} A^c Q (A^c)^\top. \qquad (13)$$

Substituting node i with node h in Lemma V.3 gives

$$P_h[k-m+1] \preceq AP_j[k-m]A^\top + Q, \quad \forall j \in \mathcal{N}_{h(0,1)}. \qquad (14)$$

When substituting (14) into (13) we have that for all $j \in \{\cup_{h\in\mathcal{N}_{i(m)}} \mathcal{N}_{h(0,1)}\}$ and $k \in \mathbb{Z}_{\geq m}$ it holds that:

$$\begin{aligned} P_{i_f}[k] &\preceq A^{m-1}\big(AP_j[k-m]A^\top + Q\big)(A^{m-1})^\top + \sum_{c=0}^{m-2} A^c Q (A^c)^\top \\ &= A^m P_j[k-m](A^m)^\top + \sum_{c=0}^{m-1} A^c Q (A^c)^\top. \end{aligned} \qquad (15)$$

Similar to the first step, $\mathcal{N}_{i(m+1)} \subset \{\cup_{h\in\mathcal{N}_{i(m)}} \mathcal{N}_{h(0,1)}\}$ holds. Therefore, the inequality of (15) also holds for all $j \in \mathcal{N}_{i(m+1)}$, which completes the proof. ■

C. Proof of Lemma V.5

The equality $P_i[1] = \Sigma_i[1]$ follows directly from the fact that the DEI and the stand-alone LKF use the same equations to determine $P_i[1]$ and $\Sigma_i[1]$, respectively, and $P_i[0] = \Sigma_i[0]$.

The inequality $P_i[k] \preceq \Sigma_i[k]$, for all $k \in \mathbb{Z}_{\geq 2}$, is proven by induction. From Algorithm V.1 and (7) one obtains that

$$\begin{aligned} P_i^{-1}[k] &= \Big(AP_{i_f}[k-1]A^\top + Q\Big)^{-1} + C_i^\top R_i^{-1} C_i, \\ \Sigma_i^{-1}[k] &= \Big(A\Sigma_i[k-1]A^\top + Q\Big)^{-1} + C_i^\top R_i^{-1} C_i. \end{aligned} \qquad (16)$$

The first is step to prove that $P_i[2] \preceq \Sigma_i[2]$. From $P_i[1] = \Sigma_i[1]$ together with Lemma V.3, i.e. $P_{i_f}[k] \preceq P_i[k]$, we have that $P_{i_f}[1] \preceq \Sigma_i[1]$. Hence, it also holds that $(AP_{i_f}[1]A^\top + Q)^{-1} \succeq \big(A\Sigma_i[1]A^\top + Q\big)^{-1}$. Using this inequality in (16) results in $P_i^{-1}[2] \succeq \Sigma_i^{-1}[2]$ and thus, $P_i[2] \preceq \Sigma_i[2]$.

The second step is to prove that $P_i[k] \preceq \Sigma_i[k]$, if $P_i[k-1] \preceq \Sigma_i[k-1]$ holds. Lemma V.3, i.e., $P_{i_f}[k] \preceq P_i[k]$, gives that the latter inequality results in $P_{i_f}[k-1] \preceq \Sigma_i[k-1]$. Hence, it also holds that $(AP_{i_f}[k-1]A^\top + Q)^{-1} \succeq \big(A\Sigma_i[k-1]A^\top + Q\big)^{-1}$. Using this inequality in (16) gives that $P_i^{-1}[k] \succeq \Sigma_i^{-1}[k]$ and thus, $P_i[k] \preceq \Sigma_i[k]$, which completes the proof. ■