14th International Conference on Information Fusion
Chicago, Illinois, USA, July 5-8, 2011

Empirical case-studies of state fusion via ellipsoidal intersection

Joris Sijs
TNO Technical Sciences
Delft, The Netherlands
Email: joris.sijs@tno.nl

Mircea Lazar
Eindhoven University of Technology
Eindhoven, The Netherlands
Email: m.lazar@tue.nl

Abstract— **This article presents a practical assessment of the recently developed state fusion method *ellipsoidal intersection* and focusses on distributed state estimation in sensor networks. It was already proven that this fusion method combines strong fundamental properties with attractive features in accuracy and computational requirements. However, these features were derived for linear processes with observability of the state vector in at least one of the local measurements. Therefore, several empirical case-studies are performed to assess ellipsoidal intersection with respect to three real-life limitations. A scenario of cooperative adaptive cruise control is used to analyze the *absence of observability* in any local measurement. Furthermore, the Van-der-Pol oscillator and a benchmark application of tracking shockwaves on highways assess the fusion method for *nonlinear process models*. The latter example is also used in a set-up where the employed state estimation methodology differs per node, so to meet *different computational requirements per node*.**

I. Introduction

Some well known state-estimators for a process with Gaussian noise distributions are the Kalman filter (KF), extended Kalman filter (EKF) and unscented Kalman filter (UKF), as presented in [1]–[3]. Their centralized algorithms estimate the global state of a process based on all the measurements. Nowadays, measurements are often acquired by a network of sensor-nodes, also known as a sensor network, e.g., [4]. Employing a centralized state-estimator requires global communication and central data-processing, which is likely to become infeasible for large-scale sensor networks. An upcoming solution is distributed state estimation (DSE), e.g., [5]–[8]. Therein, a distributed strategy is proposed to decrease communication and computational requirements per node.

In DSE each node typically performs an estimation algorithm, such as the KF, to process the local measurement. Thereby obtaining a local estimate of the global state. Communication between nodes is used to attain the main objective of DSE: *achieve stability of local estimates and improve local estimation accuracies*. Stability in the sense of estimation refers to a bounded covariance of the modeled estimation error. A solution is for each node to fuse the local estimate of its estimation algorithm with the ones received from neighboring nodes. Among the existing fusion approaches, e.g., [9]–[13], the method *ellipsoidal intersection* of [13] guarantees an improvement in accuracy while obtaining low computational complexity. Also, the theoretical study on DSE of [14], by combining the KF and *ellipsoidal intersection*, proved stability of local estimates in every node given that the state is observable in at least one local measurement.

The main contribution of this article is to extend the theoretical analysis of *ellipsoidal intersection* with three empirical DSE case-studies. Each case is characterized by a recurring practical limitation: (*i*) absence of observability in *all* local measurements, (*ii*) nonlinear process-models and (*iii*) different computational limitations per node. A cooperative adaptive cruise control scenario is employed to assess the first limitation, i.e., the global state is not observable in any of the local measurements. The second and third case assume observability but are concerned with the nonlinear process models of a Van-der-Pol oscillator and the one for tracking shockwaves on a highway. Hence, nodes employ the EKF or UKF next to state fusion. Moreover, the third scenario analyses the feasibility of different computational requirements per node in a network of heterogenous estimators, i.e., some nodes perform the EKF while others employ the UKF. All three case-studies demonstrated the stability objective as well as a high estimation accuracy, due to which *ellipsoidal intersection* proves to be an attractive fusion method for real-life DSE.

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 $\mathbb{Z}_{\mathcal{C}} := \mathbb{Z} \cap \mathcal{C}$. Let 0 denote a zero number, or a vector or matrix with all elements equal to zero. Its dimension will be clear from the context. Similarly, I_n denotes an $n \times n$ identity matrix of appropriate dimensions. The transpose, inverse (if exists) and determinant of a matrix $A \in \mathbb{R}^{n\times n}$ are denoted as $A^\top$, A^{-1} and $|A|$ respectively. Further, $[A]_{qr}$ denotes the element on the q-th row and r-th column of A and similarly, $[x]_q$ denotes the q-th element of a vector $x \in \mathbb{R}^n$. Given that $A, B \in \mathbb{R}^{n\times n}$ are positive definite, denoted with $A \succ 0$ and $B \succ 0$ (or $A, B \succ 0$ in short), 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 $(A^{\frac{1}{2}})^{-1}$. Suppose that $A \in \mathbb{R}^{n\times n}$ is a matrix with real eigenvectors, i.e., $\nu_q(A) \in \mathbb{R}^n$, and eigenvalues, i.e., $\lambda_q(A) \in \mathbb{R}$, for all $q \in \mathbb{Z}_{[1,n]}$. Then the eigenvalue decomposition of A, i.e., $A = SDS^{-1}$, is obtained as $S := (\nu_1(A)\ \nu_2(A)\ \ldots\ \nu_n(A))$ and $D :=$

$\text{diag}(\lambda_1(A),\ldots,\lambda_n(A))$, i.e., $[S]_{qr}=[\nu_r(A)]_q$, $[D]_{qr}=\lambda_q(A)$ if $q=r$ and $[D]_{qr}=0$ if $q\neq r$, for all $q,r\in\mathbb{Z}_{[1,n]}$.

For a continuous differentiable function $f(x,y):\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^l$, the Jacobian matrix of $f(x,y)$ towards x and towards y is denoted as $\nabla_x f\in\mathbb{R}^{l\times n}$ and $\nabla_y f\in\mathbb{R}^{l\times m}$, respectively. Moreover, $\nabla_x f(a,b)\in\mathbb{R}^{l\times n}$ denotes the value of $\nabla_x f$ in case $x=a$ and $y=b$. The Gaussian function (Gaussian in short) is denoted as $G(x,\hat{x},P)$, for some $x,\hat{x}\in\mathbb{R}^n$ and $P\in\mathbb{R}^{n\times n}$. If $G(x,\hat{x},P)$ is the probability density function (PDF) of a random vector x, then by definition the mean and covariance of x are $\hat{x}$ and P, respectively. Moreover, P^{-1} is a measure for the accuracy of $\hat{x}$ as an estimated value of x. Any $G(x,\hat{x},P)$ can be represented by its unitary sub-level-set $\mathcal{E}_{\hat{x},P}\subset\mathbb{R}^n$, which is an ellipsoidal set defined as $\mathcal{E}_{\hat{x},P}:=\{x|(x-\hat{x})^\top P^{-1}(x-\hat{x})\leq 1\}$. Some abbreviations of employed state estimation set-ups are:

- cEKF, a centralized state estimation set-up of the EKF;
- cUKF, a centralized state estimation set-up of the UKF;
- dEKF, a DSE set-up as depicted in Figure 1, where each node performs the EKF as LSE;
- dUKF, a DSE set-up as depicted in Figure 1, where each node performs the UKF as LSE;
- HDSE, a heterogenous DSE set-up as depicted in Figure 1, where some nodes perform the EKF as LSE, while other nodes employ the UKF.

III. Problem formulation

Let us assume an autonomous process that is observed by a sensor network, for which $\mathcal{N}\subset\mathbb{Z}$ denotes the set of node indexes. The state-vector of the process is denoted as $x\in\mathbb{R}^n$, whereas the local measurements of a node $i\in\mathcal{N}$ are collected in the measurement-vector $y_i\in\mathbb{R}^{l_i}$, for some $l_i\in\mathbb{Z}_{\geq 1}$. In case the sample instants are denoted with $k\in\mathbb{Z}_+$, then the discrete-time process-model for any node i, given $f:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ and $h_i:\mathbb{R}^n\to\mathbb{R}^{l_i}$, is described as follows

$$x(k+1)=f(x(k),w(k)), \tag{1a}$$
$$y_i(k)=h_i(x(k))+v_i(k). \tag{1b}$$

The process-noise $w\in\mathbb{R}^m$ and the measurement-noise $v_i\in\mathbb{R}^{l_i}$ are characterized by a zero-mean Gaussian PDF, i.e.,

$$p(w(k)):=G(w(k),0,Q),\ p(v_i(k)):=G(v_i(k),0,V_i)\ \forall k\in\mathbb{Z}_+.$$

The sensor network employs a DSE strategy according to the schematic set-up of Figure 1. Therein, each node i calculates a local estimate of the state at each sample instant k by processing y_i in a "local state-estimator" (LSE). The resulting PDF is denoted as $p_i(x(k))=G(x(k),\hat{x}_i(k),P_i(k))$, for some $\hat{x}_i\in\mathbb{R}^n$ and $P_i\in\mathbb{R}^{n\times n}$. By exchanging local estimates, node i receives $p_j(x(k))$ from the nodes $j\in\mathcal{N}_i\subset\mathcal{N}$. These received PDFs are then merged with $p_i(x(k))$ in a "local state fusion" (LSF) algorithm, resulting in the PDF $p_{i_f}(x(k))=G(x(k),\hat{x}_{i_f}(k),P_{i_f}(k))$, for some $\hat{x}_{i_f}\in\mathbb{R}^n$ and $P_{i_f}\in\mathbb{R}^{n\times n}$.

The main objective of DSE is achieving stability of local estimates, i.e., a bounded covariance P_i for all $i\in\mathcal{N}$, and improving their accuracy. To that extent, a local measurement y_i should have the ability to improve the local estimate of any other node $j\in\mathcal{N}$. A DSE approach that enjoys this property is referred to as *global covariance DSE*. Whether this property is obtained depends on the fusion method, i.e., both $P_{i_f}\preceq P_i$ and $P_{i_f}\preceq P_j$ should hold.

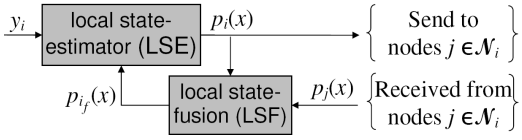

Fig. 1. Schematic set-up of the local algorithm at node i.

A common assumption for sensor networks is that correlations of PDFs will not be available, as keeping track of the shared estimates between nodes is intractable. Some examples of existing fusion methods that can cope with unknown correlations are found in [9]–[13]. A popular approach is known as *covariance intersection*, e.g., [9], which defines the fusion result as a convex combination of the original PDFs, i.e., $P_{i_f}=\omega P_i+(1-\omega)P_j$ and $\hat{x}_{i_f}=P_{i_f}(\omega P_i^{-1}\hat{x}_i+(1-\omega)P_j^{-1}\hat{x}_j)$ for some $\omega\in[0,1]$. As a result, P_{i_f} of *covariance intersection* is a conservative overapproximation of the actual covariance after fusion. The *ellipsoidal intersection* method of [13] is less conservative and satisfies the global covariance condition, i.e., $P_{i_f}\preceq P_i$ and $P_{i_f}\preceq P_j$. Moreover, since the method is computationally tractable as well, *ellipsoidal intersection* is employed as fusion method of the proposed DSE set-up.

The main issue treated in this article is how to illustrate the impact of state fusion in practical set-ups of DSE. To that extent, a detailed description of the proposed DSE algorithm will be presented next. This algorithm is then employed in three different case-studies by addressing three practical concerns, i.e., absence of state-observability in any y_i for all $i\in\mathcal{N}$, nonlinear process-models and different computational requirements per node (a network with heterogenous LSEs). The impact of *ellipsoidal intersection* is assessed is each case via the estimation accuracy and global covariance property.

IV. A distributed state-estimator

This section presents the overall algorithm of a node i according to the DSE set-up of Figure 1. A description of three different LSE-algorithms is given first, after which state fusion according to *ellipsoidal intersection* is presented. A note on estimation is that $p(x(k))$ is commonly calculated from the the previous instant $k-1$, which differs from the model of (1). Hence, an estimator is initialized for some $\hat{x}(-1)$ and $P(-1)$.

A. Local state estimation

The LSE of a node i performs an measurement update on $p_{i_f}(x(k-1))$ given $y_i(k)$, for all $k\in\mathbb{Z}_+$, in a KF, EKF or UKF.

1) (Extended) Kalman filter: Employing the KF or EKF as LSE requires a linear approximation of (1) in the state-space description. For the KF, such a process-model correspond to

$$\begin{aligned}x(k)&=Fx(k-1)+Ew(k-1),\\ y_i(k)&=H_ix(k)+v_i(k),\end{aligned} \tag{2}$$

for some $F \in \mathbb{R}^{n\times n}$, $E \in \mathbb{R}^{n\times m}$ and $H_i \in \mathbb{R}^{l_i\times n}$. Let $\hat{x}_i(k^-) \in \mathbb{R}^n$ and $P_i(k^-) \in \mathbb{R}^{n\times n}$ denote the predicted mean and error-covariance at node i at sample instant k, respectively. Then the KF calculates the updated $\hat{x}_i(k)$ and $P_i(k)$ as follows,

$$\begin{aligned}
&\hat{x}_i(k^-) = F\hat{x}_{i_f}(k-1),\\
&P_i(k^-) = FP_{i_f}(k-1)F^\top + EQE^\top,\\
&K_i(k) = P_i(k^-)H_i^\top\left(H_iP_i(k^-)H_i^\top + V_i\right)^{-1},\\
&\hat{x}_i(k) = \hat{x}_i(k^-) + K_i(k)\left(y_i(k) - H_i\hat{x}_i(k^-)\right),\\
&P_i(k) = \left(I - K_i(k)H_i\right)P_i(k^-).
\end{aligned} \tag{3}$$

Performing the above algorithm as LSE results in low computational requirements. Moreover, the KF is known to compute the optimal estimate, provided $p(w(k))$ and $p(v_i(k))$ are Gaussian and the process-model of (1) is linear. In case of nonlinear models the EKF is an alternative estimator for improving accuracy. Therein, the model-parameters depend on k and are defined via the Jacobian matrices of both nonlinear model-functions at the current working point, i.e., $F_i(k) := \nabla_x f(\hat{x}_{i_f}(k-1),0)$, $E_i(k) := \nabla_w f(\hat{x}_{i_f}(k-1),0)$ and $H_i(k) := \nabla_x h_i(\hat{x}_{i_f}(k-1))$. The algorithm of the EKF at node i is similar to (3), by substituting $F = F_i(k)$, $E = E_i(k)$ and $H = H_i(k)$, while employing $\hat{x}_i(k^-) = f(\hat{x}_{i_f}(k-1),0)$. However, accuracy of an EKF depends on the support to linearize the process-model of (1) at each sample instant. When estimation results are not satisfactory, the EKF can be replaced with the UKF.

2) Unscented Kalman filter: In case an UKF is employed as LSE, then the nonlinear model of (1) is applied to various values of $x(k-1)$ and $w(k-1)$, also referred to as "sigma-values". These values are selected from an augmented vector $\mu \in \mathbb{R}^{n+m}$ that combines the state and process noise, i.e., $\mu := \binom{x}{w}$. Since $x(k-1)$ and $w(k-1)$ at a node i are described by Gaussian PDFs, $p_i(\mu(k-1)) := G(\mu(k-1), \hat{\mu}_i(k-1), U_i(k-1))$ is also Gaussian, for some mean $\hat{\mu}_i \in \mathbb{R}^{n+m}$ and covariance $U_i \in \mathbb{R}^{(n+m)\times(n+m)}$. Values of this mean and covariance follow from $p_{i_f}(x(k-1))$ and $p(w(k-1))$, i.e.,

$$\hat{\mu}_i(k-1) := \begin{pmatrix}\hat{x}_{i_f}(k-1)\\ 0\end{pmatrix},\ U_i(k-1) := \begin{pmatrix}P_{i_f}(k-1) & 0\\ 0 & Q\end{pmatrix}.$$

The PDF $p_i(\mu(k-1))$ is then used to select $M := 2(n+m)+1$ different values of $\mu(k-1)$, which are denoted as $\hat{\mu}_{i,q}(k-1) \in \mathbb{R}^{n+m}$ for all $q \in \mathbb{Z}_{[1,M]}$. Let $\tilde{\mu}_{i,d} \in \mathbb{R}^{n+m}$ be defined as the d-th column of $U_i^{\frac{1}{2}}(k-1)$, i.e., $[\tilde{\mu}_{i,d}]_r := [U_i^{\frac{1}{2}}(k-1)]_{rd}$ for all $r,d \in \mathbb{Z}_{[1,n+m]}$. Then the "sigma-values" $\hat{\mu}_{i,q}(k-1)$, for all $q \in \mathbb{Z}_{[1,M]}$ and for some $c \in \mathbb{R}_+$, are defined as follows

$$\hat{\mu}_{i,q}(k-1) := \begin{cases}\hat{\mu}_i(k-1) + c\tilde{\mu}_{i,q} & \text{if } q \in \mathbb{Z}_{[1,n+m]},\\ \hat{\mu}_i(k-1) - c\tilde{\mu}_{i,(q-n-m)} & \text{if } q \in \mathbb{Z}_{[n+m+1,M-1]},\\ \hat{\mu}_i(k-1) & \text{if } q = M.\end{cases}$$

The process-model of (1) is performed on each "sigma-value" to obtain predictions of $x(k)$ and $y_i(k)$, for all $q \in \mathbb{Z}_{[1,M]}$, i.e.,

$$\hat{x}_{i,q}(k^-) := f\left(\hat{\mu}_{i,q}(k-1)\right) \text{ and } \hat{y}_{i,q}(k^-) := h_i\left(\hat{x}_{i,q}(k^-)\right).$$

In case $\hat{x}_i(k^-) \in \mathbb{R}^n$ and $P_i(k^-) \in \mathbb{R}^{n\times n}$ denote the predicted mean and error-covariance of $x(k)$, respectively, then the updated $\hat{x}_i(k)$ and $P_i(k)$ according to the UKF, for some $\omega_q \in \mathbb{R}_+$, $R_i(k) \in \mathbb{R}^{l_i\times l_i}$ and $S_i(k) \in \mathbb{R}^{n\times l_i}$, yields

$$\begin{aligned}
&\hat{x}_i(k^-) = \sum_{q=1}^{M}\omega_q\hat{x}_{i,q}(k^-), \quad \hat{y}_i(k^-) = \sum_{q=1}^{M}\omega_q\hat{y}_{i,q}(k^-),\\
&\hat{x}_i(k) = \hat{x}_i(k^-) + S_i(k)\left(R_i(k)+V_i\right)^{-1}\left(y_i(k) - \hat{y}_i(k^-)\right),\\
&P_i(k) = P_i(k^-) - S_i(k)\left(R_i(k)+V_i\right)S_i^\top(k).
\end{aligned} \tag{4}$$

Where,

$$\begin{aligned}
P_i(k^-) &= \sum_{q=1}^{M}\omega_q\left(\hat{x}_{i,q}(k^-) - \hat{x}_i(k^-)\right)\left(\hat{x}_{i,q}(k^-) - \hat{x}_i(k^-)\right)^\top,\\
R_i(k) &= \sum_{q=1}^{M}\omega_q\left(\hat{y}_{i,q}(k^-) - \hat{y}_i(k^-)\right)\left(\hat{y}_{i,q}(k^-) - \hat{y}_i(k^-)\right)^\top,\\
S_i(k) &= \sum_{q=1}^{M}\omega_q\left(\hat{x}_{i,q}(k^-) - \hat{x}_i(k^-)\right)\left(\hat{y}_{i,q}(k^-) - \hat{y}_i(k^-)\right)^\top.
\end{aligned}$$

Common values for the constant c and the weights ω_q, for some $\alpha \in \mathbb{R}_+$, are $c = \sqrt{n+m+\alpha}$,

$$\omega_M = \frac{\alpha}{n+m+\alpha} \text{ and } \omega_q = \frac{1}{2(n+m+\alpha)},\ \forall q \in \mathbb{Z}_{[1,M-1]}.$$

Estimating x of nonlinear processes via an UKF results in a low estimation error at the cost of high computational requirements. Therefore, a trade-off must be made between accuracy and computational complexity to decide which estimator is employed as LSE. Before the overall algorithm of a node i is given, let us first present *ellipsoidal intersection*.

B. State fusion according to ellipsoidal intersection

This section summarizes the recently developed state fusion method for two PDFs *ellipsoidal intersection*, as presented in [13]. The method fuses $p_i(x) := G(x,\hat{x}_i,P_i)$ and $p_j(x) := G(x,\hat{x}_j,P_j)$ into a single PDF that is denoted as $p_{i_f}(x)$, 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 is that correlations are parameterized via exclusive and mutual information of $p_i(x)$ and $p_j(x)$ *a priori* to deriving a fusion formula via estimation theory. Mutual implies that, for example, the same measurements or process-model parameters were used in both $p_i(x)$ and $p_j(x)$. Similarly, exclusive information refers to, for example, measurements that were used in either $p_i(x)$ or $p_j(x)$. To that extent, let us introduce the following parametrization.

- Let $p_\gamma(x) := G(x,\gamma,\Gamma)$, for some $\gamma \in \mathbb{R}^n$ and $\Gamma \in \mathbb{R}^{n\times n}$, denote the estimate of x based on the mutual information of $p_i(x)$ and $p_j(x)$;
- Let $p_{j_e}(x) := G(x,\theta_j,\Theta_j)$, for some $\theta_j \in \mathbb{R}^n$ and $\Theta_j \in \mathbb{R}^{n\times n}$, denote the estimate of x based on the exclusive information of $p_j(x)$ only.

Then $p_j(x)$ is as defined as the update of the mutual PDF $p_\gamma(x)$ with the exclusive PDF $p_{j_e}(x)$. Since $p_\gamma(x)$ and $p_{j_e}(x)$ are uncorrelated, the results of [5] give that

$$P_j = (\Gamma^{-1} + \Theta_j^{-1})^{-1} \text{ and } \hat{x}_j = P_j(\Gamma^{-1}\gamma + \Theta_j^{-1}\theta_j). \tag{5}$$

Similar developments in estimation theory define $p_{i_f}(x)$ *by updating* $p_i(x)$ *with the exclusive PDF* $p_{j_e}(x)$. Hence, the same results of [5] imply that $p_{i_f}(x)$ is characterized by $P_{i_f} = (P_i^{-1} + \Theta_j^{-1})^{-1}$ and $\hat{x}_{i_f} = P_{i_f}(P_i^{-1}\hat{x}_i + \Theta_j^{-1}\theta_j)$. Substituting the resulting θ_j and Θ_j, as obtained from (5), into these expressions of P_{i_f} and $\hat{x}_{i_f}$ gives an explicit fusion update, i.e.,

$$\begin{aligned} P_{i_f} &= \left(P_i^{-1} + P_j^{-1} - \Gamma^{-1}\right)^{-1}, \\ \hat{x}_{i_f} &= P_{i_f}\left(P_i^{-1}\hat{x}_i + P_j^{-1}\hat{x}_j - \Gamma^{-1}\gamma\right). \end{aligned} \tag{6}$$

The second step is determining the values of γ and Γ when correlation is unknown. To obtain a robust update of $p_i(x)$ with $p_{j_e}(x)$, the following hypothesis is employed: *the PDF that parameterizes the correlation of* $p_i(x)$ *and* $p_j(x)$*, i.e.,* $p_\gamma(x)$*, is as accurate as possible*. Let us start by deriving a value for the mutual covariance, before is continued with the mutual mean.

1) Mutual covariance: A higher accuracy of $p_\gamma(x)$ is equivalent to a reduction of the eigenvalues $\lambda_q(\Gamma)$, for some $q \in \mathbb{Z}_{[1,n]}$. Hence, maximizing the accuracy is equivalent to minimizing $\sum_{q=1}^{n}\lambda_q(\Gamma)$. However, the accuracy of $p_\gamma(x)$ cannot exceed the accuracy of $p_j(x)$, as the latter one is an update of $p_\gamma(x)$ with $p_{j_e}(x)$. A mathematical expression of this statement, which follows from (5) and the fact that $\Theta_j \succeq 0$ of $p_{j_e}(x)$, is that $\Gamma \succeq P_j$ holds. Similarly, $\Gamma \succeq P_i$ must also hold. Let $\mathcal{E}_{0,P_i}$, $\mathcal{E}_{0,P_j}$ and $\mathcal{E}_{0,\Gamma}$ denote the sub-level-sets that correspond to these three covariances. Then $\Gamma \succeq P_i$ and $\Gamma \succeq P_j$ can also be expressed as $\mathcal{E}_{0,P_i} \cup \mathcal{E}_{0,P_j} \subseteq \mathcal{E}_{0,\Gamma}$. All together, a formal definition of the mutual covariance is stated as follows

$$\begin{aligned} \Gamma := \arg\min_{\Upsilon \in \mathbb{R}^{n\times n}} & \sum_{q=1}^{n}\lambda_q(\Upsilon) \\ \text{subject to} \quad & \mathcal{E}_{0,P_i} \cup \mathcal{E}_{0,P_j} \subseteq \mathcal{E}_{0,\Upsilon}. \end{aligned} \tag{7}$$

Basically, the above expression of Γ defines the sub-level-set of the mutual covariance, i.e., $\mathcal{E}_{0,\Gamma}$, as the smallest ellipsoid to enclose the sub-level-sets of the original estimates, i.e., $\mathcal{E}_{0,P_i}$ and $\mathcal{E}_{0,P_j}$. To solve the minimization problem of (7), let the diagonal matrices $D_i, D_j \in \mathbb{R}^{n\times n}$ and rotational matrices $S_i, S_j \in \mathbb{R}^{n\times n}$ be introduced via the eigenvalue decompositions

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

Then an explicit formula of the mutual covariance, yields

$$\Gamma = S_i D_i^{\frac{1}{2}} S_j D_\Gamma S_j^{-1} D_i^{\frac{1}{2}} S_i^{-1}, \tag{8}$$

$$[D_\Gamma]_{qr} := \begin{cases} \max([D_j]_{qr}, 1) & \text{if} \quad q = r, \\ 0 & \text{if} \quad q \neq r. \end{cases} \tag{9}$$

2) Mutual mean: The mutual mean represents an agreement between $\hat{x}_i$ and $\hat{x}_j$. Typically, this means that γ is characterized by minimization of the Euclidian distance of $\gamma - \hat{x}_i$ and $\gamma - \hat{x}_j$. As such, a cost-function $J : \mathbb{R}^n \to \mathbb{R}_+$ is defined, for some suitable $W_i, W_j \succeq 0$, whose minimum corresponds to γ, i.e.,

$$\gamma := \arg\min_{\upsilon \in \mathbb{R}^n} J(\upsilon), \tag{10a}$$

$$J(\upsilon) := (\upsilon - \hat{x}_i)^\top W_i (\upsilon - \hat{x}_i) + (\upsilon - \hat{x}_j)^\top W_j (\upsilon - \hat{x}_j). \tag{10b}$$

The use of the weighting matrices W_i and W_j is to enable a different accuracy for each element in $\gamma - \hat{x}_i$ and $\gamma - \hat{x}_j$. Since any variation in the accuracy of γ, $\hat{x}_i$ and $\hat{x}_j$ is caused by exclusive information, the mutual mean γ is determined according to the following reasoning: *if* $p_i(x)$ *has a high exclusive accuracy, then* γ *should be close to* $\hat{x}_j$ *and, vice versa,* γ *should be close to* $\hat{x}_i$ *in case* $p_j(x)$ *has a high exclusive accuracy.* A particular definition of the weighting matrices that is in line with this reasoning is the following,

$$W_i = P_j^{-1} - \Gamma^{-1} \text{ and } W_j = P_i^{-1} - \Gamma^{-1}. \tag{11}$$

The above weights employ $\Theta_j^{-1} = P_j^{-1} - \Gamma^{-1}$ and $\Theta_i^{-1} := P_i^{-1} - \Gamma^{-1}$, which are a measure for the accuracy of exclusive information of $p_j(x)$ and $p_i(x)$, respectively. When solving (10) one obtains that $\gamma = (W_i + W_j)^{-1}(W_i\hat{x}_i + W_j\hat{x}_j)$. However, this solution is valid for a cost-function $J(\upsilon)$ that is convex, i.e., $W_i + W_j \succ 0$ holds. Therefore, a small approximation is applied to W_i and W_j of (11) in case $W_i + W_j \succeq 0$. To that extent, let $B := P_i^{-1} + P_j^{-1} - 2\Gamma^{-1}$, let $\lambda_{\min}(B) > 0$ be defined as the smallest positive eigenvalue of B and let $\beta > 0$ denote a design parameter of the approximation. Then the explicit formula of the mutual mean is given as follows

$$\begin{aligned} \gamma = & \left(P_i^{-1} + P_j^{-1} - 2\Gamma^{-1} + 2\eta I_n\right)^{-1} \times \\ & \left(\left(P_j^{-1} - \Gamma^{-1} + \eta I_n\right)\hat{x}_i + \left(P_i^{-1} - \Gamma^{-1} + \eta I_n\right)\hat{x}_j\right), \end{aligned} \tag{12}$$

$$\eta := \begin{cases} 0 & \text{if} \quad |B| \neq 0, \\ \beta \ll \lambda_{\min}(B) & \text{if} \quad |B| = 0. \end{cases} \tag{13}$$

The interested reader is referred to [13], [14] for more details on *ellipsoidal intersection* and its performance with respect to *covariance intersection*. This article continues by pointing out an important property of *ellipsoidal intersection*, after which the overall algorithm is presented.

Remark IV.1 The result of $p_{i_f}(x)$ should be the same when fusing $p_j(x)$ with $p_i(x)$, instead of $p_i(x)$ with $p_j(x)$, i.e., switch $P_i \leftrightarrow P_j$ and $\hat{x}_i \leftrightarrow \hat{x}_j$. The fusion update of (6) guarantees this property given that Γ and γ obtain the same values. This condition is met by the mutual covariance, since $\mathcal{E}_{0,P_i} \cup \mathcal{E}_{0,P_j} = \mathcal{E}_{0,P_j} \cup \mathcal{E}_{0,P_i}$ in the definition of Γ of (7). Also, the mutual mean satisfies the condition, as switching of i and j does not affect the cost-function $J(\upsilon)$ of (10b) in combination with (11).

C. Overall algorithm of a node i

The schematic set-up of Figure 1 shows that at each sample instant k a node i performs the LSE to calculate $p_i(x(k))$. 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 $p_j(x(k))$, for any $j \in \mathcal{N}_i$, 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 $L := \sharp\mathcal{N}_i$, 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 thus $p_{i_f}(x(k)) := p_{i(L)}(x)$. In case "LocalStateEst" denotes the algorithm that corresponds to one of the employed LSEs, i.e., a KF, EKF or UKF, then the algorithm that is performed by a node i, yields

Algorithm IV.2 ***DSE at node*** i

$[\hat{x}_i(k), P_i(k)] = \text{LocalStateEst}(\hat{x}_{i_f}(k-1), P_{i_f}(k-1), y_i(k))$;
$\hat{x}_{i(0)} = \hat{x}_i(k), \quad P_{i(0)} = P_i(k)$;
for $l = 1, \ldots, L$, do:
$\hat{x}_{j(l)} = \hat{x}_j(k), \quad P_{j(l)} = P_j(k), \qquad j \in \mathcal{N}_i$;
$\Gamma_{(l)} = \text{MutualCovariance}(P_{i(l-1)}, P_{j(l)}), \qquad (8)$;
$\gamma_{(l)} = \text{MutualMean}(P_{i(l-1)}, P_{j(l)}, \Gamma_{(l)}, \hat{x}_{i(l-1)}, \hat{x}_{j(l)}), \quad (12)$;
$P_{i(l)} = \left(P^{-1}_{i(l-1)} + P^{-1}_{j(l)} - \Gamma^{-1}_{(l)}\right)^{-1}$;
$\hat{x}_{i(l)} = P_{i(l)}\left(P^{-1}_{i(l-1)}\hat{x}_{i(l-1)} + P^{-1}_{j(l)}\hat{x}_{j(l)} - \Gamma^{-1}_{(l)}\gamma_{(l)}\right)$;
end
$\hat{x}_{i_f}(k) = \hat{x}_{i(L)}, \quad P_{i_f}(k) = P_{i(L)}$; □

Now that the developed DSE is completed, let us analyze the impact of *ellipsoidal intersection* from a practical point of view for the three introduced case-studies.

V. Case 1: Absence of local observability

In this case-study the process-model of (1) is assumed to be linear, i.e., it follows the description of (2) and the LSE employs a KF. A practical limitation, which can occur in a deployed sensor network, is that x is not observable in *any* of the local measurements. The criteria for local observability at a node i is that (A, H_i) is an observable-pair, in which $A \in \mathbb{R}^{n\times n}$ is defined via the time-continuous process-model of the state, i.e., $\dot{x} = Ax + w$. Not satisfying this criteria implies that some eigenvalues of the (modeled) error-covariance P_i become unbounded when node i estimates x based on y_i only. Therefore, $p_i(x)$ must exploit all the measurement-information within the network via a global covariance DSE for attaining stable local estimates at the different nodes, i.e., $\lambda_q(P_i)$ is bounded for all nodes $i \in \mathcal{N}$ and all $q \in \mathbb{Z}_{[1,n]}$. The considered application is a benchmark example of DSE for cooperative adaptive cruise controllers [15].

Case-study

Consider a platoon of four vehicles having cooperative adaptive cruise controllers. Each vehicle requires the kinematic state values of the leading vehicle in the platoon. Hence, the state-vector x is defined as the position and speed in the X-direction, i.e., $[x]_1$ and $[x]_2$, respectively, and the position and speed in the Y-direction, i.e., $[x]_3$ and $[x]_4$, respectively. The real position of vehicle 1, which starts from $(X,Y) = (10,1)$ and then drives in slalom towards $(105,15)$, is depicted in Figure 2. Therefore, in case the unknown acceleration is represented by process noise, then the discrete-time process model of (1) with a sampling time of 0.1 seconds, yields

$$x(k+1) = \begin{pmatrix} 1 & 0.1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0.1 \\ 0 & 0 & 0 & 1 \end{pmatrix} x(k) + \begin{pmatrix} 0.005 & 0 \\ 0.1 & 0 \\ 0 & 0.005 \\ 0 & 0.1 \end{pmatrix} w(k),$$
$$p(w(k)) = G(w(k), 0, 10I_2).$$

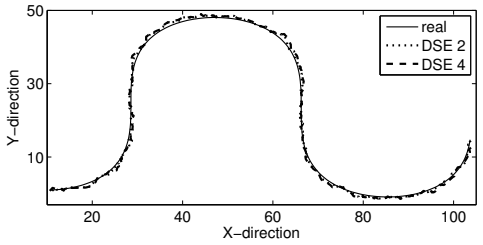

Fig. 2. Position of the leading vehicle versus the estimated position according to vehicle 2 (DSE 2) and vehicle 4 (DSE 4).

Vehicle 1 measure its X-position, whereas vehicle 3 measures the Y-position of vehicle 1. This means that only vehicles 1 and 3 have measurements that depend on the state x, due to which the following local measurements are defined,

$$y_1(k) = H_1 x(k) + v_1(k), \quad p(v_1(k)) = G\big(v_1(k), 0, 0.5\big),$$
$$y_3(k) = H_3 x(k) + v_3(k), \quad p(v_3(k)) = G\big(v_3(k), 0, 0.8\big),$$
$$H_1 = (1 \quad 0 \quad 0 \quad 0), \quad H_3 = (0 \quad 0 \quad 1 \quad 0).$$

Since for the considered example $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, (A, H_i) is not an observable-pair in both $i \in \{1,3\}$. Moreover, vehicles 2 and 4 have no measurements that depend on x, due to which their KF only performs a prediction of the state. Hence, x is not observable in *any* of the local measurements. However, the collection of measurements in the platoon does result in an observable state-vector, i.e., $\left(A, \begin{pmatrix} H_1 \\ H_3 \end{pmatrix}\right)$ is an observable-pair.

Each vehicle i performs Algorithm IV.1 and shares $p_i(x(k))$ with the front and rear vehicle, i.e., $\mathcal{N}_1 = \{2\}$, $\mathcal{N}_2 = \{1,3\}$, $\mathcal{N}_3 = \{2,4\}$ and $\mathcal{N}_4 = \{3\}$. All estimators are initialized by $\hat{x}_i(-1) = (10 \quad 6 \quad 1 \quad 0)^\top$ and $P_i(-1) = 25I_4$, for all $i \in \mathbb{Z}_{[1,4]}$. The estimation results of vehicles two and four are compared in Figure 2 and Figure 3. Figure 3 presents the sum of all eigenvalues of P_i and the squared estimation error, i.e.,

$$\sigma_i(k) := \sum_{q=1}^{4} \lambda_q(P_i(k)), \qquad \forall i \in \mathcal{N}, \qquad (14)$$
$$\Delta_i(k) := (\hat{x}_i(k) - x(k))^\top (\hat{x}_i(k) - x(k)), \quad \forall i \in \mathcal{N}. \qquad (15)$$

Figure 3 shows that the estimation error of vehicles 2 and 4 are comparable, even though both vehicles are forced to rely on neighboring vehicles for estimating x. Moreover, although the results of vehicles 2 and 4 are presented, the eigenvalues of

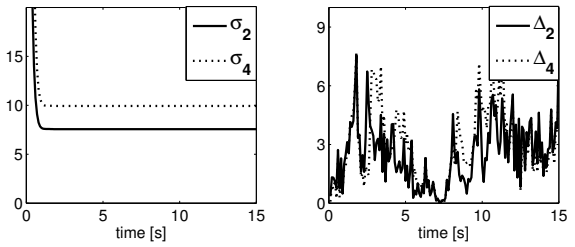

Fig. 3. Modeled estimation error, i.e., $\sigma_i(k)$, and real estimation error, i.e., $\Delta_i(k)$, of the second and fourth vehicle.

$P_i(k)$ for all vehicles converged during the simulation. Hence, the local estimates are stable, which can only occur if any $p_i(x)$ exploits the information of *all* measurements in the network. This is an indication that the developed DSE enjoys the global covariance property, also when the local observability criteria is not met by any node $i \in \mathcal{N}$.

Apart from local observability, other issues for the developed DSE are a result of the assumption that the process-model is linear. Therefore, an analysis of DSE in a set-up with nonlinear models is presented in the next sections.

VI. Case-study 2: Nonlinear process model

Let us assume that the process-model of (1) is nonlinear. Then employing the KF as LSE will result in (highly) inaccurate estimates $p_i(x(k))$ and $p_{i_f}(x(k))$. This can be solved by replacing the KF with an EKF or UKF, since these two methods are designed for nonlinear models and result in a Gaussian PDF that can be used by *ellipsoidal intersection*. The only issue is that the resulting Gaussian PDF $p_i(x(k))$ is suboptimal, since both the EKF and UKF apply an approximation on the update of x to handle nonlinearities. Hence, an empirical case-study of the developed DSE set-up is performed to analyze whether unknown correlations of the approximated PDFs are treated correctly by *ellipsoidal intersection*. To that extent, the DSE is compared to a centralized estimation set-up. Moreover, each node of the sensor network measures a different, unique state-element. Therefore, for such a sensor network consisting of only two nodes any difference between the centralized and distributed solution is then caused by an improper evaluation of correlations in *ellipsoidal intersection*.

Case-study

Let us consider a network of two nodes that observe the two states of a Van-der-Pol oscillator, i.e., $[x(k)]_1$ and $[x(k)]_2$. The discrete-time process-model of (1), with $\delta \in \mathbb{R}_+$ defined as the sampling time, yields

$$x(k+1) = \begin{pmatrix} 1 & \delta \\ 0 & 1+0.5\delta \end{pmatrix} x(k) + \begin{pmatrix} 0 \\ f_2(x(k)) \end{pmatrix} + w(k),$$

where

$$f_2\big(x(k)\big) := \delta \cdot [x(k)]_1 \left(0.5\,[x(k)]_1\,[x(k)]_2 - 1\right).$$

Figure 4 depicts the state values of the Van-der-Pol oscillator in case $x(0) = \left(\begin{smallmatrix} 0.5 \\ 0 \end{smallmatrix}\right)$ and $p(w(k)) = G(w(k), 0, 10^{-3} I_2)$. Furthermore, the local measurements are defined as follows

$$y_1(k) = \begin{pmatrix} 1 & 0 \end{pmatrix} x(k) + v_1(k), \quad p(v_1(k)) = G(v_1(k), 0, 0.8),$$
$$y_2(k) = \begin{pmatrix} 0 & 1 \end{pmatrix} x(k) + v_2(k), \quad p(v_2(k)) = G(v_2(k), 0, 0.5).$$

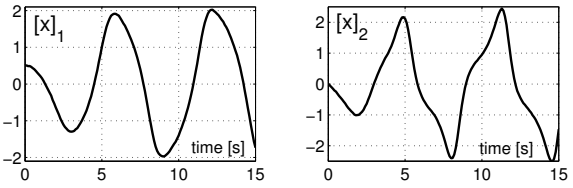

Fig. 4. State values of the Van-der-Pol oscillator.

In the centralized set-up both y_1 and y_2 are sent to a central state-estimator, which can be either an EKF, denoted as cEKF, or an UKF, denoted as cUKF. The resulting PDF of the centralized estimators is denoted as $p(x(k)) = G(x(k), \hat{x}(k), P(k))$, for some $\hat{x} \in \mathbb{R}^n$ and $P \in \mathbb{R}^{n \times n}$. Their performance is compared to the corresponding distributed set-ups, i.e., the developed DSE of Algorithm IV.1. In case both nodes employ the EKF as LSE, then the distributed set-up is denoted as dEKF, whereas dUKF denotes the developed DSE such that the UKF algorithm is performed as LSE. All estimators start with an initial mean of $\left(\begin{smallmatrix} 2 \\ -0.3 \end{smallmatrix}\right)$ and a error-covariance that is equal to $5I_2$. Furthermore, the cUKF and dUKF define $Q = 10^{-3} I_2$. However, since the EKF derives a Jacobian-form of the nonlinear model, the method employs an approximation of process dynamics with a higher inaccuracy. This inaccuracy is modeled via an increased process noise for the cEKF and dEKF, i.e., $Q = 10^{-1} I_2$. The resulting squared estimation error, i.e., $\Delta_i(k)$ of (15) for the dEKF and dUKF, which for the cEKF and cUKF is defined as $\Delta(k) = (\hat{x}(k) - x(k))^\top (\hat{x}(k) - x(k))$, are depicted in Figure 5. Therein, only the error of the first node is presented for each DSE set-up, since *ellipsoidal intersection* guarantees that node 1 and node 2 have equivalent estimates after each fusion step (see Remark IV.1).

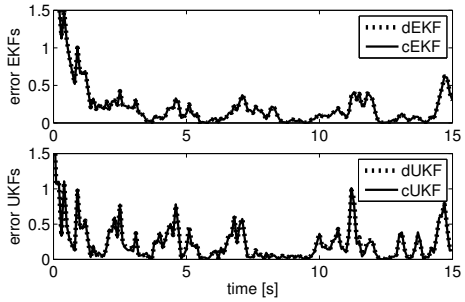

Fig. 5. The squared estimation error $\Delta(k)$ of the cEKF and cUKF (solid lines) and $\Delta_1(k)$ of the dEKF and dUKF (dashed lines).

Figure 5 shows that the centralized and the distributed estimation set-ups have an equivalent performance in accuracy.

This means that the approximation into suboptimal Gaussian PDFs has a negligible effect on *ellipsoidal intersection* as a state fusion method. In both cases the correlation of the original estimates is treated correctly.

A different performance measure than accuracy is the correlation coefficient matrix $\rho(\hat{x},y) \in \mathbb{R}^{n\times l}$ between an estimated state $\hat{x} \in \mathbb{R}^n$ and a measurement $y \in \mathbb{R}^l$. The elements of this matrix are defined as follows

$$[\rho(\hat{x},y)]_{rq} := \frac{cov([\hat{x}]_r,[y]_q)}{\sqrt{cov([\hat{x}]_r)cov([y]_q)}}, \quad \forall q \in \mathbb{Z}_{[1,n]},\ r \in \mathbb{Z}_{[1,l]}.$$

Each element $[\rho(\hat{x},y)]_{rq} \in \mathbb{R}_{[-1,1]}$, for some suitable r and q, is a measure of the correlation between the elements $[\hat{x}]_r$ and $[y]_q$. A value of 1 indicates that the two elements are equivalent, whereas a value of 0 corresponds to no similarity at all. Let us define $y(k) := (y_1(k) \quad y_2(k))^\top$, then the correlation coefficient matrices for the different estimators give the following results,

$$\text{cEKF: } \rho(\hat{x},y) = \begin{pmatrix} 0.85 & 0.07 \\ 0.12 & 0.94 \end{pmatrix},\ \text{dEKF: } \rho(\hat{x}_1,y) = \begin{pmatrix} 0.85 & 0.07 \\ 0.12 & 0.95 \end{pmatrix},$$
$$\text{cUKF: } \rho(\hat{x},y) = \begin{pmatrix} 0.90 & 0.09 \\ 0.11 & 0.96 \end{pmatrix},\ \text{dUKF: } \rho(\hat{x}_1,y) = \begin{pmatrix} 0.90 & 0.09 \\ 0.11 & 0.96 \end{pmatrix}.$$

These correlation coefficient matrices show that the two DSE set-ups use both measurements y_1 and y_2 in the same effective manner to estimate x as their corresponding centralized set-ups. Hence, for this small sensor network with a nonlinear process-model the global covariance property is established. An extended analysis of the developed DSE for a nonlinear process-model, where different nodes can employ different types of LSEs and thus enable different computational requirements per node, is presented next.

VII. Case 3: A network of heterogenous LSEs

Commonly, LSEs of the different nodes in a sensor network are derived from the same type of (centralized) state-estimator, e.g. [5], [6], [16], [17]. The goal of this section is to present a first analysis of a DSE where different nodes perform different types of LSEs, i.e., some nodes will perform the EKF as LSE and other nodes will employ an UKF. Such a heterogeneous DSE (HDSE) set-up allows different computational limitations per node in the network and thus enhances feasibility of DSE in sensor networks. Also, nodes that are added to an existing network can employ arbitrary LSE methodologies, while still exchanging estimates with neighboring nodes for state fusion. The benchmark application for testing this HDSE is tracking shockwaves on a highway.

Case-study

The traffic shockwave is a spatio-temporal dynamical phenomenon typically emerging from high density highway traffic. It is characterized by an increase in vehicle density and a decrease in vehicle speed. Shockwaves "travel" along the highway upstream (i.e. opposite direction to the traffic). This benchmark case-study consists of initiating a shockwave, after which the goal is to track this (simulated) shockwave using aggregated measurements of speed and density within certain road segments. To that extent, consider a stretch of a one-lane road that is divided into 20 segments of each $L = 500$ meter. A total of 5 nodes are used to monitor shockwaves on that particular road. Node 1 is located at road segment 1, node 2 at segment 5, node 3 at segment 10, node 4 at segment 15 and node 5 at road segment 20. Each node exchanges data with direct neighboring nodes, i.e., $\mathcal{N}_1 = \{2\}$, $\mathcal{N}_1 = \{1,3\}$, $\mathcal{N}_3 = \{2,4\}$, $\mathcal{N}_4 = \{3,5\}$ and $\mathcal{N}_5 = \{4\}$.

The discrete-time METANET-model of [18] is used to simulate the shockwave and the corresponding measurements. Therein, $s^n(k) \in \mathbb{R}$ and $\rho^n(k) \in \mathbb{R}$ denote the average speed and density of the n-th road segment at sample instant k. The METANET-model defines a relation of the average speed and density between neighboring segments, for some $\tau, \eta, \kappa, \rho_{crit}, \alpha, v_{free} \in \mathbb{R}$ and sampling-time $\delta \in \mathbb{R}_+$, as follows

$$\rho^n(k+1) = \rho^n(k) + \frac{\delta}{L}\left(\rho^{n-1}(k)s^{n-1}(k) - \rho^n(k)s^n(k)\right),$$
$$s^n(k+1) = s^n(k) + \frac{\delta}{\tau}\left(v_{free}e^{-\frac{1}{\alpha}\left(\frac{\rho^n(k)}{\rho_{crit}}\right)^\alpha} - s^n(k)\right)$$
$$+ \frac{\delta}{L}s^n(k)\left(s^{n-1}(k) - s^n(k)\right) - \frac{\eta\delta}{\tau L}\frac{\rho^{n+1}(k) - \rho^n(k)}{\rho^n(k) + \kappa}.$$

The model parameters that are used in this simulation, yield $\tau = 0.0039$, $\eta = 191$, $\kappa = 254$, $\rho_{crit} = 33.0$, $\alpha = 5.61$, $v_{free} = 89.9$ and $\delta = \frac{10}{3600}$. The resulting shockwave is depicted in Figure 6 and titled as "real". Notice that the wave starts at road segment 20 with an increased vehicle density and then travels towards road segment 1 in approximately 35 minutes. The sensor network set-up is such that each node measures the average speed and density of its own segment, i.e.,

$$y_i(k) = \begin{pmatrix} \rho^{q_i}(k) \\ s^{q_i}(k) \end{pmatrix} + v_i(k) \text{ and } q_i := \begin{cases} 1 & \text{if } i = 1, \\ 5(i-1) & \text{if } i \in \mathbb{Z}_{[2,5]}. \end{cases}$$

Three DSE configurations are employed to recover the average speed and density at all segments based on the five measurements. The first two configurations are the dEKF and dUKF as they were introduced in Section VI. The third configuration implements the HDSE, which is defined by the following LSEs: nodes 1, 3 and 5 employ an UKF, while nodes 2 and 4 perform an EKF-algorithm. All nodes start with equivalent initial values, i.e., $s^n(-1) = 85$ and $\rho^n(-1) = 30$, for all $n \in \mathbb{Z}_{[1,20]}$. Notice, that the METANET-model requires values for $\rho^0(k)$, $\rho^{21}(k)$ and $s^0(k)$. Since this information is not available to the dEKF, dUKF and HDSE, their values are modeled as process noise. Figure 6 shows the real and estimated vehicle density, i.e., ρ, at node 3 according to the dEKF, dUKF and HDSE. The estimated density at other nodes is similar to node 3 and therefore omitted in this section.

Figure 6 shows that the dEKF suffers from deriving a Jacobian-form of the process-model in a sense that the estimated wave tends to "die out" after it was measured. See, for example, a wave that is briefly measured at road segment 15 around 10 minutes. Only when the wave passed segment 10 the dEKF is capable of tracking the wave. Results of the HDSE show that this improper tracking of the dEKF can be solved by replacing the EKF at nodes 1, 3 and 5 with an UKF.

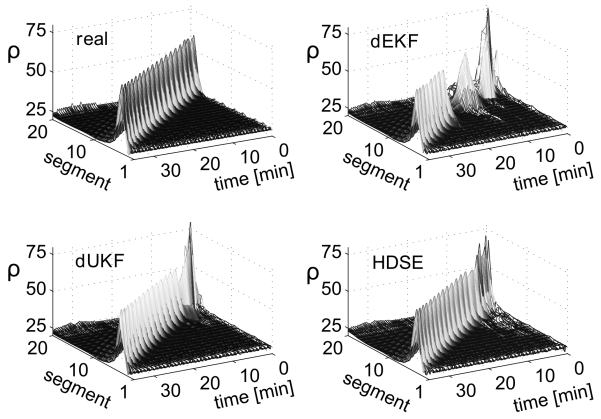

Fig. 6. The real density of all 20 segments in time and their estimated values at node 3 according to the dEKF, dUKF and HDSE.

Already after the first 5 minutes the HDSE has similar results as the dUKF. However, in the long run the dUKF showed less estimation error then the HDSE during simulation.

Notice that even the HDSE enjoys the global covariance property, since node 3, which is located at road segment 10, is able to track the shockwave already from the moment that the wave is firstly measured at node 5. This proves that node 3 uses the information that is made available by the measurement in node 5 and thus that measurements and local estimates throughout the network are correlated. During the simulation nodes that employed an EKF had an average computation-time of 5 [*ms*] per sampling instant, which increased to 20 [*ms*] for nodes that performed an UKF algorithm. Hence, from the fact that different types of LSEs can be employed in different nodes of the network, the HDSE allows to decrease the computational requirements of some nodes compared to the dUKF set-up. This, while remaining a comparable accuracy as the dUKF in the observed shockwaves.

VIII. Conclusions

In this article the impact of the state fusion method *ellipsoidal intersection* was assessed for distributed state estimation (DSE) in sensor networks. To that extent, each node performs a local state estimation algorithm based on its local measurement, e.g., KF, EKF or UKF. The resulting estimate is then fused with the estimates obtained in neighboring nodes by employing the above mentioned fusion method. Three empirical case-studies were performed to analyze *ellipsoidal intersection* on some practical limitations of sensor networks. It was shown in a cooperative adaptive cruise control scenario that the developed DSE can handle a set-up where the state-vector is not observable in any of the local measurements. Also, an illustrative example of the Van-der-Pol oscillator and a benchmark application of tracking shockwaves on highways showed that the developed DSE is suitable for nonlinear process-models. Furthermore, an extension of the latter case-study was used to assess a mixture of LSEs, i.e., some nodes perform the EKF algorithm as LSE and other nodes employ an UKF. This scenario showed that *ellipsoidal intersection* can fuse estimates from various type of state-estimation methodologies in a suitable manner.

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] R. Kandepu, B. Foss, and L. Imsland, "Applying the unscented Kalman filter for nonlinear state estimation," *Journal of Process Control, doi:10.1016/j.jprocont.2007.11.004*, 2008.

[3] S. J. Julier and J. K. Uhlmann, "A new extension of the kalman filter to nonlinear systems," in *Proceedings of AeroSense: The 11th International Symposium on Aerospace.*, Orlando, FL, USA, 1997, pp. 182–193.

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

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

[6] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proceedings of the 46th IEEE Conf. on Decision and Control*, New Orleans, LA, USA, 2007, pp. 5492 – 5498.

[7] S. Julier, "Estimating and Exploiting the Degree of Independent Information in Distributed Data Fusion," in *Proceedings of the 12th International Conference on Information Fusion*, 2009, pp. 772–779.

[8] F. Sawo, F. Beutler, and U. Hanebeck, "Decentralized state estimation of distributed phenomena based on covariance bounds," in *Proceedings of the 17th IFAC World Congress*, Seoul, Korea, 2008.

[9] S. J. Julier and J. K. Uhlmann, "A Non-divergent Estimation Algorithm in the Presence of Uknown Correlations," in *Proceedings of the American Control Conference*, Piscataway, NJ, USA, 1997, pp. 2369–2373.

[10] D. Franken and A. Hupper, "Improved Fast Covariance Intersection for Distributed Data Fusion," in *Proceedings of the 8th Int. Conf. on Information Fusion, d.o.i.: 10.1109/ICIF.2005.1591849*, Philidalphia, PA, USA, 2005.

[11] Y. Zhuo and J. Li, "Data fusion of unknown correlations using internal ellipsoidal approximations," in *Proceedings of the 17th IFAC World Congress*, 2008, pp. 2856–2860.

[12] L. Chen, P. Arambel, and R. Mehra, "Fusion under Unknown Correlation - Covariance Intersection as a Special Case," in *Proceedings of 5th IEEE Int. Conf. on Information Fusion*, 2002, pp. 905–912.

[13] J. Sijs, M. Lazar, and P. v.d. Bosch, "State fusion with unknown correlation: Ellipsoidal intersection," in *Proceedings of the American Control Conference*, Baltimore, USA, 2010, pp. 3992 – 3997.

[14] J. Sijs and M. Lazar, "On the dispersion of the Kalman filtering algorithm," in *Proceeding of the American Control Conference (to appear)*, 2011.

[15] B. van Arem, C. van Driel, and R. Visser, "The impact of cooperative adaptive cruise control on traffic-flow characteristics," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 4, 2006.

[16] U. Khan and J. Moura, "Distributed Kalman filters in sensor networks: Bipartite Fusion Graphs," in *Proceedings of the IEEE 14th Workshop on Statistical Signal Processing*, Madison, Wisconsin, USA, 2007, pp. 700–704.

[17] J. Sijs, M. Lazar, P. Van de Bosch, and Z. Papp, "An overview of non-centralized Kalman filters," in *Proceedings of the IEEE Int. Conference on Conference on Control Applications*, San Antonio, USA, 2008, pp. 739–744.

[18] A. Hegyi, B. De Schutter, and H. Hellendoorn, "Model predictive control for optimal coordination of ramp metering and variable speed limits," *Transportation Research Part C*, vol. 13, no. 3, pp. 185–209, Jun. 2005.