Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Stability of Consensus Extended Kalman Filtering for Distributed State Estimation

Giorgio Battistelli * **Luigi Chisci** *

* *Università di Firenze, Dipartimento di Ingegneria dell'Informazione (DINFO), Via di Santa Marta 3, 50139 Firenze, Italy (e-mail: giorgio.battistelli@unifi.it , luigi.chisci@unifi.it).*

Abstract: The paper addresses networked estimation of the state of a nonlinear dynamical system. It is shown how, exploiting a suitable consensus approach wherein prior and novel information are dealt with in a separate way along with the extended Kalman filter linearization paradigm, the resulting distributed nonlinear filter guarantees local stability under minimal requirements of network connectivity and system collective observability. A simulation case-study concerning target tracking with a network of nonlinear (angle and range) position sensors is worked out in order to show the effectiveness of the considered consensus filter.

Keywords: Networked systems; distributed state estimation; sensor fusion; nonlinear filters; stability analysis; consensus; Kalman filters.

1. INTRODUCTION

Consensus is a widely exploited tool for distributing computations over networks in a scalable way. An especially important application of consensus is networked state estimation given measurements provided by a wireless sensor network. The literature on the subject is quite vast and the interested reader is referred to [4, 5, 6, 10, 11, 12, 14, 15] and references therein for an overview of the different existing approaches. Recent work [1, 2] has introduced a linear consensus Kalman filter with guaranteed stability under minimal requirements of network connectivity and system collective observability, i.e. observability from the whole network but not necessarily from individual sensors. Such a networked filter carries out in each node an update of information with the local measurements followed by consensus on the posterior information and then prediction. This approach, named *consensus on information* (CI), is, however, somewhat conservative due to an unnecessary discount of novel information, undergoing consensus combined with the prior information. With this respect, it has been shown how performance can be improved by weighting differently the novel and the prior information in the consensus step [3, 9]. In particular, for the resulting consensus filter, [3] proves that stability in the linear system case still holds under network connectivity and collective observability.

Following the guidelines of [3], the consensus filters of [1, 2, 3, 9] can be readily extended to a nonlinear setting by exploiting the Extended Kalman Filter (EKF) linearization argument. The present paper provides a further contribution by proving that such a family of consensus filters guarantees local stability, under the same connectivity and collective observability assumptions, in such a more general nonlinear system setting. Thanks to this result, the proposed consensus EKF emerges as an effective tool for the solution of many practically relevant distributed nonlinear filtering problems like, e.g., distributed tracking of a moving object given measurements of angle, range and/or Doppler wireless communicating sensors spread over the area of interest; such sensors, in fact, are highly nonlinear and unable to individually guarantee observability.

The rest of the paper is organised as follows. Section 2 introduces the problem setting. Section 3 reviews the considered family of consensus EKF algorithms for distributed state estimation and Section 4 analyses their stability property. Section 5 demonstrates, via simulation experiments, the effectiveness of such a consensus filter in a nonlinear target tracking case-study. Section 6 ends the paper with concluding remarks. All the proofs are omitted due to space constraints.

2. PROBLEM SETTING

This paper addresses Distributed State Estimation (DSE) over a sensor network consisting of two types of nodes: *communication* nodes have only *processing* and *communication* capabilities, i.e. they can process local data as well as exchange data with neighboring nodes, while *sensor* nodes have also *sensing* capabilities, i.e. they can sense data from the environment. Notice that communication nodes are introduced to act as "relays" of information whenever sensor nodes are too far away to communicate. In the sequel, the network will be denoted by the triplet $(\mathcal{S}, \mathcal{C}, \mathcal{A})$ where: $\mathcal{S}$ is the set of sensor nodes, $\mathcal{C}$ the set of communication nodes, $\mathcal{N} = \mathcal{S} \bigcup \mathcal{C}$, $\mathcal{A} \subseteq \mathcal{N} \times \mathcal{N}$ is the set of arcs (connections) such that $(i, j) \in \mathcal{A}$ if node j can receive data from node i (clearly $(i, i) \in \mathcal{A}$ for all $i \in \mathcal{N}$). Further, for each node $i \in \mathcal{N}$, $\mathcal{N}^i$ will denote the set of its in-neighbors (including i itself), i.e. $\mathcal{N}^i \triangleq \{j : (j, i) \in \mathcal{A}\}$.

The DSE problem over the sensor network $(\mathcal{S}, \mathcal{C}, \mathcal{A})$ can be formulated as follows. Consider a dynamical system

$$\mathbf{x}_{t+1} = \mathbf{f}(\mathbf{x}_t) + \mathbf{w}_t \tag{1}$$

and a set of sensors $\mathcal{S}$ with measurement equations

$$\mathbf{y}_t^i = \mathbf{h}^i(\mathbf{x}_t) + \mathbf{v}_t^i, \qquad i \in \mathcal{S}. \tag{2}$$

Table 1. Information CEKF Algorithm, to be implemented at each sampling interval $t = 1, 2, \ldots$ starting from initial conditions $\hat{\mathbf{x}}_{1|0}$, $\mathbf{\Omega}_{1|0}$, $\mathbf{q}_{1|0} = \mathbf{\Omega}_{1|0}\, \hat{\mathbf{x}}_{1|0}$.

Correction (measurement-update):

$$\mathbf{C}_t^i = \frac{\partial \mathbf{h}^i}{\partial \mathbf{x}_t}\left(\hat{\mathbf{x}}_{t|t-1}\right), \qquad i \in \mathcal{S}$$

$$\mathbf{\Omega}_{t|t} = \mathbf{\Omega}_{t|t-1} + \sum_{i\in\mathcal{S}} \left(\mathbf{C}_t^i\right)^\top \mathbf{V}^i\, \mathbf{C}_t^i$$

$$\overline{\mathbf{y}}_t^i = \mathbf{y}_t^i - \mathbf{h}^i\left(\hat{\mathbf{x}}_{t|t-1}\right) + \mathbf{C}_t^i \hat{\mathbf{x}}_{t|t-1}, \quad i \in \mathcal{S}$$

$$\mathbf{q}_{t|t} = \mathbf{q}_{t|t-1} + \sum_{i\in\mathcal{S}} \left(\mathbf{C}_t^i\right)^\top \mathbf{V}^i\, \overline{\mathbf{y}}_t^i$$

Prediction (time-update):

$$\hat{\mathbf{x}}_{t|t} = \mathbf{\Omega}_{t|t}^{-1}\mathbf{q}_{t|t}, \quad \text{and} \quad \mathbf{A}_t = \frac{\partial \mathbf{f}}{\partial \mathbf{x}_t}\left(\hat{\mathbf{x}}_{t|t}\right)$$

$$\mathbf{\Omega}_{t+1|t} = \mathbf{W} - \mathbf{W}\mathbf{A}_t\left(\mathbf{\Omega}_{t|t} + \mathbf{A}_t^\top \mathbf{W}\mathbf{A}_t\right)^{-1}\mathbf{A}_t^\top \mathbf{W}$$

$$\hat{\mathbf{x}}_{t+1|t} = \mathbf{f}\left(\hat{\mathbf{x}}_{t|t}\right), \quad \text{and} \quad \mathbf{q}_{t+1|t} = \mathbf{\Omega}_{t+1|t}\, \hat{\mathbf{x}}_{t+1|t}$$

Then the objective is to have, at each time $t \in \{1, 2, \ldots\}$ and in each node $i \in \mathcal{N}$, an estimate $\hat{\mathbf{x}}_{t|t}^i$ of the state $\mathbf{x}_t$ constructed only on the basis of the local measurements (when available) and of data received from all adjacent nodes $j \in \mathcal{N}^i \backslash \{i\}$.

2.1 Centralized Extended Kalman Filter

Before describing the proposed DSE algorithm, it is convenient to briefly recall the equations of the centralized Extended Kalman Filter, which is assumed to simultaneously process all measurements $\{y_k^i, i \in \mathcal{S}\}$. Hereafter, for convenience, the *information filter* form will be adopted. The information filter propagates, instead of the estimate $\hat{\mathbf{x}}_{t|t-1}$ and covariance $\mathbf{P}_{t|t-1}$, the *information* (inverse covariance) matrices

$$\mathbf{\Omega}_{t|t-1} \stackrel{\triangle}{=} \mathbf{P}_{t|t-1}^{-1}, \qquad \mathbf{\Omega}_{t|t} \stackrel{\triangle}{=} \mathbf{P}_{t|t}^{-1}$$

and the vectors

$$\mathbf{q}_{t|t-1} \stackrel{\triangle}{=} \mathbf{P}_{t|t-1}^{-1}\hat{\mathbf{x}}_{t|t-1}, \qquad \mathbf{q}_{t|t} \stackrel{\triangle}{=} \mathbf{P}_{t|t}^{-1}\hat{\mathbf{x}}_{t|t}$$

that will be referred to as *information vectors*. Then, the recursive information filter of Table 1 can be derived [3], where $\mathbf{W}$ and $\mathbf{V}^i$, $i \in \mathcal{N}$ are given positive definite matrices. A typical choice for such matrices is to take $\mathbf{W}$ as an estimate of the inverse covariance of the process disturbance $\mathbf{w}_t$, and each $\mathbf{V}^i$ as an estimate of the inverse covariance of the measurement noise $\mathbf{v}_t^i$ affecting the i-th sensor. Notice, however, that a specific choice of such matrices is immaterial for the subsequent developments.

The algorithm of Table 1 generalizes the *Information Kalman Filter* algorithm, corresponding to $\mathbf{f}(\mathbf{x}) = \mathbf{A}_t\mathbf{x}$ and $\mathbf{h}^i(\mathbf{x}) = \mathbf{C}_t^i\mathbf{x}$, to nonlinear systems (1) and/or sensors (2) via the Extended Kalman Filter paradigm of linearizing the state and measurement equations around the current estimate. With this respect, the following assumption is needed.

A1. The functions $\mathbf{f}$ and $\mathbf{h}^i$, $i \in \mathcal{S}$, are twice continuously differentiable on $\mathbb{R}^n$, where $n = \dim(\mathbf{x})$.

Notice that, in order to streamline the presentation, here and in the following it is supposed that the functions $\mathbf{f}$ and $\mathbf{h}^i$, $i \in \mathcal{S}$, are defined over the whole $\mathbb{R}^n$. However, all the results presented hereafter could be suitably modified to account for the case when the system trajectories are confined to a given set $\mathcal{X} \subset \mathbb{R}^n$.

3. DISTRIBUTED EXTENDED KALMAN FILTER

The focus of this paper is on the family of DSE algorithms proposed in [3] and based on the idea of combining *consensus on information (CI)* and *consensus on measurements (CM)*, by performing at each time instant two parallel consensus iterations. A brief description of the resulting algorithm is described in the following. For a detailed derivation, the interested reader is referred to [3].

Let us assume that, at time t, each node $i \in \mathcal{N}$ be provided with a local information pair $\left(\mathbf{\Omega}_{t|t-1}^i, \mathbf{q}_{t|t-1}^i\right)$. Then, the CI spreads such information in the network by performing a given number L of consensus steps of the type

$$\begin{aligned} \mathbf{q}_{t|t}^i(\ell+1) &= \sum_{j\in\mathcal{N}^i} \pi^{i,j}\; \mathbf{q}_{t|t}^j(\ell) \\ \mathbf{\Omega}_{t|t}^i(\ell+1) &= \sum_{j\in\mathcal{N}^i} \pi^{i,j}\; \mathbf{\Omega}_{t|t}^j(\ell) \end{aligned} \tag{3}$$

for $\ell = 0, 1, \ldots, L-1$ with the initialization $\mathbf{q}_{t|t}^i(0) = \mathbf{q}_{t|t-1}^i$ and $\mathbf{\Omega}_{t|t}^i(0) = \mathbf{\Omega}_{t|t-1}^i$. Notice that, in each consensus iteration, each node i computes a regional average, that is a combination of the values in $\mathcal{N}^i$ with suitable consensus weights $\pi^{i,j}$, $j \in \mathcal{N}^i$. In this paper, a convex combination is adopted by supposing $\pi^{i,j} \geq 0$ and $\sum_{j\in\mathcal{N}^i}\pi^{i,j} = 1, \forall i \in \mathcal{N}$.

As for CM, the idea is to exploit consensus in order to compute in a distributed way the quantities $\partial\mathbf{\Omega}_t \stackrel{\triangle}{=} \sum_{i\in\mathcal{S}} \left(\mathbf{C}_t^i\right)^\top \mathbf{V}_t^i\mathbf{C}_t^i$ and $\partial\mathbf{q}_t \stackrel{\triangle}{=} \sum_{i\in\mathcal{S}} \left(\mathbf{C}_t^i\right)^\top \mathbf{V}_t^i\overline{\mathbf{y}}_t^i$. To this end, at each time t, L consensus steps of the type

$$\begin{aligned} \delta\mathbf{q}_t^i(\ell+1) &= \sum_{j\in\mathcal{N}^i} \pi^{i,j}\; \delta\mathbf{q}_t^j(\ell) \\ \delta\mathbf{\Omega}_t^i(\ell+1) &= \sum_{j\in\mathcal{N}^i} \pi^{i,j}\; \delta\mathbf{\Omega}_t^j(\ell) \end{aligned} \tag{4}$$

are performed , where $\ell = 0, 1, \ldots, L-1$. For each sensor node $i \in \mathcal{S}$, the initial vector $\delta\mathbf{q}_t^i(0)$ and the initial matrix $\delta\mathbf{\Omega}_t^i(0)$ are set equal to $\left(\mathbf{C}_t^i\right)^\top \mathbf{V}_t^i\overline{\mathbf{y}}_t^i$ and $\left(\mathbf{C}_t^i\right)^\top \mathbf{V}_t^i\mathbf{C}_t^i$, respectively. For each communication node $i \in \mathcal{C}$, we simply set $\delta\mathbf{q}_t^i(0) = 0$ and $\delta\mathbf{\Omega}_t^i(0) = 0$.

The consensus iteration (4) has been originally proposed in [12, 10] in a linear setting. By following the EKF paradigm, the algorithm can be readily extended to nonlinear systems. The only difference is that the linearized output matrices $\mathbf{C}_t^i$, and hence the virtual measurements $\overline{\mathbf{y}}_t^i$, have to be redefined in terms of the local state predictions $\hat{\mathbf{x}}_{t|t-1}^i$ instead of the centralized one $\hat{\mathbf{x}}_{t|t-1}$, which is not available in a distributed setting (see [3]).

Notice that consensus provides, at convergence, the averages $\partial\mathbf{\Omega}_t/|\mathcal{N}|$ and $\partial\mathbf{q}_t/|\mathcal{N}|$, $|\mathcal{N}|$ denoting cardinality of $\mathcal{N}$, while the information filter update actually requires $\partial\mathbf{\Omega}_t$ and $\partial\mathbf{q}_t$. This drawback can be partially remedied by multiplying the consensus outcome by some suitable scalar weight γ_t^i. Combining CI and CM, the consensus-

Table 2. Hybrid CMCI (HCMCI) Algorithm

Compute the local correction terms
if $i \in \mathcal{S}$ **then**
sample the measurement $\mathbf{y}_t^i$

$$\mathbf{C}_t^i = \frac{\partial \mathbf{h}^i}{\partial \mathbf{x}}\left(\hat{\mathbf{x}}_{t|t-1}^i\right)$$
$$\overline{\mathbf{y}}_t^i = \mathbf{y}_t^i - \mathbf{h}^i\left(\hat{\mathbf{x}}_{t|t-1}^i\right) + \mathbf{C}_t^i \hat{\mathbf{x}}_{t|t-1}^i$$
$$\delta\mathbf{q}_t^i = \left(\mathbf{C}_t^i\right)^\top \mathbf{V}_t^i \overline{\mathbf{y}}_t^i$$
$$\delta\mathbf{\Omega}_t^i = \left(\mathbf{C}_t^i\right)^\top \mathbf{V}^i \mathbf{C}_t^i$$

else
$\delta\mathbf{q}_t^i = \mathbf{0}$, and $\delta\mathbf{\Omega}_t^i = \mathbf{0}$
end if
Consensus:
$\delta\mathbf{q}_t^i(0) = \delta\mathbf{q}_t^i$, $\delta\mathbf{\Omega}_t^i(0) = \delta\mathbf{\Omega}_t^i$,
$\mathbf{q}_{t|t-1}^i(0) = \mathbf{q}_{t|t-1}^i$, $\mathbf{\Omega}_{t|t-1}^i(0) = \mathbf{\Omega}_{t|t-1}^i$,
for $\ell = 0, 1, \ldots, L-1$ **do**
Fuse the quantities $\delta\mathbf{q}_t^j(\ell)$ and $\delta\mathbf{\Omega}_t^j(\ell)$ as in (4)
and the quantities $\mathbf{q}_{t|t-1}^j(\ell)$ and $\mathbf{\Omega}_{t|t-1}^j(\ell)$ as in (3)
end for
Correction:

$$\mathbf{q}_{t|t}^i = \mathbf{q}_{t|t-1}^i(L) + \gamma_t^i\, \delta\mathbf{q}_t^i(L)$$
$$\mathbf{\Omega}_{t|t}^i = \mathbf{\Omega}_{t|t-1}^i(L) + \gamma_t^i\, \delta\mathbf{\Omega}_t^i(L)$$
$$\hat{\mathbf{x}}_{t|t}^i = \left(\mathbf{\Omega}_{t|t}^i\right)^{-1} \mathbf{q}_{t|t}^i$$

Prediction:

$$\hat{\mathbf{x}}_{t+1|t}^i = \mathbf{f}\left(\hat{\mathbf{x}}_{t|t}^i\right), \quad \text{and} \quad \mathbf{A}_t^i = \frac{\partial \mathbf{f}}{\partial \mathbf{x}_t}\left(\hat{\mathbf{x}}_{t|t}^i\right)$$
$$\mathbf{\Omega}_{t+1|t}^i = \mathbf{W} - \mathbf{W}\mathbf{A}_t^i\left(\mathbf{\Omega}_{t|t}^i + (\mathbf{A}_t^i)^\top \mathbf{W}\mathbf{A}_t^i\right)^{-1}(\mathbf{A}_t^i)^\top \mathbf{W}$$
$$\mathbf{q}_{t+1|t}^i = \mathbf{\Omega}_{t+1|t}^i \hat{\mathbf{x}}_{t+1|t}^i$$

based DSE algorithm of Table 2 is obtained. Hereafter, we will refer to this algorithm as *Hybrid CMCI* (HCMCI).

Actually, Table 2 provides a family of distributed filters corresponding to different choices of the scalar weigths γ_t^i. For example, when $\gamma_t^i = 1$, the CI filter of [1, 2] is retrieved. Further, in this case, it is possible to perform jointly the two parallel consensus algorithms of Table 2 so as to save bandwidth (this is true whenever the weights γ_t^i are node-independent). Another possible choice is for example $\gamma_t^i = |\mathcal{N}|$, which has the appealing feature of giving rise to a distributed algorithm converging to the centralized one as L tends to infinity. Notice that, when such a choice is adopted, the HCMCI filter coincides with the information weighted consensus of [9]. While asymptotically optimal, such a choice need not be the best one when only a finite number of consensus steps is performed. In fact, in this case, a multiplication by $|\mathcal{N}|$ could actually lead in some nodes to an overestimate of $\partial\mathbf{\Omega}_t$, a situation that one might want to avoid in order to preserve the consistency of each local filter [1]. For a discussion on this issue as well as an alternative choice for the weights γ_t^i, the interested reader is referred to [3]. For the purposes of this paper, it is just sufficient to make the following assumption.

A2. There exist two positive scalars $\underline{\gamma}$ and $\bar{\gamma}$ such that $0 < \underline{\gamma} \le \gamma_t^i \le \bar{\gamma}$, for any $i \in \mathcal{N}$ and $t \ge 0$.

[1] Recall that a filter is said to be *consistent* when its estimated error covariance is an upper bound (in the positive definite sense) of the true error covariance [7].

4. STABILITY ANALYSIS

In this section, the stability properties of the HCMCI DSE algorithm of Section 3 are analyzed. To this end, notice first that under assumption A1 the function $\mathbf{f}$ can be expanded as

$$\mathbf{f}(\mathbf{x}_t) - \mathbf{f}(\hat{\mathbf{x}}_{t|t}^i) = \mathbf{A}_t^i(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t}^i) + \boldsymbol{\varphi}(\mathbf{x}_t, \hat{\mathbf{x}}_{t|t}^i) \tag{5}$$

with $\mathbf{A}_t^i$ as in the HCMCI algorithm and $\boldsymbol{\varphi}(\cdot)$ a suitable continuous function going to zero as $\hat{\mathbf{x}}_{t|t}$ tends to $\mathbf{x}_t$. Similarly, each function $\mathbf{h}^i$, $i \in \mathcal{S}$, can be expanded as

$$\mathbf{h}^i(\mathbf{x}_t) - \mathbf{h}^i(\hat{\mathbf{x}}_{t|t-1}^i) = \mathbf{C}_t^i(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}^i) + \boldsymbol{\chi}^i(\mathbf{x}_t, \hat{\mathbf{x}}_{t|t-1}^i) \tag{6}$$

with $\mathbf{C}_t^i$ as in the HCMCI algorithm and $\boldsymbol{\chi}^i(\cdot)$ a suitable continuous function going to zero as $\hat{\mathbf{x}}_{t|t-1}^i$ tends to $\mathbf{x}_t$. By exploiting such expansions, it is possible to write the estimation error dynamics so that the linearized part is separated from the nonlinear terms. To this end, let us denote by $\mathbf{\Pi}$ the consensus matrix, whose elements are the consensus weights $\pi^{i,j}$ for any $i, j \in \mathcal{N}$. Further, let $\pi_\ell^{i,j}$ be the (i,j)-th element of $\mathbf{\Pi}^\ell$, i.e. the ℓ-th power of the consensus matrix $\mathbf{\Pi}$. Then the following result holds.

Proposition 1. Let assumptions A1-A2 hold and let the HCMCI algorithm be initialized at time $t = 1$ with positive definite information matrices $\mathbf{\Omega}_{1|0}^i$. Then, for any i and any t, the matrices $\mathbf{\Omega}_{t|t}^i$ are invertible and the estimation errors $\mathbf{e}_t^i = \mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}^i$ obey the recursion

$$\mathbf{e}_{t+1}^i = \sum_{j \in \mathcal{N}} \mathbf{\Phi}_t^{i,j} \mathbf{e}_t^j + \mathbf{r}_t^i + \mathbf{s}_t^i \tag{7}$$

where

$$\begin{aligned}
\mathbf{\Phi}_t^{i,j} &= \pi_L^{i,j}\, \mathbf{A}_t^i \left(\mathbf{\Omega}_{t|t}^i\right)^{-1} \mathbf{\Omega}_{t|t-1}^j\,, \\
\mathbf{r}_t^i &= \boldsymbol{\varphi}(\mathbf{x}_t, \hat{\mathbf{x}}_{t|t}^i) \\
&\quad + \sum_{j \in \mathcal{S}} \pi_L^{i,j} \gamma_t^i \mathbf{A}_t^i \left(\mathbf{\Omega}_{t|t}^i\right)^{-1} (\mathbf{C}_t^j)^\top \mathbf{V}^j \boldsymbol{\chi}^j(\mathbf{x}_t, \hat{\mathbf{x}}_{t|t-1}^j)\,, \\
\mathbf{s}_t^i &= \mathbf{w}_t - \sum_{j \in \mathcal{S}} \pi_L^{i,j} \gamma_t^i \mathbf{A}_t^i \left(\mathbf{\Omega}_{t|t}^i\right)^{-1} (\mathbf{C}_t^j)^\top \mathbf{V}^j \mathbf{v}_t^j\,.
\end{aligned}$$

□

In order to study the stability of the estimation error dynamics (7), the following assumption on the consensus weights is needed.

A3. The consensus matrix $\mathbf{\Pi}$ is row stochastic and primitive [2].

Notice that assumption A3 can always be satisfied provided that the network is connected. For instance, in this case, the Metropolis weights [16, 4] satisfy A3. While taking the consensus matrix $\mathbf{\Pi}$ row stochastic is sufficient for stability, a doubly stochastic $\mathbf{\Pi}$ would also ensure that all the elements of $\mathbf{\Pi}^L$ tends to $1/|\mathcal{N}|$ as $L \to +\infty$.

Let now $\mathbf{p}$ be the Perron-Frobenius left eigenvector of the matrix $\mathbf{\Pi}^L$ and let p^i denote its i-th component. Further, consider the candidate Lyapunov function

[2] Recall that a non-negative square matrix $\mathbf{\Pi}$ is row stochastic if all its rows sum up to 1. Further, it is primitive if there exists an integer m such that all the elements of $\mathbf{\Pi}^m$ are strictly positive.

$$\mathscr{V}_t(\mathbf{e}_t) = \sum_{i\in\mathcal{N}} p^i \left(\mathbf{e}_t^i\right)^\top \mathbf{\Omega}_{t|t-1}^i \mathbf{e}_t^i$$

for the overall estimation error dynamics, where $\mathbf{e}_t = \mathrm{col}\left(\mathbf{e}_t^i,\, i \in \mathcal{N}\right)$. Notice that, by virtue of assumption A3, the eigenvector $\mathbf{p}$ has strictly positive components $p^i, i \in \mathcal{N}$, and satisfies the equation $\mathbf{p}^\top \mathbf{\Pi}^L = \mathbf{p}^\top$, i.e., $\sum_{j\in\mathcal{N}} p^j \pi_L^{j,i} = p^i$.

The following result concerning the linearized part of the error dynamics can now be stated.

Lemma 1. Let assumptions A1-A3 be satisfied. Further, suppose that the following conditions hold:

i) there exist nonnegative reals $\bar{a}$ and $\bar{c}$ such that
$$\|\mathbf{A}_t^i\| \le \bar{a}, \quad \|\mathbf{C}_t^i\| \le \bar{c}$$
for any i and any t;

ii) there exist postive reals $\underline{\omega}$, $\bar{\omega}$ such that
$$0 < \underline{\omega}\mathbf{I} \le \mathbf{\Omega}_{t|t}^i \le \bar{\omega}\mathbf{I}$$
for any i and any t;

iii) the matrix $\mathbf{A}_t^i$ is invertible for any i and any t.

Then, there exists a nonegative scalar $\tilde{\beta} < 1$ such that, for any t,
$$\mathscr{V}_{t+1}(\mathbf{\Phi}_t \mathbf{e}_t) \le \tilde{\beta}\, \mathscr{V}_t(\mathbf{e}_t)$$
where $\mathbf{\Phi}_t$ is the block matrix whose block elements are given by the matrices $\mathbf{\Phi}_t^{i,j}$ defined in Proposition 1. □

In the spirit of the classical results on stability of the CEKF [13], one can see i)-iii) as conditions that can be verified on line during the state estimation process in order to assess its reliability. Of course such conditions can also be related to specific properties of system (1)-(2). For instance, condition i) automatically holds when the functions $\mathbf{f}$ and $\mathbf{h}^i$, $i \in \mathcal{S}$, are globally Lipschitz or, in view of assumption A1, when the estimated trajectories $\hat{\mathbf{x}}_{k|k}^i$, $i \in \mathcal{N}$, are bounded. Further, condition ii) is closely related to the collective observability of the state $\mathbf{x}_t$ from the measurements $\mathbf{y}_t^i$, $i \in \mathcal{S}$, collected by all the available sensors. This issue will be discussed in some detail in Section 4.1.

Let us now turn back our attention to the overall estimation error dynamics by noting that the functions $\boldsymbol{\varphi}$ and $\boldsymbol{\chi}^i$ in (5) and (6) represent the remainders of the Taylor expansion of $\mathbf{f}$ and, respectively, $\mathbf{h}^i$ and hence, under suitable assumptions, go to zero with order of convergence greater than 1. With this respect, in the lines of [13], the following assumption is made.

A4. There exist positive reals $\epsilon_\varphi, \kappa_\varphi, \epsilon_{\chi^i}, \kappa_{\chi^i}$, $i \in \mathcal{S}$, such that the nonlinear functions $\boldsymbol{\varphi}$ and $\boldsymbol{\chi}^i$ in (5) and (6), respectively, are bounded as
$$\|\boldsymbol{\varphi}(\mathbf{x}, \hat{\mathbf{x}})\| \le \kappa_\varphi \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \tag{8}$$
$$\|\boldsymbol{\chi}^i(\mathbf{x}, \hat{\mathbf{x}})\| \le \kappa_{\chi^i} \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \tag{9}$$
for any pair $\mathbf{x}, \hat{\mathbf{x}} \in \mathbb{R}^n$ with $\|\mathbf{x} - \hat{\mathbf{x}}\| \le \epsilon_\varphi$ and $\|\mathbf{x} - \hat{\mathbf{x}}\| \le \epsilon_{\chi^i}$, respectively.

By exploiting Lemma 1 and assumption A4 the following local stability result can be derived.

Theorem 1. Let assumptions A1-A4 be satisfied. Further, suppose that conditions i)-iii) of Lemma 1 hold. Then, the estimation error $\mathbf{e}_t^i$ turns out to be bounded in all the network nodes, i.e., there exists a positive real ϵ such that
$$\limsup_{t\to\infty} \|\mathbf{e}_t^i\| \le \epsilon \tag{10}$$
for any i, provided that the initial estimation errors satisfy
$$\|\mathbf{e}_1^i\| \le \epsilon_0 \tag{11}$$
for some suitable constant $\epsilon_0 > 0$ and the disturbances satisfy
$$\|\mathbf{w}_t\| \le \epsilon_w, \quad \|\mathbf{v}_t^i\| \le \epsilon_{v^i},\ i \in \mathcal{S} \tag{12}$$
for some suitable constants $\epsilon_w > 0$ and $\epsilon_{v^i} > 0$, $i \in \mathcal{S}$. □

It is worth noting that, when the disturbance amplitudes $\epsilon_w > 0$ and $\epsilon_{v^i} > 0$ decrease, the asymptotic bound ϵ decreases as well and, in particular, the following corollary to Theorem 1 holds.

Corollary 1. Let the system dynamics (1) and the measurement equations (2) be noise-free, i.e.,
$$\mathbf{w}_t = 0, \quad \mathbf{v}_t^i = 0$$
for any i and any t. Then, under the same assumptions of Theorem 1, the estimation error goes to zero in all the network nodes, i.e.,
$$\lim_{t\to\infty} \|\mathbf{e}_t^i\| = 0$$
for any i, provided that the initial estimation errors satisfy
$$\|\mathbf{e}_1^i\| \le \epsilon_0$$
for some suitable constant $\epsilon_0 > 0$. □

4.1 Connection with collective observability

This section is devoted to discussing how conditions i)-iii) of Lemma 1 can be related to specific properties of system (1)-(2). To this end, let $\mathbf{h} = \mathrm{col}(\mathbf{h}^i,\, i \in \mathcal{S})$ be the collective output function, and let $\mathbf{F}^{[M]}(\mathbf{x})$ be the collective observability mapping defined over a time window of length M, i.e.

$$\mathbf{F}^{[M]}(\mathbf{x}) = \begin{bmatrix} \mathbf{h}(\mathbf{x}) \\ \mathbf{h} \circ \mathbf{f}(\mathbf{x}) \\ \vdots \\ \mathbf{h} \circ \underbrace{\mathbf{f} \circ \cdots \circ \mathbf{f}}_{M \text{ times}}(\mathbf{x}) \end{bmatrix}$$

where $\circ$ denotes composition. In words, given a time window $\{t-M, \ldots, t\}$, $\mathbf{F}^{[M]}(\mathbf{x})$ coincides with the mapping from the state $\mathbf{x}$ at time $t - M$ to the vector made up of the noise-free collective outputs at times $t - M, \ldots, t$.

Supposing that the system trajectory lies within some compact set $\mathcal{X}$, the following assumptions are now needed.

A5. For any $\mathbf{x} \in \mathcal{X}$, $\partial\mathbf{f}(\mathbf{x})/\partial\mathbf{x}$ is non-singular.

A6. There exist a positive integer M such that, for any $\mathbf{x} \in \mathcal{X}$, $\mathrm{rank}\left\{\partial\mathbf{F}^{[M]}(\mathbf{x})/\partial\mathbf{x}\right\} = n$ where $n = \dim(\mathbf{x})$.

Notice that assumption A5 amounts to requiring that the state transition function $\mathbf{f}(\mathbf{x})$ is a diffeomorphism on $\mathcal{X}$ and, hence, reversible. Further, as well known, assumption A6 ensures that *collective observability*, in the sense of the invertibility of the mapping $\mathbf{F}^{[M]}(\mathbf{x})$, holds.

The following result can now be stated.

Lemma 2. Let the system trajectory belong to $\mathcal{X}$, i.e., $\{\mathbf{x}_t\} \subset \mathcal{X}$, and suppose that assumptions A1-A6 are satisfied and that the HCMCI algorithm is initialized at time $t = 1$ with positive definite information matrices $\mathbf{\Omega}^i_{1|0}$. Then, conditions i)-iii) of Lemma 1 hold provided that, for any i and t, the estimation errors satisfy

$$\|\mathbf{e}^i_t\| \leq \hat{\epsilon} \tag{13}$$

for some suitable constant $\hat{\epsilon}$ and the disturbances satisfy

$$\|\mathbf{w}_t\| \leq \hat{\epsilon}_w, \quad \|\mathbf{v}^i_t\| \leq \hat{\epsilon}_{v^i},\ i \in \mathcal{S} \tag{14}$$

for some suitable constants $\hat{\epsilon}_w > 0$ and $\hat{\epsilon}_{v^i} > 0$, $i \in \mathcal{S}$. □

In view of Lemma 2 and Theorem 1, it is possible to prove the following stability result which summarizes all the foregoing derivations.

Theorem 2. Let the system trajectory belong to $\mathcal{X}$, i.e., $\{\mathbf{x}_t\} \subset \mathcal{X}$, and suppose that assumptions A1-A6 are satisfied and that the HCMCI algorithm is initialized at time $t = 1$ with positive definite information matrices $\mathbf{\Omega}^i_{1|0}$. Then, the estimation error $\mathbf{e}^i_t$ turns out to be bounded in all the network nodes, i.e., there exists a positive real $\tilde{\epsilon}$ such that

$$\|\mathbf{e}^i_t\| \leq \tilde{\epsilon} \tag{15}$$

for any i, provided that the initial estimation errors satisfy

$$\|\mathbf{e}^i_1\| \leq \tilde{\epsilon}_0 \tag{16}$$

for some suitable constant $\tilde{\epsilon}_0 > 0$ and the disturbances satisfy

$$\|\mathbf{w}_t\| \leq \tilde{\epsilon}_w, \quad \|\mathbf{v}^i_t\| \leq \tilde{\epsilon}_{v^i},\ i \in \mathcal{S} \tag{17}$$

for some suitable constants $\tilde{\epsilon}_w > 0$ and $\tilde{\epsilon}_{v^i} > 0$, $i \in \mathcal{S}$. □

5. SIMULATION RESULTS

The aim of this section is to corroborate the theoretical analysis by showing the effectiveness of the HCMCI algorithm in a target tracking case study. To this end, the target motion is modelled by a linear (nearly constant velocity) model

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{w}_t$$

with

$$\mathbf{A} = \begin{bmatrix} 1 & T_s & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T_s \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} \frac{T_s^3}{3} & \frac{T_s^2}{2} & 0 & 0 \\ \frac{T_s^2}{2} & T_s & 0 & 0 \\ 0 & 0 & \frac{T_s^3}{3} & \frac{T_s^2}{2} \\ 0 & 0 & \frac{T_s^2}{2} & T_s \end{bmatrix} q$$

where: $\mathbf{x}_t = [x_t, \dot{x}_t, y_t, \dot{y}_t]^\top$ is the kinematic target state at sampling time t made up of the Cartesian coordinates of position (x_t, y_t) and of velocity $(\dot{x}_t, \dot{y}_t)$; T_s is the sampling interval; q is the variance of the random fluctuations of target speed and $\mathbf{Q}$ the covariance matrix of the disturbance $\mathbf{w}_t$. Two different simulation scenarios corresponding to two different sensor networks will be considered.

The target position is measured by two types of nonlinear sensors measuring angle or, respectively, distance. These two sensors, from now on indicated by the acronyms DOA (Direction Of Arrival) and TOA (Time Of Arrival), are characterized by the following measurement functions:

$$\mathbf{h}^i(\mathbf{x}) = \begin{cases} \operatorname{atan2}\left(x - x^i, y - y^i\right), & \text{if } i \text{ is a DOA sensor} \\ \sqrt{\left(x - x^i\right)^2 + \left(y - y^i\right)^2}, & \text{if } i \text{ is a TOA sensor} \end{cases}$$

where atan2 is the 4-quadrant inverse tangent function and (x^i, y^i) denotes the position of the i-th sensor. Overall, the network consists of 100 communication nodes, 5 TOA sensor nodes, and 5 DOA sensor nodes. A graphical representation of the sensor network is provided in Fig. 1.

The measurement noise is assumed to have $\sigma_\theta = 2°$ standard deviation for DOA sensors, and $\sigma_r = 10\ m$ standard deviation for TOA sensors. Other parameters of the simulations are fixed to sampling interval $T_s = 1$ and $q = 0.5\ m^2/s^3$. The matrices $\mathbf{W}$ and $\mathbf{V^i}$ in the HCMCI algorithm are taken as the inverses of the disturbance and, respectively, measurement noise covariances. Finally, the consensus weights used in the simulations have been set equal to the Metropolis weights.

The HCMCI algorithm described in Section 3 has been compared with the CEKF of Section 2.1, with the CI DSE algorithm of [1, 2], and with the CM DSE algorithm (originally proposed by [12] in a linear setting). Notice that, in the considered setting, the CI algorithm corresponds to the HCMCI algorithm when the weights γ^i_t are set equal to 1 for any i and any t. Further, the CM algorithm follows the same steps as in Table 2 with the only difference than no consensus on the prior information $\left(\mathbf{\Omega}^i_{t|t-1}, \mathbf{q}^i_{t|t-1}\right)$ is performed. For the sake of comparison, 200 independent Monte Carlo trials have been performed and the position root mean square error (PRMSE) has been computed as performance index. Fig. 2 shows the comparison between the considered state estimation algorithms for $L = 1$ and $L = 9$ consensus steps. As it can be seen, both the HCMCI and the CI show a stable behavior already for one consensus step. On the other hand, the CM algorithm requires a minimum number of consensus steps per iteration in order to converge (in this case, a minimum of $L = 9$ consensus steps is required). On the contrary, the HCMCI algorithm provides satisfactory performance already for a low L with better estimation accuracy as the number of consensus steps grows. Similar conclusions can be drawn also in the nonlinear sensor network case as shown in Fig. 2. In addition, in this case, the HCMCI substantially outperforms the CI both for low and high numbers of consensus steps. With this respect, notice that simulations have been performed also for the values of L ranging from 2 to 8, with similar results. The plots are not reported here due to space constraints.

6. CONCLUSIONS

Distributed state estimation for nonlinear systems has been addressed. In particular, stability analysis for a family of consensus Extended Kalman Filter algorithms has been carried out. It has been proved that such algorithms guarantee local stability under network connectivity and collective system observability. An open problem that deserves further investigation is whether similar, or even stronger, stability properties can be achieved by means of different distributed nonlinear state estimation techniques, for example based on the Unscented Kalman Filter [8].

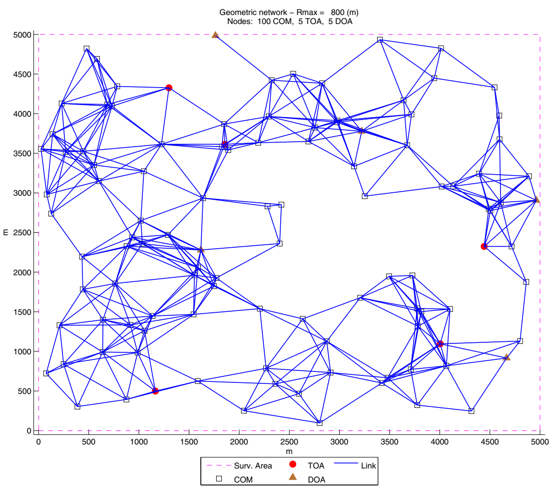

Fig. 1. Nonlinear sensor network used in the simulations.

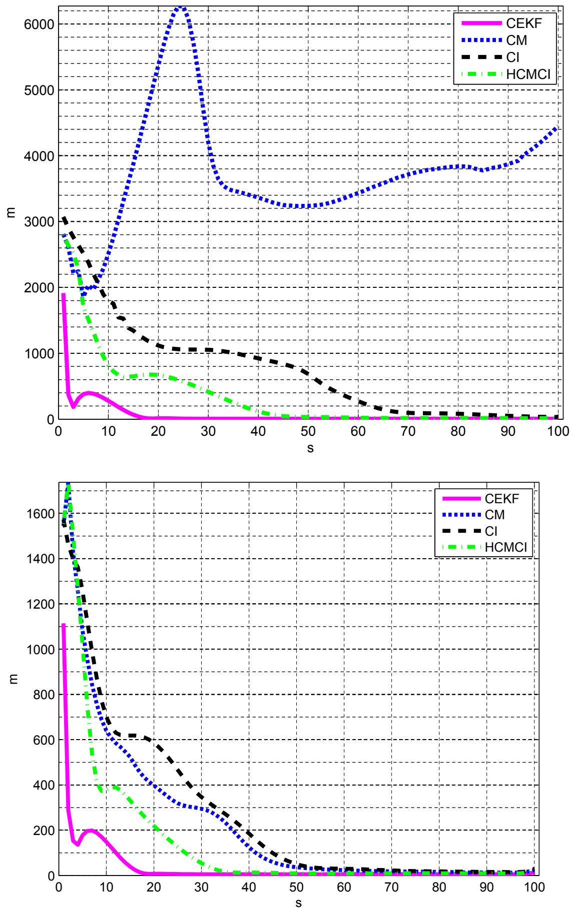

Fig. 2. PRMSE of the considered DSE algorithms for $L = 1$ (top), and $L = 9$ (bottom) consensus steps.

REFERENCES

[1] G. Battistelli and L. Chisci, "Kullback-Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability", *Automatica*, vol. 50, n. 3, pp. 707-718, 2014.

[2] G. Battistelli, L. Chisci, S. Morrocchi and F. Papi, "An information-theoretic approach to distributed state estimation", *Proc. 18th IFAC World Congress*, pp. 12477-12482, 2011.

[3] G. Battistelli, L. Chisci, G. Mugnai, A. Farina and A. Graziano, "Consensus-based algorithms for distributed filtering", *Proc. 51st IEEE Conference on Decision and Control*, pp. 794-799, 2012.

[4] G.C. Calafiore and F. Abrate, "Distributed linear estimation over sensor networks", *International Journal of Control*, vol. 82, no. 5, pp. 868-882, 2009.

[5] M. Farina, G. Ferrari-Trecate and R. Scattolini, "Distributed moving horizon estimation for linear constrained systems", *IEEE Transactions on Automatic Control*, vol. 55, no. 11, pp. 2462-2475, 2010.

[6] M. Farina, G. Ferrari-Trecate, and R. Scattolini, "Distributed moving horizon estimation for nonlinear constrained systems", *International Journal of Robust and Nonlinear Control*, vol. 22, no. 2, pp. 123-143, 2012.

[7] S.J. Julier and J.K. Uhlmann, "General decentralized data fusion with covariance intersection" in *Handbook of Data Fusion*, CRC Press, 2001.

[8] S.J. Julier and J.K. Uhlmann, "Unscented filtering and nonlinear estimation", *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401-422, 2004.

[9] A.T. Kamal, J.A. Farrell and A.K. Roy-Chowdhury, "Information weighted consensus filters and their application in distributed camera networks", *IEEE Transactions on Automatic Control*, vol. 58, no.12, pp. 3112-3125, 2013.

[10] M. Kamgarpour and C. Tomlin, "Convergence properties of a decentralized Kalman filter", *Proc. 47th IEEE Conference on Decision and Control*, pp. 3205–3210, 2008.

[11] W. Li and Y. Jia, "Distributed consensus filtering for discrete-time nonlinear systems with non-Gaussian noise", *Signal Processing*, vol. 92, no. 10, pp. 2464-2470, 2012.

[12] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks", *Proc. 46th IEEE Conference on Decision and Control*, pp. 5492-5498, 2007.

[13] K. Reif, S. Gunther, E. Yaz and R. Unbehauen, "Stochastic stability of the discrete-time extended Kalman filter", *IEEE Transactions on Automatic Control*, vol. 44, n. 4, pp. 714-728, 1999.

[14] S.S. Stankovic, M.S. Stankovic and D.M. Stipanovic, "Consensus based overlapping decentralized estimation with missing observations and communication faults", *Automatica*, vol. 45, n. 6, pp. 1397–1406, 2009.

[15] V. Ugrinovskii, "Distributed robust filtering with H_∞ consensus", *Automatica*, vol. 47, no. 1, pp. 1-13, 2011.

[16] L. Xiao, S. Boyd and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus", *Proc. 4th International Symposium on Information Processing in Sensor Networks*, pp. 63-70, 2005.