Consensus-based linear and nonlinear filtering

Consensus-based linear and nonlinear filtering

G. Battistelli, L. Chisci, *Senior Member, IEEE,* G. Mugnai, A. Farina *Fellow, IEEE* and A. Graziano,

***Abstract*—This note addresses Distributed State Estimation (DSE) over sensor networks. Two existing consensus approaches for DSE, i.e. *consensus on information (CI)* and *consensus on measurements (CM)*, are combined to provide a novel class of hybrid consensus filters (named Hybrid CMCI) which enjoy the complementary benefits of CM and CI. Novel theoretical results, limitedly to linear systems, on the guaranteed stability of the Hybrid CMCI filters under collective observability and network connectivity are proved. Finally, the effectiveness of the proposed class of consensus filters is evaluated on a target tracking case-study with both linear and nonlinear sensors.**

***Index Terms*—Distributed state estimation; sensor networks; consensus; nonlinear filtering.**

I. INTRODUCTION

Advances in wireless sensor networks are making multi-agent (distributed) architectures ubiquitous in modern monitoring and control systems. This technological breakthrough, however, needs to be supported by theoretical research work finalized to redesign distributed estimation and control algorithms that preserve as much as possible the stability, performance and robustness requirements of their centralized counterparts. The present note specifically deals with *Distributed State Estimation (DSE)* over a sensor network with no fusion centre, by exploiting the *consensus* approach [1]. This topic has recently gathered a lot of attention in the context of both parameter and state estimation. For what concerns parameter estimation, consensus-based techniques with guaranteed performance can be found for example in [2], [3].

As for state estimation, several consensus-based approaches have been proposed in the literature both in linear and nonlinear settings. A first family of techniques, hereafter referred to as *Consensus on Estimates* (CE), is based on the idea of spreading the available information over the network by performing, at each time instant, a consensus averaging of the local state estimates/predictions [4]-[11]. The choice of averaging out only the state estimates serves the purpose of keeping the information exchange between neighboring nodes as limited as possible. However, since also the covariance matrices contain valuable information which can be used to improve performance, other, more elaborated approaches have been proposed. For instance, in [12]-[14], it is proposed to perform a consensus among the local measurements and innovation covariances so as to approximate, in a distributed way, the correction step of the centralized Kalman filter. This approach is referred to in this note as *Consensus on Measurements* (CM). A different point of view has been adopted in [15], giving rise to the so-called *Consensus on Information* (CI) approach. From an algorithmic viewpoint, in the linear case CI is nothing but consensus applied to the inverse covariance (information) matrix and to the information vector. From an information-theoretic viewpoint, CI can be interpreted as consensus on *probability density functions* in the *Kullback-Leibler average* sense.

G. Battistelli, L. Chisci and G. Mugnai are with Dipartimento di Ingegneria dell'Informazione, Università di Firenze, 50139, Firenze, Italy. e-mail: {giorgio.battistelli,luigi.chisci,giovanni.mugnai}@unifi.it

A. Farina and A. Graziano are with Selex ES, Rome, Italy. e-mail: {alfonso.farina,antonio.graziano}@selex-es.com

This work was supported in part by a grant of Selex ES.

The approach presented in this note moves from the observation that CM and CI exhibit complementary features. Specifically, CI guarantees stability for any number of consensus steps (even a single one [15]), but its mean-square estimation error performance can be hampered by the fact that the fusion rule adopts a conservative point of view by assuming the correlation between the estimates coming from different nodes to be completely unknown. On the other hand, CM avoids any conservative assumption on the correlation by fusing only the novel information, but it does not guarantee stability unless the number of consensus steps is sufficiently high. Clearly, this can be problematic whenever, for reduced communication cost and improved energy efficiency, only few consensus steps can be performed in each sampling interval. In this context, a novel family of DSE algorithms, named *Hybrid CMCI* (HCMCI), is proposed wherein CM and CI are combined so as to retain the positive features of both approaches. Different choices for the combination weights are discussed, leading to different properties. Following [16], where preliminary results on the subject can be found, the nonlinear case is treated by resorting to the Extended Kalman Filter (EKF) paradigm. The main contribution of this note is a stability analysis of the proposed family of HCMCI filters in the case of linear systems. In order to substantiate the analysis, a performance evaluation of the considered consensus-based filters on a target tracking simulation case-study with both linear and nonlinear sensors is presented. As a final remark, it is pointed out that, for a particular choice of the combination weights, the proposed algorithm is equivalent to the *Information Consensus Filter* (ICF), derived independently from our work and originally presented in [17], whose stability was observed experimentally in [18].

II. PROBLEM FORMULATION

This note addresses 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 among distant sensor nodes, in order to guarantee network connectivity while keeping the transmission power (and hence the communication range) of each sensor node at a moderate intensity, hence prolonging network lifetime. 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

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

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

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

It is assumed that $\mathbf{w}_t, \mathbf{v}_t^1, \mathbf{v}_t^2, \ldots$ are mutually uncorrelated zero-mean white noises with covariances $\mathbf{Q}_t = E\left[\mathbf{w}_t \mathbf{w}_t^T\right] > 0$ and

TABLE I: **Information C(E)KF Algorithm**, to be implemented at each sampling interval $t = 1, 2, \dots$ 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):

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

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

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

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

Prediction (time-update):

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

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

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

$\mathbf{R}_t^i = E\left[\mathbf{v}_t^i\left(\mathbf{v}_t^i\right)^T\right] > 0$. Then the objective is to have, at each time $t \in \{1, 2, \dots\}$ and in each node $i \in \mathcal{N}$, an estimate $\hat{\mathbf{x}}_{t|t}$ 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 $j \in \mathcal{N}^i \backslash \{i\}$.

Centralized (Extended) Kalman Filter

A testbed for the performance of any DSE algorithm is the centralized (Extended) Kalman Filter, denoted as C(E)KF, which is assumed to simultaneously process all measurements $\{\mathbf{y}_k^i, i \in \mathcal{S}\}$. 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 *information vectors*

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

Introducing the noise information matrices $\mathbf{W}_t \stackrel{\triangle}{=} \mathbf{Q}_t^{-1}$ and $\mathbf{V}_t^i \stackrel{\triangle}{=} \left(\mathbf{R}_t^i\right)^{-1}$ and supposing the functions $\mathbf{f}_t(\cdot)$ and $\mathbf{h}_t^i(\cdot)$ to be continuously differentiable for any t and any $i \in \mathcal{S}$, the recursive information filter of Table I can be derived.

Notice that the algorithm of Table I generalizes the *Information Kalman Filter* algorithm, corresponding to $\mathbf{f}_t(\mathbf{x}) = \mathbf{A}_t\mathbf{x}$ and $\mathbf{h}_t^i(\mathbf{x}) = \mathbf{C}_t^i\mathbf{x}$, 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.

III. Distributed State Estimation Algorithms

A. Consensus on Information and Consensus on Measurements

The *covariance intersection* fusion rule [19] suggests a possible consensus approach to DSE, namely consensus on the information (matrix-vector) pair. 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 *Consensus on Information (CI)* approach to DSE is summarized by the algorithm of Table II to be carried out at each sampling interval $t \geq 1$ in each node $i \in \mathcal{N}$. 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 note, a convex combination is adopted by supposing $\pi^{i,j} \geq 0$ and $\sum_{j\in\mathcal{N}^i}\pi^{i,j} = 1, \forall i \in \mathcal{N}$.

It is worth to point out that the CI algorithm of Table II reduces to the well known *covariance intersection* [19] for a single ($L = 1$) consensus step. It represents, therefore, a generalization of covariance intersection to multiple ($L > 1$) consensus steps which can be introduced in order to improve performance. Notice that L linearly increases both computation and communication burdens; hence it should be selected as a suitable tradeoff between cost and performance.

TABLE II: **Consensus on Information (CI) Algorithm**

Correction:

if $i \in \mathcal{S}$ **then**

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

$$\mathbf{\Omega}_t^i(0) = \mathbf{\Omega}_{t|t-1}^i + \left(\mathbf{C}_t^i\right)^T \mathbf{V}_t^i\, \mathbf{C}_t^i$$

sample the measurement $\mathbf{y}_t^i$

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

$$\mathbf{q}_t^i(0) = \mathbf{q}_{t|t-1}^i + \left(\mathbf{C}_t^i\right)^T \mathbf{V}_t^i\, \overline{\mathbf{y}}_t^i$$

end if

if $i \in \mathcal{C}$ **then**

$$\mathbf{\Omega}_t^i(0) = \mathbf{\Omega}_{t|t-1}^i, \quad \text{and} \quad \mathbf{q}_t^i(0) = \mathbf{q}_{t|t-1}^i$$

end if

Consensus:

for $\ell = 0, 1, \dots, L-1$ **do**

Fuse the quantities $\mathbf{q}_t^j(\ell)$ and $\mathbf{\Omega}_t^j(\ell)$ according to

$$\mathbf{q}_t^i(\ell+1) = \sum_{j\in\mathcal{N}^i} \pi^{i,j}\ \mathbf{q}_t^j(\ell)$$

$$\mathbf{\Omega}_t^i(\ell+1) = \sum_{j\in\mathcal{N}^i} \pi^{i,j}\ \mathbf{\Omega}_t^j(\ell)$$

end for

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

Prediction:

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

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

$$\mathbf{q}_{t+1|t}^i = \mathbf{\Omega}_{t+1|t}^i\hat{\mathbf{x}}_{t+1|t}^i$$

As an alternative approach, in [12], [13] it has been proposed to exploit consensus in order to compute in a distributed way the quantities $\partial\mathbf{\Omega}_t \stackrel{\triangle}{=} \sum_{i\in\mathcal{S}}\left(\mathbf{C}_t^i\right)^T\mathbf{V}_t^i\mathbf{C}_t^i$ and $\partial\mathbf{q}_t \stackrel{\triangle}{=} \sum_{i\in\mathcal{S}}\left(\mathbf{C}_t^i\right)^T\mathbf{V}_t^i\mathbf{y}_t^i$ for distributed linear Kalman filtering. As shown in [16], this approach, that will be referred to as *consensus on measurements (CM)*, can be extended to nonlinear systems by following the EKF paradigm. To this end, the information filter of Table I can be exploited, with the only difference that the virtual measurements $\overline{\mathbf{y}}_t^j$ has to be redefined in terms of the local state predictions $\hat{\mathbf{x}}_{t|t-1}^j$ instead of the centralized one $\hat{\mathbf{x}}_{t|t-1}$ (which is not available in a distributed setting). The interest reader is referred to [16] for a detailed derivation.

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 ω_t^i. Possible choices for the weights ω_t^i will be discussed in the next section. Summing up, the consensus-based DSE algorithm of Table III is obtained.

B. Hybrid Consensus

Consensus on measurements and consensus on information have complementary positive and negative features. To see this, let us denote by $\mathbf{\Pi}$ the consensus matrix, whose elements are the consensus weights $\pi^{i,j}$, $i,j \in \mathcal{N}$. 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 correction step for the considered algorithms can be rewritten as in Table IV. Notice that, for the terms $\delta\boldsymbol{q}_t^i$ and $\delta\mathbf{\Omega}_t^i$, the summations are extended only to sensor nodes since, for all comunication nodes, one has $\delta\mathbf{q}_t^j = 0$ and $\delta\mathbf{\Omega}_t^j = 0$.

TABLE III: **Consensus on Measurements (CM) Algorithm**

Compute the local correction terms:
if $i \in \mathcal{S}$ **then**
sample the measurement $\mathbf{y}_t^i$
$\mathbf{C}_t^i = \dfrac{\partial \mathbf{h}_t^i}{\partial \mathbf{x}} \left(\hat{\mathbf{x}}_{t|t-1}^i\right)$
$\overline{\mathbf{y}}_t^i = \mathbf{y}_t^i - \mathbf{h}_t^i \left(\hat{\mathbf{x}}_{t|t-1}^i\right) + \mathbf{C}_t^i \hat{\mathbf{x}}_{t|t-1}^i$
$\delta \mathbf{q}_t^i = \left(\mathbf{C}_t^i\right)^T \mathbf{V}_t^i \, \overline{\mathbf{y}}_t^i$
$\delta \mathbf{\Omega}_t^i = \left(\mathbf{C}_t^i\right)^T \mathbf{V}_t^i \, \mathbf{C}_t^i$
else
$\delta \mathbf{q}_t^i = \mathbf{0}$, and $\delta \mathbf{\Omega}_t^i = \mathbf{0}$
end if
Consensus:
$\delta \mathbf{q}_t^i(0) = \delta \mathbf{q}_t^i$, and $\delta \mathbf{\Omega}_t^i(0) = \delta \mathbf{\Omega}_t^i$
for $\ell = 0, 1, \ldots, L-1$ **do**
Fuse the quantities $\delta \mathbf{q}_t^j(\ell)$ and $\delta \mathbf{\Omega}_t^j(\ell)$ according to:
$\delta \mathbf{q}_t^i(\ell+1) = \sum_{j \in \mathcal{N}^i} \pi^{i,j} \, \delta \mathbf{q}_t^j(\ell)$
$\delta \mathbf{\Omega}_t^i(\ell+1) = \sum_{j \in \mathcal{N}^i} \pi^{i,j} \, \delta \mathbf{\Omega}_t^j(\ell)$
end for
Correction:
$\mathbf{q}_{t|t}^i = \mathbf{q}_{t|t-1}^i \; + \; \omega_t^i \, \delta \mathbf{q}_t^i(L)$
$\mathbf{\Omega}_{t|t}^i = \mathbf{\Omega}_{t|t-1}^i \; + \; \omega_t^i \delta \mathbf{\Omega}_t^i(L)$
$\hat{\mathbf{x}}_{t|t}^i = \left(\mathbf{\Omega}_{t|t}^i\right)^{-1} \mathbf{q}_{t|t}^i$
Prediction: as in table II

TABLE IV: Expression of the information vector and matrix in the correction step after consensus for the considered algorithms

	$\mathbf{q}_{t\|t}^i$	$\mathbf{\Omega}_{t\|t}^i$
CI	$\sum_{j \in \mathcal{N}} \pi_L^{i,j} \mathbf{q}_{t\|t-1}^j + \sum_{j \in \mathcal{S}} \pi_L^{i,j} \delta \mathbf{q}_t^j$	$\sum_{j \in \mathcal{N}} \pi_L^{i,j} \mathbf{\Omega}_{t\|t-1}^j + \sum_{j \in \mathcal{S}} \pi_L^{i,j} \delta \mathbf{\Omega}_t^j$
CM	$\mathbf{q}_{t\|t-1}^i + \omega_t^i \sum_{j \in \mathcal{S}} \pi_L^{i,j} \delta \mathbf{q}_t^j$	$\mathbf{\Omega}_{t\|t-1}^i + \omega_t^i \sum_{j \in \mathcal{S}} \pi_L^{i,j} \delta \mathbf{\Omega}_t^j$
HCMCI	$\sum_{j \in \mathcal{N}} \pi_L^{i,j} \mathbf{q}_{t\|t-1}^j + \omega_t^i \sum_{j \in \mathcal{S}} \pi_L^{i,j} \delta \mathbf{q}_t^j$	$\sum_{j \in \mathcal{N}} \pi_L^{i,j} \mathbf{\Omega}_{t\|t-1}^j + \omega_t^i \sum_{j \in \mathcal{S}} \pi_L^{i,j} \delta \mathbf{\Omega}_t^j$

As it can be seen from Table IV, in CI a convex combination of the local priors is computed. As shown in [15], this is important in order to ensure boundedness of the estimation error for any number (even a single one) of consensus steps. On the contrary, in CM no consensus on the local priors is performed since only the novel information, i.e. $\delta \mathbf{q}_t^j$ and $\delta \mathbf{\Omega}_t^j$, is combined. Hence, the stability of this family of distributed filters depends in a crucial way on the number L of consensus steps per sampling interval. To see this, let $\mathcal{N}^i(L)$ be the set of nodes from which node i can be reached in L hops. Then, since $\pi_L^{i,j}$ can be different from 0 only if $j \in \mathcal{N}^i(L)$, it is immediate to see that CM can provide a bounded estimation error on a given node i only if observability (or at least detectability) from the set of nodes $\mathcal{N}^i(L)$ holds. In general, this becomes true only for a sufficiently large number of consensus steps (see, for instance, the example of Section V).

On the other hand, in CI the novel information is unavoidably underweighted, whereas in CM the multiplication by the scalar weights ω_t^i can, at least partially, remedy this drawback, thus leading to enhanced performance as the number of consensus steps increases. In order to combine the benefits and neutralize the drawbacks of both CI and CM, the two approaches can be combined so as to give rise to a hybrid consensus approach, hereafter called HCMCI (*Hybrid Consensus on Measurements - Consensus on Information*). HCMCI actually carries out in parallel, at each consensus step, the two types of consensus CI and CM described above. The resulting algorithm is detailed in Table V.

TABLE V: **Hybrid CMCI (HCMCI) Algorithm**

Compute the local correction terms
if $i \in \mathcal{S}$ **then**
sample the measurement $\mathbf{y}_t^i$
$\delta \mathbf{q}_t^i = \left(\mathbf{C}_t^i\right)^T \mathbf{V}_t^i \, \mathbf{y}_t^i$
$\delta \mathbf{\Omega}_t^i = \left(\mathbf{C}_t^i\right)^T \mathbf{V}_t^i \, \mathbf{C}^i$.
else
$\delta \mathbf{q}_t^i = \mathbf{0}$, and $\delta \mathbf{\Omega}_t^i = \mathbf{0}$
end if
Consensus:
$\delta \mathbf{q}_t^i(0) = \delta \mathbf{q}_t^i$, $\delta \mathbf{\Omega}_t^i(0) = \delta \mathbf{\Omega}_t^i$,
$\mathbf{q}_t^i(0) = \mathbf{q}_{t|t-1}^i$, $\mathbf{\Omega}_t^i(0) = \mathbf{\Omega}_{t|t-1}^i$,
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 Table III
and in parallel fuse the quantities $\mathbf{q}_t^j(\ell)$ and $\mathbf{\Omega}_t^j(\ell)$ as in Table II
end for
Correction:
$\mathbf{q}_{t|t}^i = \mathbf{q}_{t|t-1}^i(L) \; + \; \omega_t^i \, \delta \mathbf{q}_t^i(L)$
$\mathbf{\Omega}_{t|t}^i = \mathbf{\Omega}_{t|t-1}^i(L) \; + \; \omega_t^i \, \delta \mathbf{\Omega}_t^i(L)$
$\hat{\mathbf{x}}_{t|t}^i = \left(\mathbf{\Omega}_{t|t}^i\right)^{-1} \mathbf{q}_{t|t}^i$
Prediction: as in table II

Notice that, actually, Table V provides a family of distributed filters corresponding to different choices of the scalar weigths ω_t^i. With this respect, a reasonable choice consists in setting $\omega_t^i = |\mathcal{N}|$. In fact, when the consensus weights are chosen so that $\pi_L^{i,j} \to 1/|\mathcal{N}|$ as $L \to \infty$, such a choice has the appealing feature of giving rise to a distributed algorithm converging to the centralized (E)KF as L tends to infinity. Notice that, when such a choice is adopted, the resulting algorithm is equivalent to the ICF proposed in [18] in a linear setting. Notice also that, in this case, it is possible to perform jointly the two parallel consensus algorithms of Table V so as to save bandwidth [18] (this is true whenever the weights ω_t^i are node-independent).

While asymptotically optimal, the choice $\omega_t^i = |\mathcal{N}|$ may have some drawbacks when only a finite, possibly small, number of consensus steps is performed. In fact, a multiplication by $|\mathcal{N}|$ could actually lead in some nodes to an overestimation of $\partial \mathbf{\Omega}_t$, a situation that one might want to avoid in order to preserve the consistency of each local filter[1]. An alternative solution is to exploit consensus so as to compute, in a distributed way, a normalization factor able to improve the filter performance while preserving consistency of each local filter. For example, an estimate of the fraction $|\mathcal{S}|/|\mathcal{N}|$ of sensor nodes in the network can be computed via the consensus algorithm

$$b_t^i(\ell+1) = \sum_{j \in \mathcal{N}^i} \pi^{i,j} \; b_t^j(\ell), \text{ for } \ell = 0, 1, \ldots, L-1 \qquad (3)$$

with the initialization $b_t^i(0) = 1$ if $i \in \mathcal{S}$, and $b_t^i(0) = 0$ otherwise. Then, it can be seen that the choice

$$\omega_t^i = \begin{cases} 1/b_t^i(L) & \text{if } b_t^i(L) \neq 0 \\ 1 & \text{otherwise} \end{cases} \qquad (4)$$

has the desirable property of preserving the consistency of each local filter. This can be seen by looking at the last row of Table IV and noting that the novel information is never overweighted. In fact, the terms $\delta \mathbf{q}_t^j$ and $\delta \mathbf{\Omega}_t^j$, for $i \in \mathcal{S}$, are multiplied by a weight

[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 [19].

$\omega_t^i \pi_L^{i,j} = \pi_L^{i,j}/b_t^i(L)$ which is guaranteed not to exceed 1 since, by construction, $b_t^i(L) = \sum_{j\in\mathcal{S}} \pi_L^{i,j}$.

IV. STABILITY ANALYSIS

In this section, the stability properties of the proposed DSE algorithm are analyzed in a linear time-invariant setting. To this end, it is supposed that the system dynamics and measurement equations take the form

$$\begin{aligned} \mathbf{x}_{t+1} &= \mathbf{A}\mathbf{x}_t + \mathbf{w}_t && (5)\\ \mathbf{y}_t^i &= \mathbf{C}^i\mathbf{x}_t + \mathbf{v}_t^i \qquad \text{if } i \in \mathcal{S} && (6) \end{aligned}$$

Further, the process and measurement noise covariances are supposed to be time-invariant as well, i.e., $\mathbf{Q}_t = \mathbf{Q} > 0$ and $\mathbf{R}_t^i = \mathbf{R}^i > 0$ for any t and any i. The following preliminary assumptions are needed.

- **A1.** The system matrix $\mathbf{A}$ is invertible.
- **A2.** The system is collectively observable, i.e. the pair $(\mathbf{A}, \mathbf{C})$ is observable where $\mathbf{C} := \mathrm{col}\left(\mathbf{C}^i;\, i \in \mathcal{S}\right)$.
- **A3.** The consensus matrix $\mathbf{\Pi}$ is row stochastic and primitive [2].
- **A4.** There exist two positive scalars $\underline{\omega}$ and $\bar{\omega}$ such that $0 < \underline{\omega} \le \omega_t^i \le \bar{\omega}$, for any $i \in \mathcal{N}$ and $t \ge 0$.

Notice that assumption A1 is automatically satisfied in sampled-data systems wherein the matrix $\mathbf{A}$ is obtained by discretization of a continuous-time system matrix. As for assumption A2, the collective observability requirement can be relaxed to collective detectability by resorting to an observability decomposition in each network node as discussed in [15]. Assumption A3 is strictly related to the network connectivity. In fact, it can be satisfied if and only if the network is strongly connected (i.e., any node is reachable from any other node through a directed path). For instance, in this case, the Metropolis weights [2], [3] 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 $L \to +\infty$. Finally, notice that both the choices suggested in the previous section for the weights ω_t^i satisfy assumption A4.

We now state the main stability result of this note which will be proved step by step in the remaining of this section.

Theorem 1: Let assumptions A1-A4 hold. Further, let the HCMCI algorithm be adopted and let $L \ge 1$. Finally, suppose that, for every $i \in \mathcal{N}$, the a priori information matrix $\mathbf{\Omega}_{1|0}^{i}$ is positive definite. Then, the estimation error $\mathbf{e}_t^i = \mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}^i$ is asymptotically bounded in mean square, i.e., $\limsup_{t\to\infty} \mathbb{E}\left\{\|\mathbf{e}_t^i\|^2\right\} < +\infty$, for any $i \in \mathcal{N}$. $\square$

In order to derive the above stability result, a first important step is the study of the properties of the information matrices $\mathbf{\Omega}_{t|t}^i$.

Lemma 1: Let the same assumptions as in Theorem 1 hold. Then, there exist positive definite matrices $\underline{\mathbf{\Omega}}$, $\overline{\mathbf{\Omega}}$, $\underline{\mathbf{\Omega}}^+$, and $\overline{\mathbf{\Omega}}^+$ such that $0 < \underline{\mathbf{\Omega}} \le \mathbf{\Omega}_{t|t}^i \le \overline{\mathbf{\Omega}}$ and $0 < \underline{\mathbf{\Omega}}^+ \le \mathbf{\Omega}_{t+1|t}^i \le \overline{\mathbf{\Omega}}^+$ for any $i \in \mathcal{N}$ and $t \ge 1$.

Proof: Let us first focus on the matrices $\mathbf{\Omega}_{t|t}^i$. By recalling the identity in the last row of Table IV and the fact that $\delta\mathbf{\Omega}_t^j = \left(\mathbf{C}^j\right)^T \mathbf{V}^j\,\mathbf{C}^j$, it can be seen that assumption A4 implies that

$$\begin{aligned} \mathbf{\Omega}_{t|t}^i &\le \sum_{j\in\mathcal{N}} \pi_L^{i,j} \mathbf{\Omega}_{t|t-1}^i + \sum_{j\in\mathcal{S}} \pi_L^{i,j} \bar{\omega} \left(\mathbf{C}^j\right)^T \mathbf{V}^j\,\mathbf{C}^j\,, && (7)\\ \mathbf{\Omega}_{t|t}^i &\ge \sum_{j\in\mathcal{N}} \pi_L^{i,j} \mathbf{\Omega}_{t|t-1}^i + \sum_{j\in\mathcal{S}} \pi_L^{i,j} \underline{\omega} \left(\mathbf{C}^j\right)^T \mathbf{V}^i\,\mathbf{C}^j\,. && (8) \end{aligned}$$

Suppose now that a CI algorithm is carried out with the noise covariance matrices $\mathbf{R}^j$ replaced by $\mathbf{R}^j/\bar{\omega}$ and let $\overline{\mathbf{\Omega}}_{t|t}^i$ and $\overline{\mathbf{\Omega}}_{t|t-1}^i$ be the resulting information matrices. Similarly, let $\underline{\mathbf{\Omega}}_{t|t}^i$ and $\underline{\mathbf{\Omega}}_{t|t-1}^i$ denote the information matrices resulting from a CI algorithm with the matrices $\mathbf{R}^j$ replaced by $\mathbf{R}^j/\underline{\omega}$. Recalling the first row of Table IV, it can be readily seen that

$$\begin{aligned} \overline{\mathbf{\Omega}}_{t|t}^i &= \sum_{j\in\mathcal{N}} \pi_L^{i,j} \overline{\mathbf{\Omega}}_{t|t-1}^i + \sum_{j\in\mathcal{S}} \pi_L^{i,j} \bar{\omega} \left(\mathbf{C}^j\right)^T \mathbf{V}^j\,\mathbf{C}^j\,, && (9)\\ \underline{\mathbf{\Omega}}_{t|t}^i &= \sum_{j\in\mathcal{N}} \pi_L^{i,j} \underline{\mathbf{\Omega}}_{t|t-1}^i + \sum_{j\in\mathcal{S}} \pi_L^{i,j} \underline{\omega} \left(\mathbf{C}^j\right)^T \mathbf{V}^i\,\mathbf{C}^j\,. && (10) \end{aligned}$$

Hence, provided that the same initializations are adopted for all the algorithms, i.e. $\overline{\mathbf{\Omega}}_{1|0}^i = \underline{\mathbf{\Omega}}_{1|0}^i = \mathbf{\Omega}_{1|0}^i$, one can see by means of simple induction arguments that $\underline{\mathbf{\Omega}}_{t|t}^i \le \mathbf{\Omega}_{t|t}^i \le \overline{\mathbf{\Omega}}_{t|t}^i$ for any $i \in \mathcal{N}$ and $t \ge 1$ (recall also that the prediction step of the Kalman filter recursion is monotone nondecreasing in the sense that $\underline{\mathbf{\Omega}}_{t|t}^i \le \mathbf{\Omega}_{t|t}^i \le \overline{\mathbf{\Omega}}_{t|t}^i$ implies $\underline{\mathbf{\Omega}}_{t+1|t}^i \le \mathbf{\Omega}_{t+1|t}^i \le \overline{\mathbf{\Omega}}_{t+1|t}^i$). Then, the first part of the proof is concluded by noting that, under assumptions A1-A3, there exist two positive definite matrices $\underline{\mathbf{\Omega}}$ and $\overline{\mathbf{\Omega}}$ such that $\underline{\mathbf{\Omega}}_{t|t}^i \ge \underline{\mathbf{\Omega}}$ and $\overline{\mathbf{\Omega}}_{t|t}^i \le \overline{\mathbf{\Omega}}$ (by virtue of Theorem 2 of [15] and the fact that $\mathbf{\Omega}_{1|0}^i > 0$). As for the bounds on $\mathbf{\Omega}_{t+1|t}^i$, one can simply take $\overline{\mathbf{\Omega}}^+ = \mathbf{W}$ and $\underline{\mathbf{\Omega}}^+ = \left(\mathbf{A}\underline{\mathbf{\Omega}}^{-1}\mathbf{A}^T + \mathbf{Q}\right)^{-1}$. $\square$

Lemma 1 ensures that the matrices $\mathbf{\Omega}_{t|t}^i$ are non-singular and, hence, that the estimation errors $\mathbf{e}_t^i$ are well defined. As it can be seen from the proof, this property crucially depends on network connectivity (assumption A3) and collective observability (assumption A2). The following result can now be stated.

Proposition 1: Let the same assumptions as in Theorem 1 hold. Then the estimation errors $\mathbf{e}_t^i$ obey the recursion

$$\mathbf{e}_{t+1}^i = \sum_{j\in\mathcal{N}} \mathbf{\Phi}_t^{i,j}\mathbf{e}_t^j + \sum_{j\in\mathcal{N}} \mathbf{\Gamma}_t^{i,j}\mathbf{v}_t^j + \mathbf{w_t} \tag{11}$$

for any $i \in \mathcal{N}$, where $\mathbf{\Phi}_t^{i,j} = \pi_L^{i,j}\,\mathbf{A}\left(\mathbf{\Omega}_{t|t}^i\right)^{-1}\mathbf{\Omega}_{t|t-1}^j$ and $\mathbf{\Gamma}_t^{i,j} = \pi_L^{i,j}\omega_t^i, \mathbf{A}\left(\mathbf{\Omega}_{t|t}^i\right)^{-1}\mathbf{V}_t^j$.

Proof: Notice first that $\mathbf{e}_{t+1}^i = \mathbf{A}(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t}^i) + \mathbf{w}_t$. Further, the estimate $\hat{\mathbf{x}}_{t|t}^i$ can be expressed as $\hat{\mathbf{x}}_{t|t}^i = \left(\mathbf{\Omega}_{t|t}^i\right)^{-1}\left[\sum_{j\in\mathcal{N}}\pi_L^{i,j}\mathbf{\Omega}_{t|t-1}^i\hat{\mathbf{x}}_{t|t-1}^j + \omega_t^i\sum_{j\in\mathcal{S}}\pi_L^{i,j}\left(\mathbf{C}^j\right)^T\mathbf{V}^j\,\mathbf{y}_t^j\right]$ with $\mathbf{y}_t^j = \mathbf{C}^j\mathbf{x}_t + \mathbf{v}_t^j$ (recall again the last row of Table IV). Then, equation (11) can be derived with straightforward calculations by recalling the expression for $\mathbf{\Omega}_{t|t}^i$ in the last row of Table IV and noting that the identity $\mathbf{x}_t = \left(\mathbf{\Omega}_{t|t}^i\right)^{-1}\left[\sum_{j\in\mathcal{N}}\pi_L^{i,j}\mathbf{\Omega}_{t|t-1}^i\mathbf{x}_t + \omega_t^i\sum_{j\in\mathcal{S}}\pi_L^{i,j}\left(\mathbf{C}^j\right)^T\mathbf{V}^j\mathbf{C}^{\mathbf{j}}\mathbf{x}_t\right]$ holds . $\square$

Consider now the noise-free collective dynamics of the estimation errors

$$\mathbf{e}_{t+1} = \mathbf{\Phi}_t\mathbf{e}_t \tag{12}$$

where $\mathbf{e}_t = \mathrm{col}\left(\mathbf{e}_t^i,\, i \in \mathcal{N}\right)$ and $\mathbf{\Phi}_t$ is a block matrix whose block elements are given by the matrices $\mathbf{\Phi}_t^{i,j}$ defined in Proposition 1. Then, the following result holds.

Lemma 2: Let the same assumptions as in Theorem 1 hold. Then the time-varying system (12) is uniformly exponentially stable.

Proof: Let $\mathbf{p}$ be the Perron-Frobenius left eigenvector of the matrix $\mathbf{\Pi}^L$. Notice that, by virtue of Assumption A3, such an eigenvector has strictly positive components $p^i, i \in \mathcal{N}$, and satisfies the equation $\mathbf{p}^T\mathbf{\Pi}^L = \mathbf{p}^T$, i.e., $\sum_{j\in\mathcal{N}} p^j\pi_L^{j,i} = p^i$. Consider now the candidate Lyapunov function

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

[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.

for system (12). In view of Lemma 1, it is immediate to see that there exist suitable strictly positive constants α_1 and α_2 such that $\alpha_1\|\mathbf{e}_t\|^2 \leq \mathscr{V}_t(\mathbf{e}_t) \leq \alpha_2\|\mathbf{e}_t\|^2$. Moreover, exploiting the fact that $\boldsymbol{\Omega}^i_{t+1|t} \leq \tilde{\beta}\mathbf{A}^{-T}\boldsymbol{\Omega}^i_{t|t}\mathbf{A}^{-1}$ for some positive real $\tilde{\beta} < 1$ (see point iii) in Lemma 1 of [15]), it turns out that

$$\begin{aligned}\left(\mathbf{e}^i_{t+1}\right)^T \boldsymbol{\Omega}^i_{t+1|t}\mathbf{e}^i_{t+1} &= \left(\sum_{j\in\mathcal{N}} \boldsymbol{\Phi}^{i,j}_t \mathbf{e}^j_t\right)^T \boldsymbol{\Omega}^i_{t+1|t} \sum_{j\in\mathcal{N}} \boldsymbol{\Phi}^{i,j}_t \mathbf{e}^j_t \\ &\leq \tilde{\beta}\left(\sum_{j\in\mathcal{N}} \pi^{i,j}_L \boldsymbol{\Omega}^j_{t|t-1}\mathbf{e}^j_t\right)^T \left(\boldsymbol{\Omega}^i_{t|t}\right)^{-1} \sum_{j\in\mathcal{N}} \pi^{i,j}_L \boldsymbol{\Omega}^j_{t|t-1}\mathbf{e}^j_t \\ &\leq \tilde{\beta} \sum_{j\in\mathcal{N}} \pi^{i,j}_L \left(\mathbf{e}^j_t\right)^T \boldsymbol{\Omega}^j_{t|t-1}\mathbf{e}^j_t\end{aligned}$$

where the latter inequality follows from the fact that $\boldsymbol{\Omega}^i_{t|t} \geq \sum_{j\in\mathcal{N}} \pi^{i,j}_L \boldsymbol{\Omega}^j_{t|t-1}$ and from Lemma 2 of [15]. As a consequence, it is possible to write

$$\begin{aligned}\mathscr{V}_{t+1}(\mathbf{e}_{t+1}) &= \sum_{i\in\mathcal{N}} p^i \left(\mathbf{e}^i_{t+1}\right)^T \boldsymbol{\Omega}^i_{t+1|t}\mathbf{e}^i_{t+1} \\ &\leq \tilde{\beta} \sum_{i,j\in\mathcal{N}} p^i \pi^{i,j}_L \left(\mathbf{e}^j_t\right)^T \boldsymbol{\Omega}^j_{t|t-1}\mathbf{e}^j_t \\ &= \tilde{\beta} \sum_{j\in\mathcal{N}} p^j \left(\mathbf{e}^j_t\right)^T \boldsymbol{\Omega}^j_{t|t-1}\mathbf{e}^j_t = \tilde{\beta}\,\mathscr{V}_t(\mathbf{e}_t)\end{aligned}$$

from which uniform exponential stability of (12) follows at once. □

Hence, the statement of Theorem 1 readily follows from Lemma 2 above and the fact that the terms $\sum_{j\in\mathcal{N}} \boldsymbol{\Gamma}^{i,j}_t \mathbf{v}^j_t + \mathbf{w_t}$ in (11) are bounded in mean square by virtue of Lemma 1.

Remark 1: Notice that the agreement on the matrices $\boldsymbol{\Omega}^i$ plays a fundamental role in the Lyapunov stability analysis, in that such matrices are directly associated with the possibility of defining the candidate Lyapunov function used in the proof of Theorem 1. Hence their presence is important even if, in general, a direct interpretation of such matrices as inverse covariances is not possible.

V. SIMULATION EXAMPLES

The aim of this section is to investigate how different types of consensus-based distributed state estimators compare with each other and with the centralized state estimator. To this end, a single-target tracking case study will be considered, where the target motion is modelled by a linear (nearly constant velocity) model (see [15]) where $\mathbf{x}_t = [x_t, \dot{x}_t, y_t, \dot{y}_t]^T$ is the kinematic target state at sampling time t made up of the Cartesian coordinates of position (x_t, y_t) and of velocity $(\dot{x}_t, \dot{y}_t)$; the sampling interval has been fixed to $1\ s$ and the variance of the random fluctuations of target speed to $0.25\ m^2/s^3$. Two different simulation scenarios corresponding to two different sensor networks will be considered.

In the first scenario, the network is composed of 100 communication nodes and 5 linear position sensor nodes characterized by the measurement equation (6) with

$$\mathbf{C}^i \;=\; \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \forall i \in \mathcal{S}\,.$$

Conversely, in the second scenario two types of nonlinear position sensors measuring angle or, respectively, distance have been considered. 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:

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

TABLE VI: Percentage performance degradation

		HCMCI-1				HCMCI-2			
P_d	P_L	1	2	3	4	1	2	3	4
0.9	0	5.5	5.5	4.9	3.5	5.8	5.9	5.6	5.4
1	0.2	7.7	2.5	1.6	1.7	4.5	1.5	1.2	1.1
0.9	0.2	12	6.6	6	4.6	9.8	6.7	6.1	5.8

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. Graphical representations of the two sensor networks are provided in Fig. 1. The measurement noise will be assumed to have $\sigma_x = \sigma_y = 10\ m$ standard deviation for linear sensors, $\sigma_\theta = 2^\circ$ standard deviation for DOA sensors, and $\sigma_r = 10\ m$ standard deviation for TOA sensors. Finally, the consensus weights used in the simulations have been set equal to the Metropolis weights.

In each of the considered simulation settings 200 independent Monte Carlo trials have been performed and the position root mean square error (PRMSE), averaged over time and over all the network nodes, has been computed as performance index. Fig 2 shows the comparison between the following state estimators: C(E)KF, CM with weights ω^i_t as in (4), CI, HCMCI with weights ω^i_t as in (4) (denoted hereafter as HCMCI-1), HCMCI with weights $\omega^i_t = |\mathcal{N}|$ (denoted hereafter as HCMCI-2). Notice that the minimum numbers of consensus steps for which the CM algorithm exhibits a stable behavior for the two networks are $L = 4$ and $L = 5$, respectively. Accordingly the PRMSEs of the CM for lower values of L are omitted in the plots since out of scale. The simulation results confirm the theoretical analysis by showing that HCMCI, for both the considered weight choices, ensures stability already for $L = 1$ and outperforms CI thanks to the presence of the scalars ω^i_t weighting the novel information. Further, by comparing the PRMSEs of HCMCI-1 and HCMCI-2, it can be seen that, at least in the considered set-up, the choice $\omega^i_t = |\mathcal{N}|$, which is asymptotically optimal as $L \to +\infty$, is not necessarily the best one when only a limited number of consensus steps per iteration can be performed. Of course, different choices of the weights ω^i_t can be convenient depending on the particular setup.

Additional simulations have been carried out by considering for each sensor node a probability P_d of missed measurement and for each network link a probability P_L of data loss. Table VI reports the percentage performance degradation of HCMCI-1 and HCMCI-2 (for different values of L) relative to the ideal case $P_d = 1$ and $P_L = 0$. It can be seen that missed measurements and data dropouts affect in a similar way the two filters and, as expected, the largest performance degradation occurs for $L = 1$.

VI. CONCLUDING REMARKS

While the stability analysis carried out in this note pertains only to the LTI case, it turns out that it could be used as a starting point for studying the stability of the estimation error dynamics also in a nonlinear setting. In fact, under suitable continuity assumptions it is possible to write the estimation error dynamics in a suitable way so that the linearized part is separated from the nonlinear (higher-order) terms Then, an analysis similar to the one of Section IV could be carried out to prove stability of the linearized part of the estimation error dynamics, which in turn is the key for achieving a local stability result for the overall dynamics. Results in this direction can be found in [20] to which the interested reader is referred.

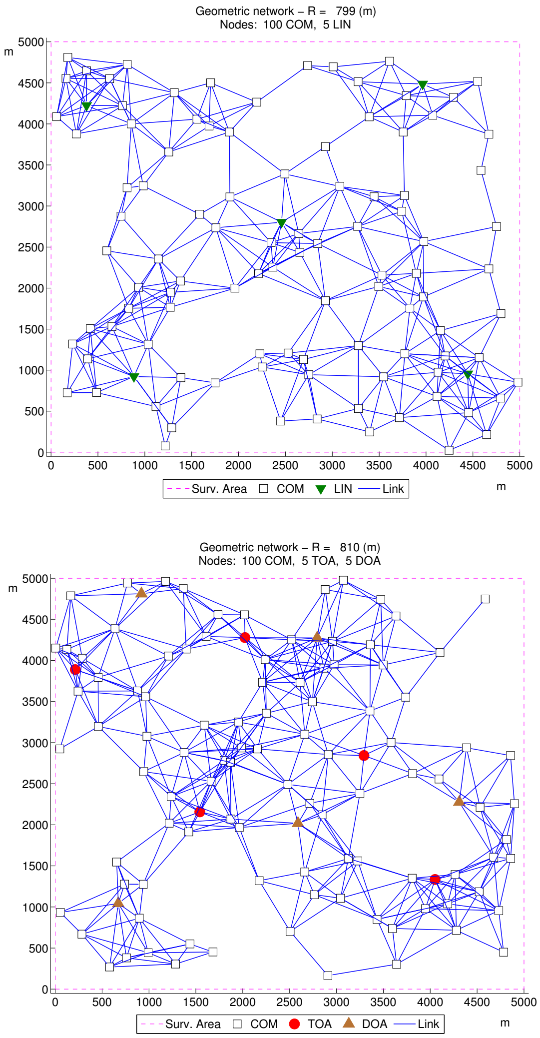

Fig. 1: Linear (top) and nonlinear (bottom) sensor networks used in the simulations.

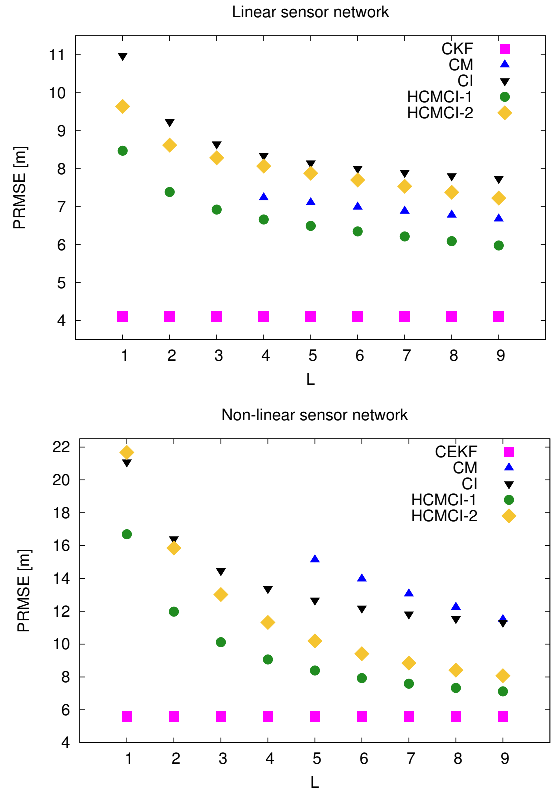

Fig. 2: PRMSEs of the considered filters on the linear (top) and nonlinear (bottom) networks for different values of L.

References

[1] R. Olfati-Saber and R.M. Murray, "Consensus problems in networks of agents with swtiching topology and time-delays", *IEEE Trans. on Automatic Control*, vol. 49, no. 9, pp. 1520-1533, 2004.

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

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

[4] R. Carli, A. Chiuso, L. Schenato and S. Zampieri, "Distributed Kalman filtering based on consensus strategies", *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 622–633, 2008.

[5] R. Olfati-Saber, "Kalman-consensus filter : Optimality, stability, and performance", *Proc. Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, pp. 7036–7042, 2009.

[6] S.S. Stankovic, M.S. Stankovic and D.M. Stipanovic, "Consensus based overlapping decentralized estimation with missing observations and communication faults", *Automatica*, vol. 45, no. 6, pp. 1397–1406, 2009.

[7] F.S. Cattivelli and A.H. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing", *IEEE Trans. on Automatic Control*, vol. 55, no. 9, pp. 2069–2084, 2010.

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

[9] F.S. Cattivelli and A.H. Sayed, "Distributed nonlinear Kalman filtering with applications to wireless localization", *Proc. 2010 IEEE International Conference on Acoustics Speech and Signal Processing*, pp. 3522-3525, 2010.

[10] 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.

[11] U.A. Khan and A. Jadbabaie, "Networked estimation under information constraints", *CoRR*, abs/1111.4580, 2011.

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

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

[14] 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.

[15] G. Battistelli and L. Chisci, "Kullback-Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability", *Automatica*, vol. 50, no. 3, pp. 707–718, 2014.

[16] 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.

[17] A.T. Kamal, J.A., Farrell and A.K Roy-Chowdhury, "Information weighted consensus", *Proc. 51st IEEE Conference on Decision and Control*, pp. 2732–2737, 2012.

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

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

[20] G. Battistelli and L. Chisci, "Stability of consensus extended Kalman filtering for distributed state estimation", *Preprints 19th IFAC World Congress*, pp. 5520-5525, 2014.