Distributed moving horizon estimation for nonlinear constrained systems

Marcello Farina [*] **Giancarlo Ferrari-Trecate** [**] **Riccardo Scattolini** [*]

[*] *Dipartimento di Elettronica e Informazione, Politecnico di Milano, Via Ponzio 34/5, 20133, Milan, Italy.*
[**] *Dipartimento di Informatica e Sistemistica, Università degli Studi di Pavia, Via Ferrata 1, 27100 Pavia, Italy.*

Abstract: In this paper we consider a nonlinear constrained system observed by a sensor network and propose a distributed state estimation scheme based on Moving Horizon Estimation (MHE). In order to embrace the case where the whole system state cannot be reconstructed from data available to individual sensors, we resort to the notion of MHE-detectability for nonlinear systems, and add to the MHE problems solved by each sensor a consensus term for propagating information about estimates through the network. Under some suitable assumptions we prove convergence to zero and stability of the state estimation error provided by any sensor.

Keywords: Distributed estimation, moving horizon estimation, consensus algorithms, nonlinear systems detectability.

1. INTRODUCTION

State estimation for nonlinear systems based on distributed sensing schemes is a challenging problem, the solution of which is of great importance in many fields. Distributed monitoring, exploration, surveillance and tracking of moving objects over specific regions are topical applications, due to the wide diffusion of sensor networks in the last decade. Sensor networks are collections of small, low power consuming and possibly cheap sensing devices, with communication and computation capabilities.

Available methods for distributed state estimation rely on local state estimators for linear systems combined with consensus and sensor-fusion algorithms. Typically, each sensor provides an estimate of the system state based on local data and consensus schemes are employed either to provide a wider set of measurement data for each individual sensor (*i.e.*, *consensus on measurements*) or to correct individual state estimates by comparison with neighboring nodes information (*i.e.*, *consensus on estimates*).

Approaches based on Kalman filters are discussed, e.g., in Carli et al. (2008), Alriksson and Rantzer (2006), Olfati-Saber (2005), Olfati-Saber (2007), Kamgarpour and Tomlin (2008), Olfati-Saber (2009). The algorithm described in Olfati-Saber (2005) relies on consensus on measurements, while in Olfati-Saber (2007) a solution based on consensus on estimates is proposed. Recently, convergence in mean of the state estimates obtained with the algorithm presented in Olfati-Saber (2005) has been proved in Kamgarpour and Tomlin (2008) and a stability analysis of the state estimator presented in Olfati-Saber (2007) is provided in Olfati-Saber (2009).
A two-step optimization procedure relying on consensus on

estimates is used in Alriksson and Rantzer (2006) and in Carli et al. (2008).

Methods based on Kalman filtering may become suboptimal or even unstable when constraints on noise and state variables are present. This motivated the development of centralized MHE schemes for linear (Rao et al., 2001; Alessandri et al., 2003), nonlinear (Rao et al., 2003; Rawlings and Mayne, 2009; Alessandri et al., 2008, 1999) and hybrid (Ferrari-Trecate et al., 2002) systems, capable to guarantee observer convergence and/or stability in a constrained setting.
A distributed MHE (DMHE) method for linear constrained systems has been proposed by the authors of the present paper in Farina et al. (2009a) and Farina et al. (2009b).

In this paper we generalize our previous results to the nonlinear setting with the goal of providing a Nonlinear DMHE (NDMHE) scheme enjoying stability properties. In order to characterize states that can and cannot be recovered by each sensor without communication we exploit the notion of MHE detectability (Rawlings and Mayne, 2009). Moreover we use a consensus-on-estimates penalty term in local MHE problems to let each sensor learn locally MHE-undetectable parts of the state from other sensors.

The paper is structured as follows. In Section 2 we introduce the observed dynamical system, the structure of the sensor network, and we recall notions of detectability for nonlinear systems. In Section 3 we describe the distributed state estimation algorithm. In Section 4 we investigate the stability and convergence properties of the presented observer and Section 5 reports some concluding remarks. For the sake of clarity the proofs are collected in the Appendix.

Notation. I_n and $\mathbf{0}_{\nu \times \mu}$ denote the $n \times n$ identity matrix and the $\nu \times \mu$ matrix of zero elements, respectively. The symbol $\otimes$ denotes the Kronecker product, and $\mathbb{1}_M$ is the *M*-dimensional

column vector whose entries are all equal to 1. The matrix $\text{diag}(M_1,\ldots,M_s)$ is block-diagonal with blocks M_i. We use the short-hand $\mathbf{v} = (v_1,\ldots,v_s)$ to denote a column vector with s (not necessarily scalar) components. For a discrete-time signal $w(t)$ and $a,b \in \mathbb{N}$, $a \leq b$, we denote $(w(a),w(a+1),\ldots,w(b))$ with $w_{[a:b]}$. For the definition of positive-definite, $\mathscr{K}$, $\mathscr{K}_\infty$ and $\mathscr{K}\mathscr{L}$ functions we defer the reader to Rawlings and Mayne (2009). Finally, the notation $\|z\|_S^2$ stands for $z^T S z$, where S is a symmetric positive-semidefinite matrix.

2. THE SYSTEM AND ITS OBSERVABILITY PROPERTIES

2.1 System and sensor network

We assume that the observed process obeys to the dynamics

$$x_{t+1} = f(x_t, w_t) \tag{1}$$

where $x_t \in \mathbb{X} \subseteq \mathbb{R}^n$ is the state vector and the term $w_t \in \mathbb{W} \subseteq \mathbb{R}^n$ represents an unknown disturbance term. We assume that the sets $\mathbb{X}$ and $\mathbb{W}$ are convex and that $\mathbb{W}$ contains the origin. Furthermore, $f(x,w)$ has continuous partial derivatives with respect to the components w^j of w, $j = 1,\ldots,n$, and is globally Lipschitz with respect to w i.e., $\exists l > 0 : \forall x \in \mathbb{X}$ and $\forall w_1, w_2 \in \mathbb{W}$

$$\|f(x,w_1) - f(x,w_2)\| \leq l\|w_1 - w_2\| \tag{2}$$

Measurements on the state vector are performed by M sensors, according to the sensing models (in general different from sensor to sensor)

$$y_t^i = h^i(x_t) + v_t^i, \ i = 1,\ldots,M \tag{3}$$

where the term $v_t^i \in \mathbb{R}^{p_i}$ represents an unknown measurement error.

The communication network among sensors is modeled by the directed graph $\mathscr{G} = (\mathscr{V},\mathscr{E})$, where the nodes in $\mathscr{V} = \{1,2,\ldots,M\}$ are sensors and an edge (j,i) in the set $\mathscr{E} \subseteq \mathscr{V} \times \mathscr{V}$ models that sensor j can transmit information to sensor i. We assume $(i,i) \in \mathscr{E}$, $\forall i \in \mathscr{V}$. We denote with $\mathscr{V}_i$ the set of the neighbors to node i, i.e., $\mathscr{V}_i = \{j \in \mathscr{V} : (j,i) \in \mathscr{E}\}$.

We associate to the graph $\mathscr{G}$ the stochastic matrix $K \in \mathbb{R}^{M \times M}$, with entries

$$k_{ij} \geq 0 \text{ if } (j,i) \in \mathscr{E} \tag{4a}$$
$$k_{ij} = 0 \text{ otherwise} \tag{4b}$$
$$\sum_{j=1}^M k_{ij} = 1, \ \forall i = 1,\ldots,M \tag{4c}$$

Any matrix K with entries satisfying (4) is said to be *compatible* with $\mathscr{G}$.

At a generic time instant t, sensor i collects measurements produced by itself and its neighboring sensors. Moreover, each sensor transmits and receives information once within a sampling interval i.e., measurements available to node i are y_t^j, with $j \in \mathscr{V}_i$.

Three types of quantities can be distinguished: *individual*, *regional*, and *collective*. Specifically, a quantity is referred to as: (*a*) *individual* (with respect to sensor i) when it is related to the node i solely; (*b*) *regional* (with respect to sensor i) if it is related to the nodes in $\mathscr{V}_i$; (*c*) *collective*, if it is related to the whole network. For the sake of clarity, we use different notations for individual, regional and collective variables. Namely, given a variable z, z^i, $\bar{z}^i$ and $\mathbf{z}$ represent its individual, regional and collective version, respectively. For instance, we refer to y_t^i

in (3) as individual measurement. On the other hand, if $\mathscr{V}_i = \{j_1^i, \ldots, j_{v_i}^i\}$, the regional measurement of node i is given by

$$\bar{y}_t^i = \bar{h}^i(x_t) + \bar{v}_t^i \tag{5}$$

where $\bar{y}_t^i = (y_t^{j_1^i}, \ldots, y_t^{j_{v_i}^i})$, $\bar{h}^i(x_t) = (h^{j_1^i}(x_t), \ldots, h^{j_{v_i}^i}(x_t))$, and $\bar{v}_t^i = (v_t^{j_1^i}, \ldots, v_t^{j_{v_i}^i})$. The dimension of vectors $\bar{y}_t^i$ and $\bar{v}_t^i$ is $\bar{p}_i = \sum_{k=1}^{v_i} p_{j_k^i}$.

2.2 Detectability properties

In the sequel we employ the notion of MHE detectability for nonlinear systems introduced in Rawlings and Mayne (2009)

Definition 1. The system $x_{k+1} = f(x_k, w_k^1)$, $y_k = h(x_k)$ is *MHE detectable* if the system augmented with an extra disturbance w_k^2

$$x_{k+1} = f(x_k, w_k^1) + w_k^2 \tag{6a}$$
$$y_k = h(x_k) \tag{6b}$$

is incrementally input-output-to-state-stable (δIOSS) with respect to the augmented disturbances $\tilde{w}_k = (w_k^1, w_k^2)$. Namely, there exist functions $\beta_D \in \mathscr{K}\mathscr{L}$, $\gamma_1, \gamma_2 \in \mathscr{K}$ such that, for every two initial states z and z^* and two disturbance sequences $\tilde{w}_{[0:k]}$ and $\tilde{w}_{[0:k]}^*$ and, given the corresponding output sequences $y_{[0:k]}$ and $y_{[0:k]}^*$, it holds that

$$\|x_t - x_t^*\| \leq \beta_D(\|z - z^*\|, t) + \gamma_1(\|\tilde{w}_k - \tilde{w}_k^*\|_{[0:t-1]}) + \\ + \gamma_2(\|y_k - y_k^*\|_{[0:t]})$$

where x_k and x_k^* are the state sequences stemming (through system (6a)) from z, $\tilde{w}_{[0:k]}$ and from z^*, $\tilde{w}_{[0:k]}^*$, respectively. $\square$

According to the terminology presented in the previous section, three different MHE detectability notions can be introduced.

Definition 2. The system is *individually MHE detectable* by sensor i (sensor i is individually MHE detectable) if the system $x_{k+1} = f(x_k, w_k)$, $y_k^i = h^i(x_k)$ is MHE detectable. The system is *regionally MHE detectable* by sensor i (sensor i is regionally MHE detectable) if the system $x_{k+1} = f(x_k, w_k)$, $\bar{y}_k^i = \bar{h}^i(x_k)$ is MHE detectable. The system is *collectively MHE detectable* if the system $x_{k+1} = f(x_k, w_k)$, $\mathbf{y}_k^* = \mathbf{h}^*(x_k)$, where $\mathbf{y}_k^* = (y_k^1, \ldots, y_k^M)$ and $\mathbf{h}^*(x_k) = (h^1(x_k), \ldots, h^M(x_k))$ is MHE detectable. $\square$

Notice that, for a given sensor i, individual detectability implies regional detectability, and regional detectability of any sensor implies collective detectability, while all opposite implications are false.

In the sequel, we consider the case where the system state is not necessarily MHE detectable by individual sensors, and that there exist, for each sensor node, a suitable change of coordinates allowing regionally detectable and undetectable states to be clearly identified. Given the regional model (1),(5), for sensor i we assume that there exists a diffeomorphism $T_i : \mathbb{R}^n \to \mathbb{R}^n$, $T_i^{-1} : x_t \mapsto \xi_t^i = T_i^{-1}(x_t)$ such that, by changing coordinates, and being $\xi_t^i = (\xi_t^{UD,i}, \xi_t^{D,i})$ the state of the equivalent system, one has

$$\xi_{t+1}^{UD,i} = f^{UD,i}(\xi_t^{UD,i}, \xi_t^{D,i}, w_t^i) \tag{7a}$$
$$\xi_{t+1}^{D,i} = f^{D,i}(\xi_t^{D,i}, w_t^i) \tag{7b}$$
$$\bar{y}_t^i = \tilde{h}^i(\xi_t^{D,i}) + \bar{v}_t^i \tag{7c}$$

where the subsystem (7b), (7c) is MHE detectable. It follows that $\xi_t^{UD,i}$ and $\xi_t^{D,i}$ denote regionally MHE undetectable and regionally MHE detectable components of ξ_t^i, respectively. We define $f_\xi^i = (f^{UD,i}, f^{D,i})$. Furthermore, we will assume that the functions T_i satisfy the following assumption.

Assumption 1. For all $i \in \mathcal{V}$, there exist $\alpha_T, \alpha_{T^{-1}} > 0$ such that

$$\left\| \left. \frac{\partial T_i(\xi)}{\partial \xi} \right|_{\xi^i} \right\| \leq \alpha_T \tag{8a}$$

$$\left\| \left. \frac{\partial T_i^{-1}(x)}{\partial x} \right|_{\tilde{x}} \right\| \leq \alpha_{T^{-1}} \tag{8b}$$

for $\tilde{\xi}^i = T_i^{-1}(\tilde{x})$, for all $\tilde{x} \in \mathbb{R}^n$.

Let $\bar{P}_{UD}^i$ and $\bar{P}_D^i$ be the $n \times n$ orthonormal projection matrices defined in such a way that $\bar{P}_{UD}^i \xi_t^i = (\xi_t^{UD,i}, 0)$ and $\bar{P}_D^i \xi_t^i = (0, \xi_t^{D,i})$, respectively. Furthermore, let the map $\mathbf{T} : \mathbb{R}^{M \cdot n} \to \mathbb{R}^{M \cdot n}$ be such that, for $\boldsymbol{\xi}_t = (\xi_t^1, \dots, \xi_t^M)$, one has $\mathbf{T}(\boldsymbol{\xi}_t) = (T^1(\xi_t^1), \dots, T^1(\xi_t^M))$. Finally, we define $\mathbf{P}_{UD} = \text{diag}(\bar{P}_{UD}^1, \dots, \bar{P}_{UD}^M)$ and $\mathbf{P}_D = \text{diag}(\bar{P}_D^1, \dots, \bar{P}_D^M)$. Note that, if the system is regionally MHE-detectable by sensor i, then $\bar{P}_{UD}^i = \mathbf{0}_{n \times n}$ and $\bar{P}_D^i = I_n$.

3. THE DISTRIBUTED ESTIMATION ALGORITHM

Our aim is to design, for a generic sensor $i \in \mathcal{V}$, an algorithm for computing an estimate of the system state based on regional measurements $\bar{y}_t^i$ and further pieces of information provided by sensors $j \in \mathcal{V}_i$. The proposed solution relies on the use of Moving Horizon Estimation (MHE), (Rao et al., 2001, 2003; Rao and Rawlings, 2000; Alessandri et al., 1999; Goodwin et al., 2005; Rao et al., 1999), in view of its capability to handle state and noise constraints. More specifically, we propose a Distributed MHE scheme for nonlinear systems (NDMHE) where each sensor solves an individual MHE problem.

3.1 The individual minimization problem

Each node $i \in \mathcal{V}$, for a given estimation horizon $N \geq 1$, at time t determines the estimates $\hat{x}^i$ and $\hat{w}^i$ of x and w, respectively, by solving the constrained minimization problem *MHE-i* defined as

$$\Theta_t^{*i} = \min_{\hat{x}_{t-N}^i, \{\hat{w}_k^i\}_{k=t-N}^t} J^i(t-N, t, \hat{x}_{t-N}^i, \hat{w}^i, \hat{\bar{v}}^i, \Gamma_{t-N}^i) \tag{9}$$

under the constraints

$$\hat{x}_{k+1}^i = f(\hat{x}_k^i, \hat{w}_k^i), \quad k = t-N, \dots, t \tag{10a}$$

$$\bar{y}_k^i = \bar{h}^i(\hat{x}_k^i) + \hat{\bar{v}}_k^i \tag{10b}$$

$$\hat{w}_k^i \in \mathbb{W} \tag{10c}$$

$$\hat{x}_k^i \in \mathbb{X} \tag{10d}$$

Let t_1 verify $t - N \leq t_1 \leq t$. We define the *transit cost* of a generic state $z \in \mathbb{R}^n$ at time t_1, computed at instant t as

$$\Xi_{t_1/t}^i(z) = \min_{\hat{x}_{t-N}^i, \{\hat{w}_k^i\}_{k=t-N}^t} \left\{ J^i(t-N, t, \hat{x}_{t-N}^i, \hat{w}^i, \hat{\bar{v}}^i, \Gamma_{t-N}^i) \right.$$
$$\left. \text{subject to (10) and } \hat{x}_{t_1}^i = z \right\} \tag{11}$$

Note that the associated optimization problem is feasible for all $z \in \mathcal{Z} = f(\mathbb{X}, \mathbb{W}) \cap \mathbb{X}$ and therefore $\mathcal{Z}$ is the domain of $\Xi_{t_1/t}^i(z)$. The individual cost function J^i is given by

$$J^i(t-N, t, \hat{x}_{t-N}^i, \hat{w}^i, \hat{\bar{v}}^i, \Gamma_{t-N}^i) = \sum_{k=t-N}^t L^i(\hat{\bar{v}}_k^i, \hat{w}_k^i) + \Gamma_{t-N}^i(\hat{x}_{t-N}^i)$$
$$\Gamma_{t-N}^i(\hat{x}_{t-N}^i) = \Gamma_{t-N}^{C,i}(\hat{x}_{t-N}^i; \hat{\bar{x}}_{t-N/t-1}^i) +$$
$$+ \Gamma_{t-N}^{0,i}(\hat{x}_{t-N}^i; \hat{x}_{t-N/t-1}^i) + \Theta_{t-1}^{*i} \tag{12}$$

In (12), the function L^i is the stage cost, $\Gamma_{t-N}^{C,i}$ is the consensus initial penalty and $\Gamma_{t-N}^{0,i}$ is the regularization initial cost. They should be defined in order to satisfy the following assumption.

Assumption 2. The stage costs L^i and the initial penalties $\Gamma_{t-N}^{C,i}$ and $\Gamma_{t-N}^{0,i}$ are continuous, bounded, positive definite and they satisfy the following inequalities for all $w \in \mathbb{R}^n$, $\bar{v} \in \mathbb{R}^{\bar{p}_i}$, $\hat{x}_0^i, \hat{\bar{x}}_{0/N-1}^i \in \mathbb{R}^n$

$$\underline{\gamma}_L(\|(\bar{v}, w)\|) \leq L^i(\bar{v}, w) \tag{13a}$$

$$\Gamma_0^{C,i}(\hat{x}_0^i; \hat{\bar{x}}_{0/N-1}^i) \leq \gamma_0(\|\hat{x}_0^i - \hat{\bar{x}}_{0/N-1}^i\|) \tag{13b}$$

$$\Gamma_0^{0,i}(\hat{x}_0^i; \hat{x}_{0/N-1}^i) \leq \gamma_0(\|\hat{x}_0^i - \hat{x}_{0/N-1}^i\|) \tag{13c}$$

where $\underline{\gamma}_L$ and γ_0 are suitable $\mathcal{K}_\infty$ functions.

We denote with $\hat{x}_{t-N/t}^i$ and with $\left\{\hat{w}_{k/t}^i\right\}_{k=t-N}^t$ the optimizers to (9) and with $\hat{x}_{k/t}^i$, $k = t-N, \dots, t$ the individual state sequence stemming from $\hat{x}_{t-N/t}^i$ and $\left\{\hat{w}_{k/t}^i\right\}_{k=t-N}^t$. Furthermore

$$\hat{\bar{x}}_{k/t}^i = \sum_{j=1}^M k_{ij} \hat{x}_{k/t}^j \tag{14}$$

denotes the weighted average of the state estimates produced by sensors $j \in \mathcal{V}^i$.

Similarly to Rao et al. (2003) and Farina et al. (2009a), suitable choices of $\Gamma_{t-N}^{C,i}$ and $\Gamma_{t-N}^{0,i}$ fulfilling Assumption 2 are the quadratic functions $\Gamma_{t-N}^{C,i} = \|\hat{x}_{t-N}^i - \hat{\bar{x}}_{t-N/t-1}^i\|^2_{(\Pi_{t-N}^{C,i})^{-1}}$ and $\Gamma_{t-N}^{0,i} = \|\hat{x}_{t-N}^i - \hat{x}_{t-N/t-1}^i\|^2_{(\Pi_{t-N}^{0,i})^{-1}}$ where the matrices $\Pi_{t-N}^{C,i}$ and $\Pi_{t-N}^{0,i}$ must be definite positive and bounded. Furthermore, they must be assigned in order to satisfy conditions guaranteeing stability, which will be discussed later on.

The penalty term $\Gamma_{t-N}^{C,i}$ embodies a *consensus-on-estimates* term, in the sense that it penalizes the deviation of $\hat{x}_{t-N}^i$ from $\hat{\bar{x}}_{t-N/t-1}^i$. Consensus, besides increasing accuracy of the individual estimates, is fundamental to guarantee convergence of the state estimates to the state of the observed system even if regional MHE detectability does not hold. In other words, it allows sensor i to reconstruct components of the state that cannot be estimated by the i-th regional model.

Finally notice that, since the cost (12) and the constraints (10) depend only upon regional variables, the overall estimation scheme is decentralized.

3.2 The collective minimization problem

The individual estimation problem (9) can be given a collective form. To this end, let $\mathbf{J}$ be the collective cost function given by

$$\mathbf{J}(\cdot) = \sum_{i=1}^M J^i(t-N, t, \hat{x}_{t-N}^i, \hat{w}^i, \hat{\bar{v}}^i, \Gamma_{t-N}^i) \tag{15}$$

Define the collective vectors $\hat{\mathbf{x}}_t = (\hat{x}_t^1, \ldots, \hat{x}_t^M)$, $\hat{\bar{\mathbf{v}}}_t = (\hat{\bar{v}}_t^1, \ldots, \hat{\bar{v}}_t^M)$, $\hat{\mathbf{w}}_t = (\hat{w}_t^1, \ldots, \hat{w}_t^M)$, the matrix $\mathbf{K} = K \otimes I_n$, the quantity $\boldsymbol{\Theta}_t^* = \sum_{i=1}^M \Theta_t^{*i}$ and the collective costs

$$\mathbf{L}(\hat{\bar{\mathbf{v}}}_k, \hat{\mathbf{w}}_k) = \sum_{i=1}^M L^i(\hat{\bar{v}}_k^i, \hat{w}_k^i) \tag{16a}$$

$$\boldsymbol{\Gamma}_{t-N}^C(\hat{\mathbf{x}}_{t-N}; \mathbf{K}\hat{\mathbf{x}}_{t-N/t-1}) = \sum_{i=1}^M \Gamma_{t-N}^{C,i}(\hat{x}_{t-N}^i; \hat{\bar{x}}_{t-N/t-1}^i) \tag{16b}$$

$$\boldsymbol{\Gamma}_{t-N}^0(\hat{\mathbf{x}}_{t-N}; \hat{\mathbf{x}}_{t-N/t-1}) = \sum_{i=1}^M \Gamma_{t-N}^{0,i}(\hat{x}_{t-N}^i; \hat{x}_{t-N/t-1}^i) \tag{16c}$$

$$\boldsymbol{\Gamma}_{t-N}(\hat{\mathbf{x}}_{t-N}) = \boldsymbol{\Gamma}_{t-N}^C(\hat{\mathbf{x}}_{t-N}; \mathbf{K}\hat{\mathbf{x}}_{t-N/t-1}) + \tag{16d}$$
$$+ \boldsymbol{\Gamma}_{t-N}^0(\hat{\mathbf{x}}_{t-N}; \hat{\mathbf{x}}_{t-N/t-1}) + \boldsymbol{\Theta}_{t-1}^*$$

Then, the collective cost function $\mathbf{J}$ can be rewritten as

$$\mathbf{J}(t-N, t, \hat{\mathbf{x}}_{t-N}, \hat{\mathbf{w}}, \hat{\bar{\mathbf{v}}}, \boldsymbol{\Gamma}_{t-N}) = \sum_{k=t-N}^t \mathbf{L}(\hat{\bar{\mathbf{v}}}_k, \hat{\mathbf{w}}_k) + \boldsymbol{\Gamma}_{t-N}(\hat{\mathbf{x}}_{t-N}) \tag{17}$$

Defining $\mathbf{f}(\hat{\mathbf{x}}_k, \hat{\mathbf{w}}_k) = (f(\hat{x}_k^1, \hat{w}_k^1), \ldots, f(\hat{x}_k^M, \hat{w}_k^M))$, $\bar{\mathbf{y}}_k = (\bar{y}_k^1, \ldots, \bar{y}_k^M)$ and $\mathbf{h}(\hat{\mathbf{x}}_k) = (\bar{h}^1(\hat{x}_k^1), \ldots, \bar{h}^M(\hat{x}_k^M))$, also the constraints (10) can be written in the following collective form

$$\hat{\mathbf{x}}_{k+1} = \mathbf{f}(\hat{\mathbf{x}}_k, \hat{\mathbf{w}}_k), \ k = t-N, \ldots, t \tag{18a}$$

$$\bar{\mathbf{y}}_k = \mathbf{h}(\hat{\mathbf{x}}_k) + \hat{\bar{\mathbf{v}}}_k^i \tag{18b}$$

$$\hat{\mathbf{w}}_k \in \mathbb{W}^M \tag{18c}$$

$$\hat{\mathbf{x}}_k \in \mathbb{X}^M \tag{18d}$$

It is easy to show that solving the problem

$$\boldsymbol{\Theta}_t^* = \min_{\hat{\mathbf{x}}_{t-N}, \{\hat{\mathbf{w}}_k\}_{k=t-N}^t} \left\{ \mathbf{J}(t-N, t, \hat{\mathbf{x}}_{t-N}, \hat{\mathbf{w}}, \hat{\bar{\mathbf{v}}}, \boldsymbol{\Gamma}_{t-N}) \text{ subj. to (18)} \right\} \tag{19}$$

is equivalent to solve the *MHE* $-i$ problems (9), in the sense that $\hat{x}_{t-N/t}^i, \{\hat{w}_{k/t}^i\}_{k=t-N}^t$ is a solution to (9) if and only if $\hat{\mathbf{x}}_{t-N/t}, \{\hat{\mathbf{w}}_{k/t}\}_{k=t-N}^t$ is a solution to (19), where $\hat{\mathbf{w}}_{k/t} = (\hat{w}_{k/t}^1, \ldots, \hat{w}_{k/t}^M)$.

The *transit cost* of a generic state $\mathbf{x} \in \mathbb{R}^{nM}$ at time t_1, computed at instant t is defined as

$$\boldsymbol{\Xi}_{t_1/t}(\mathbf{x}) = \min_{\hat{\mathbf{x}}_{t-N}, \{\hat{\mathbf{w}}_k\}_{k=t-N}^t} \left\{ \mathbf{J}(t-N, t, \hat{\mathbf{x}}_{t-N}, \hat{\mathbf{w}}, \hat{\bar{\mathbf{v}}}, \boldsymbol{\Gamma}_{t-N}) \text{ subject to (18) and } \hat{\mathbf{x}}_{t_1} = \mathbf{x} \right\} \tag{20}$$

and it holds that

$$\boldsymbol{\Xi}_{t_1/t}(\hat{\mathbf{x}}_{t_1}) = \sum_{i=1}^M \Xi_{t_1/t}^i(\hat{x}_{t_1}^i) \tag{21}$$

In view of Assumption 2 and (16a), the cost function $\mathbf{L}$ is continuous, bounded, positive definite, and satisfies the following inequalities for all $\mathbf{w} \in \mathbb{R}^{M \cdot n}$, $\bar{\mathbf{v}} \in \mathbb{R}^{\sum_{i=1}^M \bar{p}_i}$

$$\underline{\boldsymbol{\gamma}}_L(\|(\bar{\mathbf{v}}, \mathbf{w})\|) \le \mathbf{L}(\bar{\mathbf{v}}, \mathbf{w}) \tag{22}$$

where $\underline{\boldsymbol{\gamma}}_L \in \mathcal{K}_\infty$.

Furthermore, the initial penalties $\Gamma_{t-N}^{C,i}$ and $\Gamma_{t-N}^{0,i}$ must be defined in order to fulfill the following collective condition.

Assumption 3. There exists $\underline{\boldsymbol{\gamma}}_0 \in \mathcal{K}_\infty$ such that the following inequalities are verified

$$\underline{\boldsymbol{\gamma}}_0(\|\mathbf{x} - \mathbf{K}\hat{\mathbf{x}}_{t-N/t-1}\|) \le \boldsymbol{\Gamma}_{t-N}^C(\mathbf{x}; \mathbf{K}\hat{\mathbf{x}}_{t-N/t-1}), \forall \mathbf{x} \in \mathbb{X}^M \tag{23a}$$

$$\underline{\boldsymbol{\gamma}}_0(\|\mathbf{x} - \hat{\mathbf{x}}_{t-N/t-1}\|) \le \boldsymbol{\Gamma}_{t-N}^0(\mathbf{x}; \hat{\mathbf{x}}_{t-N/t-1}), \forall \mathbf{x} \in \mathbb{X}^M \tag{23b}$$

$$\boldsymbol{\Gamma}_{t-N}(\mathbf{z}) \le \boldsymbol{\Xi}_{t-N/t-1}(\mathbf{z}), \forall \mathbf{z} = \mathbb{1}_M \otimes z, z \in \mathcal{Z} \tag{23c}$$

Assumption 3 is similar to Assumption 4.17 in Rawlings and Mayne (2009). However, there are two key differences. First, inequalities in (23c) must hold only for the consensus states $\mathbf{z}$. In particular, we highlight that if $\hat{\mathbf{x}}_{t-N/t-1} = \mathbf{z}$ then $\mathbf{K}\hat{\mathbf{x}}_{t-N/t-1} = \mathbf{z}$ and hence $\boldsymbol{\Theta}_{t-1}^*$ is a global lower bound to $\boldsymbol{\Gamma}_{t-N}$. Second, similarly to Farina et al. (2009a), as an upper bound to $\boldsymbol{\Gamma}_{t-N}$ we use the transit cost instead of the arrival cost (see Definition 4.16 in Rawlings and Mayne (2009)).

Note that, guaranteeing that (23) is verified is a challenging issue, which is still an open problem in the centralized Rawlings and Mayne (2009), as well as in the decentralied context. In the special case when (1), (3) is a linear system, if the stage and initial penalty cost functions are quadratic, as shown in (Farina et al., 2009a, 2010a), it is possible to provide recursive distributed equations for updating the penalty weighting matrices $\Pi_{t-N}^{C,i}$ and $\Pi_{t-N}^{0,i}$ in order to satisfy Assumption 3, and conditions to guarantee that these matrices remain bounded, in such a way that Assumption 2 is not violated. In the nonlinear context, empirical solutions can be either to compute $\Pi_{t-N}^{C,i}$ and $\Pi_{t-N}^{0,i}$ on the basis of quadratic local approximations of the transit costs or to assign constant values to $\Pi_{t-N}^{C,i}$ and $\Pi_{t-N}^{0,i}$.

4. COLLECTIVE STABILITY PROPERTIES OF NDMHE

The main purpose of this section is to extend the stability results of Rawlings and Mayne (2009) for centralized MHE to the proposed NDMHE scheme.

Definition 3. Let Σ be system (1) with $w = 0$ and denote by $x_\Sigma(t, x_0)$ the state reached by Σ at time t starting from initial condition x_0. Assume that the trajectory $x_\Sigma(t, x_0)$ is feasible, *i.e.*, $x_\Sigma(t, x_0) \in \mathbb{X}$ for all t. Define also the collective vectors $\mathbf{x}_0 = \mathbb{1}_M \otimes x_0$ and $\mathbf{x}_\Sigma(t, x_0) = \mathbb{1}_M \otimes x_\Sigma(t, x_0) \in \mathbb{X}^M$. NDMHE is *collectively stable* if, for all $\varepsilon > 0$, there exists $\delta > 0$ such that $\|\hat{\mathbf{x}}_{0/N-1} - \mathbf{x}_0\| < \delta$ implies that $\|\hat{\mathbf{x}}_{t-N/t} - \mathbf{x}_\Sigma(t-N, x_0)\| < \varepsilon$ $\forall t \ge N$. Also, NDMHE is *collectively asymptotically stable* if it is stable and asymptotically convergent, *i.e.*

$$\|\hat{\mathbf{x}}_{t-N/t} - \mathbf{x}_\Sigma(t-N, x_0)\| \xrightarrow{t \to \infty} 0 \tag{24}$$

$\square$

Notice that the condition (24) is equivalent to *individual* convergence for all the nodes estimates, *i.e.*

$$\|\hat{x}_{t-N/t}^i - x_\Sigma(t-N, x_0)\| \xrightarrow{t \to \infty} 0 \tag{25}$$

for all $i \in \mathcal{V}$.

Moreover, as in Rao et al. (2001), convergence is defined assuming that the model generating the data is noiseless, but the possible presence of noise is taken into account in the state estimation algorithm.

Before to state our main result, we need to introduce the following dynamical system, describing the dynamics of the variable $\boldsymbol{\eta}_t \in \mathbb{R}^{n \cdot M}$

$$\boldsymbol{\eta}_t = \mathbf{P}_{UD}\mathbf{T}^{-1}\left\{ \mathbf{K}\left[\mathbf{f}\left(\mathbf{T}\left(\boldsymbol{\eta}_{t-1} + \mathbf{P}_D\boldsymbol{\xi}_\Sigma(t-N-1, x_0) + \boldsymbol{\alpha}_{t-1}^\xi\right), 0\right) + \boldsymbol{\alpha}_t^w\right] + \boldsymbol{\alpha}_t^C \right\} \tag{26}$$

where $\boldsymbol{\xi}_\Sigma(t,x_0) = \mathbf{T}^{-1}(\mathbf{x}_\Sigma(t,x_0))$, and $\boldsymbol{\alpha}_t^w$, $\boldsymbol{\alpha}_t^\xi$, $\boldsymbol{\alpha}_t^C$ and $\boldsymbol{\xi}_\Sigma(t - N - 1, x_0)$ are input terms. In the following we resort to the definition of incremental input-to-state stability, (Angeli, 2002).

Definition 4. System (26) is incrementally input-to-state stable (δISS) with respect to the input triplet $(\boldsymbol{\alpha}_t^\xi, \boldsymbol{\alpha}_t^w, \boldsymbol{\alpha}_t^C)$, if there exist $\beta \in \mathcal{KL}$, $\sigma_\alpha \in \mathcal{K}_\infty$ such that, for any $t \geq 0$, any pair of initial conditions $\boldsymbol{\eta}_{j,0}$, $j = 1,2$ and any pair of input triplets $(\boldsymbol{\alpha}_{j,t}^\xi, \boldsymbol{\alpha}_{j,t}^w, \boldsymbol{\alpha}_{j,t}^C)$, $j = 1,2$ one has

$$\begin{aligned}
\|\boldsymbol{\eta}_{1,t} - \boldsymbol{\eta}_{2,t}\| &\leq \beta(\|\boldsymbol{\eta}_{1,0} - \boldsymbol{\eta}_{2,0}\|, t) + \\
&+ \sigma_\alpha(\|(\boldsymbol{\alpha}_{1,k}^\xi, \boldsymbol{\alpha}_{1,k}^w, \boldsymbol{\alpha}_{1,k}^C) - (\boldsymbol{\alpha}_{2,k}^\xi, \boldsymbol{\alpha}_{2,k}^w, \boldsymbol{\alpha}_{2,k}^C)\|_{[0:t]})
\end{aligned} \quad (27)$$

We are now in the position to state the main result.

Theorem 1. Under Assumptions 1, 2, 3, if the system (26) is δISS with respect to the input triplet $(\boldsymbol{\alpha}_t^\xi, \boldsymbol{\alpha}_t^w, \boldsymbol{\alpha}_t^C)$ then NDMHE is collectively asymptotically stable.

Note that, if the system is regionally observable by any sensor, $\bar{P}_{ND}^i = \mathbf{0}_{n\times n}$ for all $i \in \mathcal{V}$, and hence $\mathbf{P}_{UD} = \mathbf{0}_{nM\times nM}$. Therefore, from equation (26) one has that $\boldsymbol{\eta}_t = 0$, and the δISS condition required by Theorem 1 is trivially satisfied.

As shown in the proof of Lemma 3 in the Appendix, the dynamics (26) governs the regionally undetectable components of the state estimates in the $\boldsymbol{\xi}$-coordinates. Therefore δISS of (26) implies that the regionally undetectable components of the estimation error vanish when $(\boldsymbol{\alpha}_t^w, \boldsymbol{\alpha}_t^\xi, \boldsymbol{\alpha}_t^C)$ tends to zero. For linear constrained systems δISS of (26) is implied by the much simpler condition that a suitably defined matrix (matrix Φ in (23) of Farina et al. (2009b)) is Schur. In the nonlinear context, the system (26) can be viewed as the interconnection of M dynamically coupled subsystems. Current research is focusing on the application of the small gain theorem for networks (Dashkovskiy et al., 2007) for guaranteeing δISS of (26) on the basis of the δISS properties of individual sensors and suitable conditions on their interconnections, see Farina et al. (2010b), where an example of application of DMHE to a nonlinear case study is also provided.

5. CONCLUSIONS

In this paper we have proposed a distributed state estimation scheme for nonlinear constrained systems, based on Moving Horizon Estimators. Convergence to zero and stability of the state estimation error provided by any sensor can be guaranteed, provided that suitable technical assumptions are verified.
While in the linear setting these assumptions translate in fairly simple conditions on the observed system dynamics and on the network topology, in the nonlinear framework they are quite difficult to verify. For this reason, further work will be devoted by the authors to their analysis, in order to provide simpler and more tractable conditions allowing to check the validity of these key assumptions in a general setting.

Appendix A. PROOF OF THEOREM 1

The following lemmas are needed for the proof of Theorem 1.
Lemma 2. If Assumption 3 holds then

$$\sum_{k=t-N}^{t} \mathbf{L}(\hat{\mathbf{v}}_{k/t}, \hat{\mathbf{w}}_{k/t}) \xrightarrow{t\to\infty} 0 \quad (A.1a)$$

$$\underline{\boldsymbol{\gamma}}_0(\hat{\mathbf{x}}_{t-N/t} - \mathbf{K}\hat{\mathbf{x}}_{t-N/t-1}) \xrightarrow{t\to\infty} 0 \quad (A.1b)$$

$$\underline{\boldsymbol{\gamma}}_0(\hat{\mathbf{x}}_{t-N/t} - \hat{\mathbf{x}}_{t-N/t-1}) \xrightarrow{t\to\infty} 0 \quad (A.1c)$$

Lemma 3. Under Assumptions 1, 2 and 3, if the system (26) is δISS with respect to the input triplet $(\boldsymbol{\alpha}_t^\xi, \boldsymbol{\alpha}_t^w, \boldsymbol{\alpha}_t^C)$ then NDMHE is asymptotically convergent.
Lemma 4. Under the assumptions of Lemma 3, then NDMHE is collectively stable.

For lack of space, here we provide the proof of Lemma 3, while the proofs of Lemmas 2 and 4 can be found in Farina et al. (2010b).

Proof of Lemma 3
We create, for each sensor node i, a single estimate sequence by concatenating MHE sequences for the equivalent system (7). This gives the state sequences $\bar{\boldsymbol{\xi}}_k^i$ and the corresponding augmented disturbance sequences $\bar{w}_k^i = (\bar{w}_k^{i,1}, \bar{w}_k^{i,2})$

$$\bar{\xi}_{t+1}^i = f_\xi(\bar{\xi}_t^i, \bar{w}_t^{i,1}) + \bar{w}_t^{i,2} \quad (A.2a)$$

$$\bar{y}_t^i = \tilde{h}^i(\bar{\xi}_t^i) + \hat{v}_{t/t+N}^i \quad (A.2b)$$

where $\bar{\xi}_t^i = T_i^{-1}(\hat{x}_{t/t+N}^i)$ and

$$\bar{w}_t^{i,1} = \hat{w}_{t/t+N}^i \quad (A.3a)$$

$$\bar{w}_t^{i,2} = T_i^{-1}(\hat{x}_{t+1/t+N+1}^i) - T_i^{-1}(\hat{x}_{t+1/t+N}^i) \quad (A.3b)$$

According to the introduced notation used in (7), we denote $\bar{\xi}_t^i = (\bar{\xi}_t^{UD,i}, \bar{\xi}_t^{D,i})$. Define a sequence

$$\alpha_t^{\xi,i} = \bar{P}_D^i\left(\bar{\xi}_{t-N}^i - T_i^{-1}(x_\Sigma(t-N,x_0))\right) \quad (A.4)$$

In view of (7) and Definition 1, one has

$$\begin{aligned}
\|\alpha_t^{\xi,i}\| &\leq \beta_D(\|\bar{\xi}_0^i - T_i^{-1}(x_0)\|, t-N) + \\
&+ \gamma_1(\|\bar{w}_k^i\|_{[0:t-N-1]}) + \gamma_2(\|\hat{v}_{k/k+N}^i\|_{[0:t-N]})
\end{aligned} \quad (A.5)$$

This, according to (A.3), implies that there exist functions $\gamma_{11}, \gamma_{12} \in \mathcal{K}$ such that

$$\begin{aligned}
\|\alpha_t^{\xi,i}\| &\leq \beta_D(\|\bar{\xi}_0^i - T_i^{-1}(x_0)\|, t-N) + \\
&+ \gamma_{11}(\|\hat{w}_{k/k+N}^i\|_{[0:t-N-1]}) + \gamma_2(\|\hat{v}_{k/k+N}^i\|_{[0:t-N]}) + \\
&+ \gamma_{12}(\|T_i^{-1}(\hat{x}_{k+1/k+N+1}^i) - T_i^{-1}(\hat{x}_{k+1/k+N}^i)\|_{[0:t-N-1]})
\end{aligned}$$

We define $\xi_\Sigma^i(t,x_0) = T_i^{-1}(x_\Sigma(t,x_0))$, $\boldsymbol{\xi}_\Sigma(t,x_0) = (\xi_\Sigma^1(t,x_0), \ldots, \xi_\Sigma^M(t,x_0))$, $\bar{\boldsymbol{\xi}}_t = (\bar{\xi}_t^1, \ldots, \bar{\xi}_t^M)$ and $\boldsymbol{\alpha}_t^\xi = (\alpha_t^{\xi,1}, \ldots, \alpha_t^{\xi,M})$. Collectively (A.4) results in

$$\mathbf{P}_D\left(\bar{\boldsymbol{\xi}}_{t-N} - \boldsymbol{\xi}_\Sigma(t-N,x_0)\right) = \boldsymbol{\alpha}_t^\xi \quad (A.6)$$

Furthermore, applying the mean value theorem for vector functions (see Appendix A in Rawlings and Mayne (2009)), we can write

$$\hat{x}_{k+1/t}^i = f(\hat{x}_{k/t}^i, 0) + \alpha_{k/t}^{w,i} \quad (A.7)$$

where

$$\|\alpha_{k/t}^{w,i}\| \leq l\|\hat{w}_{k/t}^i\| \quad (A.8)$$

l being the Lipschitz constant in (2). We define $\boldsymbol{\alpha}_t^w = (\alpha_{t-N-1/t-1}^{w,1}, \ldots, \alpha_{t-N-1/t-1}^{w,M})$ and $\mathbf{f}(\hat{\mathbf{x}}_k) = (f(\hat{x}_k^1), \ldots, f(\hat{x}_k^M))$. Collectively we write (A.7) as

$$\hat{\mathbf{x}}_{t-N/t-1} = \mathbf{f}(\hat{\mathbf{x}}_{t-N-1/t-1}, 0) + \boldsymbol{\alpha}_t^w \quad (A.9)$$

From Lemma 2 (A.1a) holds and together with (22) one has that $\|(\hat{\mathbf{w}}_{t-N/t}, \hat{\bar{\mathbf{v}}}_{t-N/t})\| \to 0$ as $t \to +\infty$. Hence, $\|(\hat{w}^i_{t-N/t}, \hat{\bar{v}}^i_{t-N/t})\| \to 0$ as $t \to +\infty$ for all $i \in \mathscr{V}$. Similarly, from (A.1c), $\|\hat{x}^i_{t+1/t+N+1} - \hat{x}^i_{t+1/t+N}\| \to 0$ as $t \to +\infty$ for all $i \in \mathscr{V}$. In view of (A.6) and Proposition 4.2 in Rawlings and Mayne (2009) (convergence of the state under δIOSS), this implies that $\alpha^{x,i}_t \to 0$ as $t \to +\infty$, i.e.

$$\bar{P}^i_D\left(\bar{\xi}^i_{t-N} - \xi^i_\Sigma(t-N, x_0)\right) \to 0 \,\forall i \in \mathscr{V} \qquad \text{(A.10)}$$

Moreover

$$\boldsymbol{\alpha}^w_t \xrightarrow{t \to +\infty} 0 \qquad \text{(A.11)}$$

Finally, from (A.1b) and (23), we obtain that

$$\hat{\mathbf{x}}_{t-N/t} = \mathbf{K}\hat{\mathbf{x}}_{t-N/t-1} + \boldsymbol{\alpha}^C_t \qquad \text{(A.12)}$$

where $\boldsymbol{\alpha}^C_t \to 0$ as $t \to +\infty$.

According to (A.9) and (A.12) the term $\mathbf{P}_{UD}\bar{\boldsymbol{\xi}}_{t-N}$ can be written as

$$\begin{aligned}
\mathbf{P}_{UD}\bar{\boldsymbol{\xi}}_{t-N} &= \mathbf{P}_{UD}\mathbf{T}^{-1}\left(\mathbf{K}\hat{\mathbf{x}}_{t-N/t-1} + \boldsymbol{\alpha}^C_t\right)\\
&= \mathbf{P}_{UD}\mathbf{T}^{-1}\left[\mathbf{K}\left(\mathbf{f}(\hat{\mathbf{x}}_{t-N-1/t-1}, 0) + \boldsymbol{\alpha}^w_t\right) + \boldsymbol{\alpha}^C_t\right]
\end{aligned}$$

where, using (A.6), we can write

$$\hat{\mathbf{x}}_{t-N-1/t-1} = \mathbf{T}\left(\mathbf{P}_D\boldsymbol{\xi}_\Sigma(t-N-1, x_0) + \mathbf{P}_{UD}\bar{\boldsymbol{\xi}}_{t-N-1} + \boldsymbol{\alpha}^\xi_{t-1}\right)$$

Hence, we obtain that the dynamics of $\mathbf{P}_{UD}\bar{\boldsymbol{\xi}}_{t-N}$ evolves according to (26). Analogously we obtain that the dynamics of variable $\mathbf{P}_{UD}\boldsymbol{\xi}_\Sigma(t-N, x_0)$ is given by (26), with $\boldsymbol{\alpha}^w_t, \boldsymbol{\alpha}^\xi_t, \boldsymbol{\alpha}^C_t = 0$. We define $\hat{\boldsymbol{\eta}}_t = \mathbf{P}_{UD}\bar{\boldsymbol{\xi}}_{t-N}$ and $\boldsymbol{\eta}_\Sigma(t, x_0) = \mathbf{P}_{UD}\boldsymbol{\xi}_\Sigma(t-N, x_0)$. We introduce the function $\mathbf{F}$, so that the dynamical equations for $\hat{\boldsymbol{\eta}}_t$ and $\boldsymbol{\eta}_\Sigma(t, x_0)$ can be written as

$$\hat{\boldsymbol{\eta}}_t = \mathbf{F}(\hat{\boldsymbol{\eta}}_{t-1}, \mathbf{P}_D\boldsymbol{\xi}_\Sigma(t-N-1, x_0), \boldsymbol{\alpha}^\xi_{t-1}, \boldsymbol{\alpha}^w_t, \boldsymbol{\alpha}^C_t) \qquad \text{(A.13a)}$$

$$\boldsymbol{\eta}_\Sigma(t, x_0) = \mathbf{F}(\boldsymbol{\eta}_\Sigma(t-1, x_0), \mathbf{P}_D\boldsymbol{\xi}_\Sigma(t-N-1, x_0), 0, 0, 0) \qquad \text{(A.13b)}$$

According to Definition 4, if the system (A.13a) is δISS, then there exist $\beta \in \mathscr{KL}$, $\sigma_\alpha \in \mathscr{K}_\infty$ such that

$$\begin{aligned}
\|\hat{\boldsymbol{\eta}}_t - \boldsymbol{\eta}_\Sigma(t, x_0)\| \leq &\beta(\|\hat{\boldsymbol{\eta}}_0 - \boldsymbol{\eta}_\Sigma(0, x_0)\|, t) +\\
&+ \sigma_\alpha(\|(\boldsymbol{\alpha}^\xi_k, \boldsymbol{\alpha}^w_k, \boldsymbol{\alpha}^C_k)\|_{[0:t]}) \qquad \text{(A.14)}
\end{aligned}$$

If (A.14) holds then $\|\hat{\boldsymbol{\eta}}_t - \boldsymbol{\eta}_\Sigma(t, x_0)\| \to 0$ because $(\boldsymbol{\alpha}^\xi_t, \boldsymbol{\alpha}^w_t, \boldsymbol{\alpha}^C_t) \to 0$ as $t \to +\infty$. This, together with (A.11), implies that $\hat{\mathbf{x}}_{t-N/t} \to \hat{\mathbf{x}}_\Sigma(t-N, x_0)$ as $t \to +\infty$. ∎

REFERENCES

Alessandri, A., Baglietto, M., and Battistelli, G. (2003). Receding-horizon estimation for discrete-time linear systems. *IEEE Trans. on Automatic Control*, 48(3), 473 – 478.

Alessandri, A., Baglietto, M., and Battistelli, G. (2008). Moving-horizon state estimation for nonlinear discrete-time systems: new stability results and approximation schemes. *Automatica*, 44, 1753 – 1765.

Alessandri, A., Baglietto, M., Parisini, T., and Zoppoli, R. (1999). A neural state estimator with bounded errors for nonlinear systems. *IEEE Trans. on Automatic Control*, 44(11), 2028 – 2042.

Alriksson, P. and Rantzer, A. (2006). Distributed Kalman filtering using weighted averaging. In *Proc. of the 17th International Symposium on Mathematical Theory of Networks and Systems*. Kyoto, Japan.

Angeli, D. (2002). A Lyapunov approach to incremental stability properties. *IEEE Trans. on Automatic Control*, 47, 410 – 421.

Carli, R., Chiuso, A., Schenato, L., and Zampieri, S. (2008). Distributed Kalman filtering based on consensus strategies. *IEEE Journal on Selected Areas In Communications*, (4), 622 – 633.

Dashkovskiy, S., Rüffer, B.S., and Wirth, F. (2007). An ISS small gain theorem for iss general networks. *Mathematics of Control, Signals, and Systems*, 19(2), 93–122.

Farina, M., Ferrari-Trecate, G., and Scattolini, R. (2009a). Distributed moving horizon estimation for sensor networks. In *Proc. 1st IFAC Workshop on Estimation and Control of Networked Systems, Venice, Italy*, 126 – 131.

Farina, M., Ferrari-Trecate, G., and Scattolini, R. (2009b). A moving horizon scheme for distributed state estimation. In *48th IEEE Conference on Decision and Control and 28th Chinese Control Conference, Shanghai, China*, 1818 – 1823.

Farina, M., Ferrari-Trecate, G., and Scattolini, R. (2010a). Distributed moving horizon estimation for linear constrained systems. *IEEE Trans. on Automatic Control*. In press.

Farina, M., Ferrari-Trecate, G., and Scattolini, R. (2010b). Distributed moving horizon estimation for nonlinear constrained systems. Technical Report 2010.7, Dipartimento di Elettronica e Informazione, Politecnico di Milano. Available at http://home.dei.polimi.it/farina/NDMHE2010.7.pdf.

Ferrari-Trecate, G., Mignone, D., and Morari, M. (2002). Moving horizon estimation for hybrid systems. *IEEE Trans. on Automatic Control*, 47(10), 1663 – 1676.

Goodwin, G., Seron, M., and De Doná, J. (2005). *Constrained Control and Estimation*. Springer, New Jersey.

Kamgarpour, M. and Tomlin, C. (2008). Convergence properties of a decentralized Kalman filter. *Proc. 47th IEEE Conference on Decision and Control*, 3205 – 3210.

Olfati-Saber, R. (2005). Distributed Kalman filter with embedded consensus filters. *Proc. 44th IEEE Conference on Decision and Control - European Control Conference*, 8179 – 8184.

Olfati-Saber, R. (2007). Distributed Kalman filtering for sensor networks. *Proc. 46th IEEE Conference on Decision and Control*, 5492 – 5498.

Olfati-Saber, R. (2009). Kalman-consensus filter: Optimality, stability and performance. *Proc. 48th IEEE Conference on Decision and Control*, 7036 – 7042.

Rao, C.V. and Rawlings, J. (2000). Nonlinear moving horizon state estimation. *in F. Allgöwer and A. Zheng, editors, Nonlinear Model Predictive Control, Progress in Systems and Control Theory, Birkhauser*, 45–70.

Rao, C.V., Rawlings, J., and Lee, J. (1999). Stability of constrained linear moving horizon estimation. *Proc. American Control Conference*, 3387 – 3391.

Rao, C.V., Rawlings, J., and Lee, J. (2001). Constrained linear state estimation - a moving horizon approach. *Automatica*, 37, 1619 – 1628.

Rao, C.V., Rawlings, J., and Mayne, D. (2003). Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE Trans. on Automatic Control*, 48(2), 246 – 258.

Rawlings, J.B. and Mayne, D.Q. (2009). *Model predictive control: theory and design*. Nob Hill Publishing, LLC.