High gain observer based on a triangular structure

Hassan Hammouri, Boubekeur Targui, F. Armanet

1
Université de Lyon, Lyon, F-69003, France; Université Lyon 1;
CNRS UMR 5007 LAGEP (Laboratoire d'Automatique et de GEnie des Procédés),
43 bd du 11 novembre, 69100 Villeurbanne, France
Tel +33 (0) 4 72 43 18 45 - Fax +33 (0) 4 72 43 16 99
http://www-lagep.univ-lyon1.fr/ http://www.univ-lyon1.fr http://www.cnrs.fr

HIGH GAIN OBSERVER BASED ON A TRIANGULAR STRUCTURE

HASSAN HAMMOURI, BOUBEKEUR TARGUI,
AND FREDERIC ARMANET
Laboratoire d'Automatique et de Génie des Procédés, UPRES-A Q 5007, LAGEP, CPE-Lyon, Bat. 308G, UCBL I, 69622 Villeurbanne cedex, France.

SUMMARY
A high gain observer based on a triangular structure of nonlinear systems is proposed . An algorithm permitting to calculate a gain of the observer is given. This observer synthesis is then extended to a class of multi-output nonlinear systems which contains the model of binary distillation columns. Finally, we illustrate the performance of the estimator using numerical simulations of a methanol-ethanol distillation column.

KEY WORDS Nonlinear systems, observability, observer

1. INTRODUCTION

The knowledge of state variables is often required in order to apply the advanced concepts of control and diagnosis to practical applications. One way permitting to obtain such variables, consists of combining a priori knowledge about physical system with experimental data to provide on-line estimator (observer). This estimator is generally a dynamic system obtained from the nominal model by adding a correction term which is proportional to some output deviation. In other words, given a nominal model:

$$\begin{cases} \dot{x}(t) = f(x(t), u(t)) \\ y(t) = h(x(t)) \end{cases} \tag{0}$$

The state $x(t)$ belongs to an open subset V of $\mathcal{R}^n$, the input $u(t)$ belongs to a Borelian subset U of $\mathcal{R}^m$ and the output $y(t) \in \mathcal{R}^p$. An observer for (1) is generally a dynamic system of the form:

$$\begin{cases} \dot{\hat{x}}(t) = f(\hat{x}(t), u(t)) - k(t)(h(\hat{x}(t)) - y(t)) \\ \dot{r}(t) = F(r(t), u(t), y(t), \hat{x}(t)) \\ k(t) = \varphi(r(t)) \end{cases}$$

$r(t)$ and $k(t)$ are called indifferently the gain of the observer. For some particular systems, the gain $k(t)$ does not depend on the input (see for instance the Luenberger observer). For general nonlinear systems the observer's gain depend on the input. This comes from the fact that the observability concept generally depends on the inputs: given an input u defined on some interval $[0, T]$, we say that u render system (1) **observable** if for every two initial states $x \neq x'$; there exist $t \in [0, T]$ such that $h(x_u(t)) \neq h(x'_u(t))$, where $x_u(t)$ and $x'_u(t)$ are respectively the trajectories associated with u and issued from the initial states x, x'. Generally a nonlinear system may be observable for some input and unobservable for an other one. For more details, see [7,9].

An interesting class of nonlinear systems consists of those which are observable for every input, called **uniformly observable** systems. For this class of nonlinear systems, we can design an observer whose gain does not depend on the inputs (see, [1,2,3,5,6,8,10]). For such systems a canonical (triangular) form is designed in order to design an observer. To ensure the mathematical convergence, a particular high gain is required. However, the use of large gain may generate the so-called the peak phenomena (overshoot problem), moreover the estimator becomes noise sensitive. Due to nonlinearity of the system, the choice of the gain which gives the best compromise between fast convergence, the noise rejection and the attenuation of the peak phenomena becomes a difficult task, and only simulations allow to determine a plausible gain.

This paper is organized as follows: in section 2, we extend the observer synthesis stated in [5] and [6] to a class of multi-output uniformly observable systems. In section 3, we apply this theoretical result to a binary distillation column.

2. HIGH GAIN OBSERVER

Consider the following control affine nonlinear system:

$$\begin{cases} \dot{x}(t) = f(x(t), u(t)) = f_0(x(t)) + \sum_{i=1}^{m} u_i(t) f_i(x(t)) \\ y(t) = h(x(t)) \end{cases} \quad (0)$$

where $u = (u_1, ..., u_m)$.
In the single output case ($p = 1$) and when $U = \mathcal{R}^m$, the authors in [4] and [5] have shown that if in addition system (2) is uniformly observable, then,

$$(z_1, ..., z_n) = (h(x), L_{f_0}(h(x)), ..., L_{f_0}^{n-1}(h(x)))$$

becomes a local system of coordinates (almost everywhere) in which, system (2) takes the following canonical form:

$$\begin{cases} \dot{z}(t) = Az(t) + F_0(z(t)) + \sum_{i=1}^{m} u_i(t) F_i(z(t)) \\ y(t) = Cz(t) \end{cases} \quad (0)$$

where L_{f_0} denotes the Lie derivative, A, C and the F_j's are given by:

$$A = \begin{bmatrix} 0 & 1 & 0 & & 0 \\ \vdots & & 1 & & \\ 0 & & & \ddots & 1 \\ 0 & \dots & & 0 & 0 \end{bmatrix}; C = \begin{bmatrix} 1 & 0 & . & . & 0 \end{bmatrix}$$

$F_0 = [0, \dots, 0, F_{0n}]^T$; for $1 \leq i \leq m$, $F_i = [F_{i1}, \dots, F_{in}]^T$ and $F_{ij} = F_{ij}(z_1, .., z_j)$.
Moreover, if the F_i's are global Lipschitz (or if the concerned trajectories of system (3) are bounded), then an exponential observer for system (3) takes the following form:

$$\dot{\hat{z}}(t) = A\hat{z}(t) + F_0(\hat{z}(t)) + \sum_{i=1}^{m} u_i(t) F_i(\hat{z}(t)) - S_\theta^{-1} C^T (C\hat{z}(t) - y(t))$$

where S_θ is the symmetric positive definite (S.P.D.) matrix satisfying:

$$\theta S_\theta + A^T S_\theta + S_\theta A = C^T C.$$

For single output non control affine nonlinear system (1), in [6], the authors have shown that if a single output system (1) is uniformly observable then, a similar transformation as above, transforms the system into the following form:

$$\begin{cases} \dot{z}_1(t) = F_1(z_1(t), z_2(t), u(t)) \\ \dot{z}_2(t) = F_2(z_1(t), z_2(t), z_3(t), u(t)) \\ \vdots \\ \dot{z}_{n-1}(t) = F_{n-1}(z_1(t), \ldots, z_n(t), u(t)) \\ \dot{z}_n(t) = F_n(z_1(t), \ldots, z_n(t), u(t)) \\ y(t) = Cz(t) = z_1(t) \end{cases} \tag{0}$$

with, the additional condition:

C1) $$\frac{\partial F_i}{\partial z_{i+1}}(z,u) \neq 0; \forall (z,u).$$

Now set $F = [F_1, \ldots, F_n]^T$, if the following assumption holds:
H_1)
i) The F is a global Lipschitz function:

$$\|\frac{\partial F(u,z)}{\partial z}(u,z)\| \text{ is uniformly bounded.}$$

ii) $\exists \alpha > 0$ $s.t.$ $\forall (u,z) \in (U \times R^n)$ we have:

$$\frac{\partial F_i(u,z)}{\partial z_{i+1}} \geq \alpha$$

then, we can find a constant vector K, such that the following system:

$$\dot{\hat{z}} = F(\hat{z}, u) + \Delta_\theta K(C\hat{z} - y) \tag{0}$$

becomes an exponential observer for system (4), where:

Δ_θ is the $(n \times n)$ diagonal matrix:

$$\Delta_\theta = diag(\theta, \theta^2, \ldots, \theta^n).$$

Remark 1

In the case where only condition $C1$ holds and that U is a bounded connected set, and the trajectories of (4) lie into a compact set, then hypothesis H_1 can be omited (see Remark 2 below).

In what follows, we give a constructive algorithm permitting to calculate such gain K, and we extend this observer synthesis to a class of multi-output systems.
To do so, we need some preliminary results:

Consider the following $k \times k$ matrix,

$$A_k(t) = \begin{bmatrix} 0 & a_1(t) & 0 & & 0 \\ \vdots & & a_2(t) & & \\ 0 & & & \ddots & a_{k-1}(t) \\ 0 & \ldots & & 0 & 0 \end{bmatrix} \tag{0}$$

where, the a_i's may be unknown and satisfying the following constraint:

H_2) $\forall t \geq 0, \alpha_1 \leq a_i(t) \leq \alpha_2$, for some constants $\alpha_1, \alpha_2 > 0$.

Let S_k be the $k \times k$ symmetric matrix of the form:

$$S_k = \begin{bmatrix} s_{11} & s_{12} & 0 & & 0 \\ s_{12} & s_{22} & \ddots & & \vdots \\ 0 & \ddots & & \ddots & 0 \\ \vdots & & & \ddots & s_{k-1k} \\ 0 & \dots & 0 & s_{k-1k} & s_{kk} \end{bmatrix} \qquad (0)$$

and denote by C_k the k-row vector:

$$C_k = [1, 0, .., 0]$$

we then obtain the following:

Lemma 1

Assume that H_2 holds, then for every $\rho > 0$, we can find $\eta > 0$ and a symmetric positive definite (S.P.D.) matrix S_k of the form (6) such that:

$$\forall t \geq 0, A_k^T(t)S_k + S_k A_k(t) - \rho C_k^T C_k \leq -\eta I_k.$$

Moreover S_k depends only on the bounds α_1, α_2 and not on the knowledge of $A_k(t)$. The proof of lemma 1 will be given below.

Now, let $A_n(t)$ be the $n \times n$ matrix of the form (6) in which $a_i(t)$ is replaced by $\dfrac{\partial F_i}{\partial z_{i+1}}(\hat{z}(t) + \omega(t), u(t))$, where $\omega(t)$ is any vector of R^n. Then from hypothesis H_1, it follows that the a_i's satisfy H_2, here, $\alpha_1 = \alpha$ and α_2 is the Lipschitz constant of F, given by $sup\|\dfrac{\partial F}{\partial z}(z(t), u(t))\|$, $(z, u) \in R^n \times U$. We can state the following:

Theorem 1.

Assume that H_1 holds, then system (5) in which K is replaced by $\Delta_\theta S_n^{-1} C^T$ is an exponential observer.

The proof is similar to this given in [6] (the extension of this result, is given in Theorem 2 below).

In what follows, we will extend this observer synthesis to the following class of nonlinear systems which contains the model of binary distillation columns:

$$\begin{cases} \dot{x}^1(t) = f^1(x(t), u(t)) + d^1(t) \\ \dot{x}^2(t) = f^2(x(t), u(t)) + d^2(t) \\ y(t) = (y_1(t), y_2(t))^T = (C_{n_1} x^1(t), C_{n_2} x^2(t))^T \end{cases} \qquad (0)$$

where $x = \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} \in \mathcal{R}^n$; $x^i = \begin{bmatrix} x^i_1 \\ \vdots \\ x^i_{n_i} \end{bmatrix} \in \mathcal{R}^{n_i}$ for $i = 1, 2$ $(n = n_1 + n_2)$; $y_i = C_{n_i} x^i = x^i_1$ $(C_{n_i} = [1, 0, \ldots, 0])$; u is an known signal such that $\forall t$, $u(t) \in U$, the d^i's are unknown and bounded disturbances, with $sup_{t\geq 0} \left\| d^i(t) \right\| = d < +\infty$.
Finally, the nonlinear dynamics satisfy the following triangular structure:

$$f^1(x,u) = \begin{bmatrix} f^1_1(x^1_1, x^1_2, u) \\ f^1_2(x^1_1, x^1_2, x^1_3, u) \\ \vdots \\ f^1_{n_1-1}(x^1, u) \\ f^1_{n_1}(x, u) \end{bmatrix}; \; f^2(x,u) = \begin{bmatrix} f^2_1(x^2_1, x^2_2, u) \\ f^2_2(x^2_1, x^2_2, x^2_3, u) \\ \vdots \\ f^2_{n_2-2}(x^2_1, \ldots, x^2_{n_2-1}, u) \\ f^2_{n_2-1}(x, u) \\ f^2_{n_2}(x, u) \end{bmatrix}$$

The following lemma gives a sufficient condition which guarantee the uniform observability.

Lemma 2

If for every $(x, u) \in R^n \times U$, $\frac{\partial f^i_j}{\partial x^i_{j+1}}(x, u) \neq 0$; then system (8) is uniformly observable.

Proof. The uniform observability means that for every initial states $x \neq \bar{x}$ and every input from any $[0, T]$ into U, the associated output $y(t) = h(x(t))$ and $\bar{y}(t) = h(\bar{x}(t))$ are not identically equal on $[0, T]$, where $x(t)$ (respectively $\bar{x}(t)$) is the trajectory corresponding to the input u and initial state x (resp. $\bar{x}$).
To prove the Lemma 2, it suffices to show that if for every $t \in [0, T], h(x(t)) = h(\bar{x}(t))$, then $x = \bar{x}$.
Here, $x = \begin{bmatrix} x^1 \\ x^2 \end{bmatrix}$ and $h(x) = \begin{bmatrix} x^1_1 \\ x^2_1 \end{bmatrix}$.
Assume that $h(x(t)) = h(\bar{x}(t)), \forall t \in [0, T]$.
We obtain

$$x^1_1(t) = \bar{x}^1_1(t), \forall t \in [0, T].$$

Differentiating this last equality, we get:

$$f^1_1(x^1_1(t), x^1_2(t), u(t)) = f^1_1(x^1_1(t), \bar{x}^1_2(t), u(t))$$

Using the mean value theorem, there exists a $\omega_1(t) \in [0, 1]$ s.t.

$$\frac{\partial f^1_1}{\partial x^1_2}(x^1_1(t), x^1_2(t) + \omega_1(t)(x^1_2(t) - \bar{x}^1_2(t)), u(t))(x^1_2(t) - \bar{x}^1_2(t)) = 0$$

But

$$\frac{\partial f^1_1}{\partial x^1_2}(x, u) \neq 0, \forall (x, u) \in R^n \times U.$$

Thus

$$x^1_2(t) = \bar{x}^1_2(t), \forall t \in [0, T].$$

Using the triangular structure of f^1 and the fact that $\frac{\partial f^1_i}{\partial x_{i+1}}(x, u) \neq 0; \forall (x, u) \in R^n \times U$, an induction proof gives $x^1(t) = \bar{x}^1(t), \forall t \in [0, T]$.

In similar manner, using the triangular structure of f^2 and the fact that $x^1(t) = \bar{x}^1(t)$, and $x_1^2(t) = \bar{x}_1^2(t)$ we can show by similar induction proof that $\forall t \geq 0, x^2(t) = \bar{x}^2(t)$. This ends the proof of the lemma 2.

As in the single output case (see hypothesis H_1), the design of an observer for (8), requires the following assumption :
H_3)
i) $\exists c > 0; \forall (x,u), \left\| \frac{\partial f^i}{\partial x}(x,u) \right\| \leq c$

ii) $\exists \alpha > 0; \forall x; \forall u, \frac{\partial f_j^i}{\partial x_{j+1}^i}(x,u) \geq \alpha$, for $i = 1, 2$ and for $1 \leq j \leq n_i - 1$.

Remark 2

If U is a connected compact subset of R^m and that the trajectories of system (8) lie into a connected compact subset K of R^n, then hypothesis H_3 can be replaced by:

C2)
$$\frac{\partial f^i}{\partial x_{j+1}^i}(x,u) \neq 0, \forall (x,u) \in (K \times U).$$

Proof of remark 2. Indeed, From $C2$, it follows that $\exists \alpha > 0; \forall (x,u) \in (K \times U)$ we have $|\frac{\partial f_j^i}{\partial x_{j+1}^i}(x,u)| \geq \alpha$.
Now, since $K \times U$ is a connected set, it follows that $\frac{\partial f_j^i}{\partial x_{j+1}^i}$ keeps a constant sign.
Using the simple change of coordinates $\tilde{x}_j^i = \varepsilon_{ij} x_j^i$ where $\varepsilon_{ij} = sign(\frac{\partial f_j^i}{\partial x_{j+1}^i})$, system (8) becomes:

$$\begin{cases} \dot{\tilde{x}}^1(t) = \tilde{f}^1(\tilde{x}(t), u(t)) + \tilde{d}^1(t) \\ \dot{\tilde{x}}^2(t) = \tilde{f}^2(\tilde{x}(t), u(t)) + \tilde{d}^2(t) \\ y(t) = (C_{n_1}\tilde{x}^1(t), C_{n_2}\tilde{x}^2(t))^T \end{cases}$$

Moreover, it has a similar triangular structure (8) and,

$$\frac{\partial \tilde{f}_j^i}{\partial \tilde{x}_{j+1}^i}(\tilde{x},u) \geq \alpha, \forall (\tilde{x},u) \in (\tilde{K} \times U)$$

where $\tilde{K} = \{\tilde{x} / \tilde{x}_j^i = \varepsilon_{ij} x_j^i \ and \ x \in K\}$.
Now as in [5], we can always find global Lipschitz functions having similar triangular structure as $\breve{f}^1, \breve{f}^2$ and satisfying:
a) $\breve{f}^i(\tilde{x},u) = \tilde{f}^i(\tilde{x},u)$, for $i = 1, 2$ and for every $(\tilde{x},u) \in (\tilde{K} \times U)$
b) $\frac{\partial \breve{f}_j^i}{\partial \tilde{x}_{j+1}^i}(\tilde{x},u) \geq \alpha, \forall (\tilde{x},u) \in (R^n \times U)$
This ends the proof of Remark 2.

Noticing that if U is not a connected set, then condition $H_3 - i)$ together with $C2$ are not sufficient for the existence of an observer of constant gain.

Conter-example

Consider the following system:

$$(S) \quad \begin{cases} \dot{x} = \begin{bmatrix} 0 & u \\ 0 & 0 \end{bmatrix} x \\ y(t) = x_1 = Cx \end{cases}$$

with $u \in U = \{+1, -1\}$.
Clearly system (S) has a triangular structure and satisfies $H_3 - i$), and $C2$.
However (S) cannot admit an observer with constant gain.

Proof of the Conter-example. Assuming that (S) admits an observer of constant gain:

$$\dot{\hat{x}} = \begin{bmatrix} 0 & u \\ 0 & 0 \end{bmatrix} \hat{x} + K(C\hat{x} - y)$$

with $K = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$.
Thus the obtained error equation:

$$\dot{e} = \begin{bmatrix} k_1 & u \\ k_2 & 0 \end{bmatrix} e$$

becomes simultaneous asymptotically stable for every input $u : R^+ \to \{1, -1\}$.
Now set $A_1 = \begin{bmatrix} k_1 & 1 \\ k_2 & 0 \end{bmatrix}$ and $A_2 = \begin{bmatrix} k_1 & -1 \\ k_2 & 0 \end{bmatrix}$, in particular the two systems:

$$\dot{e} = A_i e \quad (i = 1, 2)$$

becomes asymptotically stable.
This implies that the spectrums $Sp(A_i)$ are in $\mathcal{C} = \{\lambda \in \mathcal{C} : Re(\lambda) < 0\}$.
Let us examine this last fact:
The respective characteristic polynomial of A_1 and A_2 are $P_1 = \lambda^2 - k_1\lambda - k_2$ and $P_2 = \lambda^2 - k_1\lambda + k_2$.
Set r_{11} , r_{12} (resp. r_{21}, r_{22}) the roots of P_1 (resp. P_2), we obtain:

$$\begin{cases} r_{11} + r_{12} = -k_1 \\ r_{11} r_{12} = -k_2 \\ r_{21} + r_{22} = -k_1 \\ r_{21} r_{22} = k_2 \end{cases}$$

Thus the stability of A_1 (resp. of A_2) is equivalent to $k_1 < 0$ and $k_2 < 0$ (resp. $k_1 < 0$ and $k_2 > 0$). Consequently, there exist no k_1, k_2 which give rise to both simultaneous asymptotic stability of A_1 and A_2.

Now, we can state our main result:

Theorem 2.
Let $\delta_1 > 0$, $\delta_2 > 0$ be two constants satisfying:

$$\frac{2n_1 - 1}{2n_2 - 1}\delta_1 < \delta_2 < \frac{2n_1 + 1}{2n_2 - 1}\delta_1$$

set, $\Delta_{\theta^{\delta_i}} = diag(\theta^{\delta_i}, \theta^{2\delta_i}, \ldots, \theta^{n_i\delta_i}), (i = 1, 2)$, and assume that H_3 holds for system (8). Then there exist S.P.D. matrices S_{n_i} $(i = 1, 2)$ of the form (7) such that the following system:

$$\begin{cases} \dot{\hat{x}}^1(t) = f^1(\hat{x}(t), u(t)) - r_1\Delta_{\theta^{\delta_1}} S_{n1}^{-1} C_{n1}^T (C_{n1}\hat{x}(t) - y_1(t)) \\ \dot{\hat{x}}^2(t) = f^2(\hat{x}(t), u(t)) - r_2\Delta_{\theta^{\delta_2}} S_{n2}^{-1} C_{n2}^T (C_{n2}\hat{x}(t) - y_2(t)) \end{cases} \quad (0)$$

becomes an exponential estimator:
$\exists r_1, r_2 > 0; \exists \theta_0 > 0; \forall \theta \geq \theta_0; \exists \lambda_1, \lambda_2 > 0; \exists \sigma > 0$,such that for every $\hat{x}(0), x(0)$, we have: $\|\hat{x}(t) - x(t)\| \leq \lambda_1 e^{-\sigma t} \|\hat{x}(0) - x(0)\| + \lambda_2 d$, here d is the upper bound of $\|d^i(t)\|$, $i = 1, 2$. This means that the ball $B(0, \lambda_2 d)$ is exponentially asymptotically attracts the error of estimation.

Remark 3

Noticing that for θ large σ becomes large and hence the speed of convergence of the observer may be chosed arbitrary . However, for θ sufficiently large, λ_1 may become large and hence the disturbance may affect the performance of the estimator. Consequently, a good choice of θ which not affect the performances of the estimator is necessary. Due to nonlinearity, this problem becomes very difficult and only simulations allows to obtain a compromise between the fast convergence of the observer and the disturbance sensitivity. First, let us give the proof of Lemme 1.

*Proof of Lemma 1.*We will give the proof by induction.

For $k = 2$; consider $A_2(t) = \begin{bmatrix} 0 & a_1(t) \\ 0 & 0 \end{bmatrix}$, with $0 < \alpha_1 \leq a_1(t) \leq \alpha_2$, $\forall t \geq 0$.
Let $S_2 = \begin{bmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{bmatrix}$ be symmetric positive definite (S.P.D), and set;

$$P_2(t) = A_2^T(t)S_2 + S_2A_2(t) - \rho C_2^T C_2 = \begin{bmatrix} -\rho & s_{11}a_1(t) \\ s_{11}a_1(t) & 2s_{12}a_1(t) \end{bmatrix}, \quad (0)$$

for some $\rho > 0$ and where $X_2 = [x_1, x_2]^T$.
We have:

$$X_2^T P_2(t) X_2 = -\rho x_1^2 + 2s_{11}a_1(t)x_1x_2 + 2s_{12}a_1(t)x_2^2$$

Now choosing $s_{11} > 0$, $s_{22} > 0$ and $s_{12} < 0$, and using inequality $0 < \alpha_1 \leq a_1(t) \leq \alpha_2$, we obtain:

$$X_2^T P_2(t) X_2 \leq -\rho x_1^2 + 2s_{11}\alpha_2 x_1x_2 - 2s_{12}\alpha_1 x_2^2 \quad (0)$$

Now, choosing $s_{12} < 0$ and $s_{22} > 0$ such that:

$$|s_{12}| > \frac{s_{11}^2\alpha_2^2}{2\rho\alpha_1} \quad (0)$$

$$s_{22} > \frac{s_{12}^2}{s_{11}} \quad (0)$$

Inequality (12), implies that:

$$X_2^T P_2(t) X_2 \leq -\eta \|X_2\|^2, \text{ where } \eta = (1 - \frac{2s_{11}\alpha_2}{2\rho\alpha_1 |s_{12}|}) \min\{\rho, 2|s_{12}|\alpha_1\} > 0,$$

and inequality (13) imposes that S_2 remains S.P.D. Hence lemma 1 is proved for $k=2$. Now assume that for every $\rho > 0$; there exist $\eta_{k-1} > 0$ and a S.P.D matrix S_{k-1} of the form (7) such that:

$$\forall t \geq 0, A_{k-1}^T(t)S_{k-1} + S_{k-1}A_{k-1}(t) - \rho C_{k-1}^T C_{k-1} \leq -\eta_{k-1} I_{k-1} \tag{0}$$

Let us show that for any $\rho > 0$; there exists $\eta > 0$ and a S.P.D matrix S_k, such that inequality (14) holds for S_k and C_k.
Set, $S_k = \begin{bmatrix} S_{k-1} & F_{k-1} \\ F_{k-1}^T & s_{kk} \end{bmatrix}$, where S_{k-1} satisfies (15) and F_{k-1} is a (k-1)-column vector of the form $F_{k-1} = [0, ..., 0, s_{k-1k}]^T$.
Set,

$$A_k(t) = \begin{bmatrix} A_{k-1} & v_{k-1}(t) \\ 0 & 0 \end{bmatrix}$$

where v_{k-1} is the (k-1)-column vector $v_{k-1} = [0, ..., 0, a_{k-1}(t)]^T$.
Consider the symmetric matrix $P_k(t) = S_k A_k(t) + A_k(t)^T S_k - \rho C_k^T C_k$. As for $k=2$, we will show that we can choose S_k such that for every $X_k \in \mathcal{R}^k$, $X_k^T P_k(t) X_k \leq -\eta \|X_k\|^2$.
A simple computation gives:

$$P_k(t) = \begin{bmatrix} P_{k-1}(t) & S_{k-1}v_{k-1}(t) + A_{k-1}^T(t)F_{k-1} \\ F_{k-1}^T A_{k-1}(t) + v_{k-1}^T(t)S_{k-1} & 2F_{k-1}^T v_{k-1}(t) \end{bmatrix}$$

hence,

$$X_k^T P_k(t) X_k = X_{k-1}^T P_{k-1}(t) X_{k-1} + 2X_{k-1}^T S_{k-1} v_{k-1}(t) x_k + 2s_{k-1k} a_{k-1}(t) x_k^2.$$

Now, choosing $s_{k-1k} < 0$ and using the fact that $0 < \alpha_1 \leq a_{k-1}(t) \leq \alpha_2$, we get:

$$X_k^T P_k(t) X_k = X_{k-1}^T P_{k-1}(t) X_{k-1} + 2\alpha_2 \|S_{k-1}\| \|X_{k-1}\| \, |x_k| + 2s_{k-1k}\alpha_1 x_k^2 \tag{0}$$

Inequalities (14) and (15) yield to:

$$X_k^T P_k(t) X_k \leq -\eta_{k-1} \|X_{k-1}\|^2 + 2\alpha_2 \|S_{k-1}\| \|X_{k-1}\| \, |x_k| + 2s_{k-1k}\alpha_1 x_k^2 \tag{0}$$

Now, consider $s_{k-1k} < 0$ and such that:

$$\frac{\alpha_2 \|S_{k-1}\|}{\sqrt{2\,|s_{k-1k}|\,\alpha_1 \eta_{k-1}}} < 1$$

or equivalently;

$$s_{k-1k} < -\frac{\alpha_2^2 \|S_{k-1}\|^2}{2\alpha_1 \eta_{k-1}}$$

Combining this last inequality with (16), we get:

$$X_k^T P_k(t) X_k \leq -\eta \|X_k\|^2, \text{ where } \eta = (1 - \frac{\alpha_2 \|S_{k-1}\|}{\sqrt{2\,|s_{k-1k}|\,\alpha_1 \eta_{k-1}}}) \min\{\eta_{k-1}, 2s_{k-1k}\alpha_1\}$$

To end the proof, we will choose s_{kk} so that S_k remains S.P.D.
From the above notations, we have:

$$\begin{aligned} X_k^T S_k X_k &= X_{k-1}^T S_{k-1} X_{k-1} + 2F_{k-1}^T X_{k-1} x_k + s_{kk} x_k^2 \\ &\geq X_{k-1}^T S_{k-1} X_{k-1} - 2\|F_{k-1}\| \|X_{k-1}\| \, |x_k| + s_{kk} x_k^2 \\ &\geq X_{k-1}^T S_{k-1} X_{k-1} - 2\,|s_{k-1k}|\, \|X_{k-1}\| \, |x_k| + s_{kk} x_k^2 \text{ (since } \|F_{k-1}\| = |s_{k-1k}|\,) \end{aligned}$$

From the induction hypothesis, we know that S_{k-1} is S.P.D. Hence there exist a constant $\tilde{\nu}_{k-1} > 0$ s.t. $X_{k-1}^T S_{k-1} X_{k-1} \geq \tilde{\nu}_{k-1} \|X_{k-1}\|^2$.
Thus,

$$X_k^T S_k X_k \geq \tilde{\nu}_{k-1} \|X_{k-1}\|^2 - 2\,|s_{k-1k}|\,\|X_{k-1}\|\,|x_k| + s_{kk} x_k^2$$

Now, choosing s_{kk} such that

$$s_{kk} > \frac{s_{k-1k}^2}{\nu_{k-1}}$$

or equivalently;

$$\frac{|s_{k-1k}|}{\sqrt{s_{kk}\nu_{k-1}}} < 1$$

we obtain:

$$X_k^T S_k X_k \geq \nu_k \|X_k\|^2 \text{ where } \nu_k = \left(1 - \frac{|s_{k-1k}|}{\sqrt{s_{kk}\nu_{k-1}}}\right) \inf\{\tilde{\nu}_{k-1}, s_{kk}\} > 0$$

This ends the proof of lemma 1.

proof of Theorem 2. Let α, c be the positive constants given in assumption H_3 and consider two any unknown matrices:

$$A_{n_i}(t) = \begin{bmatrix} 0 & a_1^i(t) & 0 & \dots & 0 \\ \vdots & & a_2^i(t) & & \vdots \\ 0 & & & \ddots & a_{n_i-1}^i(t) \\ 0 & \dots & & 0 & 0 \end{bmatrix} \text{ for } i = 1, 2$$

such that : $\alpha \leq a_j^i(t) \leq c$, for $i = 1, 2$ and $1 \leq j \leq n_i$.
From lemma 1 in which α_1, α_2 are respectively replaced by α and c, we know that there exists two matrices S_{n_1} and S_{n_2} which only depend on α and c such that:

$$\forall t \geq 0, A_{n_i}^T(t) S_{n_i} + S_{n_i} A_{n_i}(t) - \rho_i C_{n_i}^T C_{n_i} \leq -\eta_i I_{n_i} \qquad (0)$$

for some constants $\rho_i > 0$ and $\eta_i > 0$, $i = 1, 2$.

In what follows, we will show that system (9) in which the S_{n_i}'s are those satisfying (17), is an exponential estimator.

Set $e(t) = \hat{x}(t) - x(t) = \begin{bmatrix} e^1(t) \\ e^2(t) \end{bmatrix} = \begin{bmatrix} \hat{x}^1(t) - x^1(t) \\ \hat{x}^2(t) - x^2(t) \end{bmatrix}$, where $x(t)$ and $\hat{x}(t)$ are two respective trajectories of systems (8).
Differentiating $e(t)$, we obtain:

$$\left\{ \begin{array}{l} \dot{e}(t)^1(t) = f^1(\hat{x}(t), u(t)) - f^1(x(t), u(t)) \\ -r_1 \Delta_{\theta^{\delta_1}} S_{n_1}^{-1} C_{n_1}^T (C_{n_1} \hat{x}^1(t) - y_1(t)) + d^1(t) \\ \dot{e}^2(t) = f^2(\hat{x}(t), u(t)) - f^2(x(t), u(t)) \\ -r_2 \Delta_{\theta^{\delta_2}} S_{n_2}^{-1} C_{n_2}^T (C_{n_2} \hat{x}^2(t) - y_2(t)) + d^2(t) \end{array} \right. \qquad (0)$$

Recall that:

$$\left\{ \begin{array}{l} f_j^1 = f_j^1(x_1^1(t), \dots, x_{j+1}^1(t), u(t)); j = 1, \dots, n_1 - 1 \\ f_{n_1}^1 = f_{n_1}^1(x, u) \\ f_j^2 = f_j^2(x_1^2, \dots, x_{j+1}^2, u); j = 1, \dots, n_2 - 2 \\ f_{n_2-1}^2 = f_{n_2-1}^2(x, u) \\ f_{n_2}^2 = f_{n_2}^2(x, u) \end{array} \right.$$

Using the following notations:

$$\left\{\begin{array}{l}\delta f_j^1 = f_j^1(\hat{x}_1^1,\ldots,\hat{x}_j^1,x_{j+1}^1,u) - f_j^1(x_1^1,\ldots,x_{j+1}^1,u); j=1,\ldots,n_1-1\\ \delta f_{n_1}^1 = f_{n_1}^1(\hat{x},u) - f_{n_1}^1(x,u)\\ \delta f_j^2 = f_j^2(\hat{x}_1^2,\ldots,\hat{x}_j^2,x_{j+1}^2,u) - f_j^2(x_1^2,\ldots,x_{j+1}^2,u); j=1,\ldots,n_2-2\\ \delta f_{n_2-1}^2 = f_{n_2-1}^2(\hat{x}^1,\hat{x}_1^2,\ldots,\hat{x}_{n_2-1}^2,x_{n_2}^2,u) - f_{n_2-1}^2(x^1,x_1^2,\ldots,x_{n_2}^2,u)\\ \delta f_{n_2}^2 = f_{n_2}^2(\hat{x},u) - f_{n_2}^2(x,u)\end{array}\right. \quad (0)$$

we obtain:
for $1 \le j \le n_1 - 1$;

$$\begin{aligned} f_j^1(\hat{x},u) - f_j^1(x,u) &= f_j^1(\hat{x}_1^1,\ldots,\hat{x}_{j+1}^1,u) - f_j^1(\hat{x}_1^1,\ldots,\hat{x}_j^1,x_{j+1}^1,u) + \delta f_j^1\\ &= \frac{\partial f_j^1}{\partial x_{j+1}^1}(\hat{x}_1^1,\ldots,\hat{x}_j^1,\sigma_j^1(t),u(t))e_{j+1}^1(t) + \delta f_j^1\end{aligned}$$

where $\sigma_j^1(t) = \hat{x}_{j+1}^1(t) + \tau_j(t)e_{j+1}^1(t)$ for some $\tau_j(t)$, $0 \le \tau_j(t) \le 1$.
Similarly, we have:

$$f_j^2(\hat{x},u) - f_j^2(x,u) = \frac{\partial f_j^2}{\partial x_{j+1}^2}(\hat{x}_1^2,\ldots,\hat{x}_j^2,\sigma_j^2(t),u(t))e_{j+1}^2(t) + \delta f_j^2;\ \text{for } 1 \le j \le n_2 - 2$$

and,

$$f_{n_2-1}^2(\hat{x},u) - f_{n_2-1}^2(x,u) = \frac{\partial f_{n_2-1}^2}{\partial x_{n_2}^2}(\hat{x}^1,\hat{x}_1^2,\ldots,\hat{x}_{n_2-1}^2,\sigma_j^2(t),u(t))e_{n_2}^2(t) + \delta f_{n_2-1}^2$$

where $\sigma_j^2(t) = \hat{x}_{j+1}^2(t) + \tau_j'(t)e_{j+1}^2(t)$ for some $\tau_j'(t)$, $0 \le \tau_j'(t) \le 1$.
Now set:

$$a_j^1(t) = \frac{\partial f_j^1}{\partial x_{j+1}^1}(\hat{x}_1^1(t),\ldots,\hat{x}_j^1(t),\sigma_j^1(t));\ \text{for } j=1,\ldots,n_1-1$$

$$a_j^2(t) = \frac{\partial f_j^2}{\partial x_{j+1}^2}(\hat{x}_1^2(t),\ldots,\hat{x}_j^2(t),\sigma_j^2(t));\ \text{for } j=1,\ldots,n_2-2$$

$$a_{n_2-1}^1(t) = \frac{\partial f_j^2}{\partial x_{n_2}^2}(\hat{x}^1,\hat{x}_1^2,\ldots,\hat{x}_{n_2-1}^2,\sigma_{n_2}^2(t),u(t))$$

$$A_{n_i}(t) = \begin{bmatrix} 0 & a_1^i(t) & 0 & \ldots & 0\\ \vdots & & a_2^i(t) & & \vdots\\ 0 & & & \ddots & a_{n_i-1}^i(t)\\ 0 & \ldots & & 0 & 0\end{bmatrix}$$

The error equation (18) becomes:

$$\left\{\begin{array}{l}\dot{e}^1(t) = A_{n_1}(t)e^1(t) + \delta f^1 - r_1\Delta_{\theta^{\delta_1}}S_{n_1}^{-1}C_{n_1}^TC_{n_1}e^1(t) + d^1(t)\\ \dot{e}^2(t) = A_{n_2}(t)e^2(t) + \delta f^2 - r_2\Delta_{\theta^{\delta_2}}S_{n_2}^{-1}C_{n_2}^TC_{n_2}e^2(t) + d^2(t)\end{array}\right. \quad (0)$$

where $\delta f^1 = [\delta f_1^1, ..., \delta f_{n_1}^1]^T$ and $\delta f^2 = [\delta f_1^2, ..., \delta f_{n_2}^2]^T$.
As in [5], we use the following change of coordinates:

$$\begin{cases} \varepsilon_j^1 = \theta^{-j\delta_1} e_j^1 \text{ for } j = 1, ..., n_1 \\ \varepsilon_j^2 = \theta^{-j\delta_2} e_j^2 \text{ for } j = 1, ..., n_2 \end{cases}$$

Set $\varepsilon^1 = [\varepsilon_1^1, ..., \varepsilon_{n_1}^1]^T$, $\varepsilon^2 = [\varepsilon_1^2, ..., \varepsilon_{n_2}^2]^T$ and $\varepsilon = \begin{bmatrix} \varepsilon^1 \\ \varepsilon^2 \end{bmatrix}$, this change of coordinates takes the vectorial form:

$$\varepsilon = \begin{bmatrix} \varepsilon^1 \\ \varepsilon^2 \end{bmatrix} = \begin{bmatrix} \Delta_{\theta^{\delta_1}}^{-1} e^1 \\ \Delta_{\theta^{\delta_2}}^{-1} e^2 \end{bmatrix}$$

A simple calculation shows that the error equation (20) is equivalent to:

$$\begin{cases} \dot{\varepsilon}^1(t) = \theta^{\delta_1}(A_{n_1}(t) - r_1 S_{n_1}^{-1} C_{n_1}^T C_{n_1})\varepsilon^1(t) + \Delta_{\theta^{\delta_1}}^{-1}(\delta f^1 + d^1(t)) \\ \dot{\varepsilon}^2(t) = \theta^{\delta_2}(A_{n_2}(t) - r_2 S_{n_2}^{-1} C_{n_2}^T C_{n_2})\varepsilon^2(t) + \Delta_{\theta^{\delta_2}}^{-1}(\delta f^2 + d^2(t)) \end{cases} \tag{0}$$

To end the proof of theorem 2, we only need to show that for θ sufficiently large, we have:

$$\|\varepsilon(t)\| \leq c_1 e^{-\sigma t} + c_2 d$$

for some positive constants c_1, c_2 and σ.
To do so, we will show that the quadratic form $V(\varepsilon) = (\varepsilon^1)^T S_{n_1} \varepsilon^1 + (\varepsilon^2)^T S_{n_2} \varepsilon^2$ is a Lyapunov function for system (21). Let $\varepsilon(t)$ be a trajectory of (21), and differentiate $V(\varepsilon(t))$, we get:

$$\begin{aligned} \dot{V}(\varepsilon(t)) &= \theta^{\delta_1}(\varepsilon^1(t))^T [A_{n_1}(t)^T S_{n_1} + S_{n_1} A_{n_1}(t) - 2r_1 C_{n_1}^T C_{n_1}]\varepsilon^1(t) \\ &+ \theta^{\delta_2}(\varepsilon^2(t))^T [A_{n_2}(t)^T S_{n_2} + S_{n_2} A_{n_2}(t) - 2r_2 C_{n_2}^T C_{n_2}]\varepsilon^2(t) \\ &+ 2(\varepsilon^1(t))^T S_{n_1} \Delta_{\theta^{\delta_1}}^{-1}(\delta f^1 + d^1(t)) + 2(\varepsilon^2(t))^T S_{n_2} \Delta_{\theta^{\delta_2}}^{-1}(\delta f^2 + d^2(t)) \end{aligned}$$

Now, choose $r_i = \frac{\rho_i}{2}$ and using inequality (17), we get:

$$\begin{aligned} \dot{V}(\varepsilon(t)) &\leq -\theta^{\delta_1}\eta_1 \|\varepsilon^1(t)\|^2 - \theta^{\delta_2}\eta_2 \|\varepsilon^2(t)\|^2 \\ &+ 2(\varepsilon^1(t))^T S_{n_1} \Delta_{\theta^{\delta_1}}^{-1} \delta f^1 + 2(\varepsilon^2(t))^T S_{n_2} \Delta_{\theta^{\delta_2}}^{-1} \delta f^2 \\ &+ 2(\varepsilon^1(t))^T S_{n_1} \Delta_{\theta^{\delta_1}}^{-1} d^1(t) + 2(\varepsilon^2(t))^T S_{n_2} \Delta_{\theta^{\delta_2}}^{-1} d^2(t) \\ &\leq -\theta^{\delta_1}\eta_1 \|\varepsilon^1(t)\|^2 - \theta^{\delta_2}\eta_2 \|\varepsilon^2(t)\|^2 \\ &+ 2\sqrt{(\varepsilon^1(t))^T S_{n_1} \varepsilon^1(t)} \sqrt{(\Delta_{\theta^{\delta_1}}^{-1} \delta f^1)^T S_{n_1} \Delta_{\theta^{\delta_1}}^{-1} \delta f^1} \\ &+ 2\sqrt{(\varepsilon^2(t))^T S_{n_2} \varepsilon^2(t)} \sqrt{(\Delta_{\theta^{\delta_2}}^{-1} \delta f^2)^T S_{n_2} \Delta_{\theta^{\delta_2}}^{-1} \delta f^2} \\ &+ 2\sqrt{(\varepsilon^1(t))^T S_{n_1} \varepsilon^1(t)} \sqrt{(\Delta_{\theta^{\delta_1}}^{-1} d^1(t))^T S_{n_1} \Delta_{\theta^{\delta_1}}^{-1} d^1(t)} \\ &+ 2\sqrt{(\varepsilon^2(t))^T S_{n_2} \varepsilon^2(t)} \sqrt{(\Delta_{\theta^{\delta_2}}^{-1} d^2(t))^T S_{n_2} \Delta_{\theta^{\delta_2}}^{-1} d^2(t)} \end{aligned}$$

Recall that the following inequalities hold for any S.P.D matrix S:

$$\lambda_{min}(S)\|X\|^2 \leq X^T S X \leq \lambda_{max}(S)\|X\|^2$$

where $\lambda_{min}(S)$ and $\lambda_{max}(S)$ are respectively the smallest and the largest eigenvalues of S. Applying these inequalities, we obtain:

$$\begin{aligned}\dot{V}(\varepsilon(t)) \leq & -\frac{\theta^{\delta_1}\eta_1}{\lambda_{max}(S_{n_1})}(\varepsilon^1(t))^T S_{n_1}\varepsilon^1(t) - \frac{\theta^{\delta_2}\eta_2}{\lambda_{max}(S_{n_2})}(\varepsilon^2(t))^T S_{n_2}\varepsilon^2(t) \\ & + 2\sqrt{\lambda_{max}(S_{n_1})}\sqrt{(\varepsilon^1(t))^T S_{n_1}\varepsilon^1(t)}\|\Delta^{-1}_{\theta^{\delta_1}}\delta f^1\| \\ & + 2\sqrt{\lambda_{max}(S_{n_2})}\sqrt{(\varepsilon^2(t))^T S_{n_2}\varepsilon^2(t)}\|\Delta^{-1}_{\theta^{\delta_2}}\delta f^2\| \\ & + 2\sqrt{\lambda_{max}(S_{n_1})}\sqrt{(\varepsilon^1(t))^T S_{n_1}\varepsilon^1(t)}\|\Delta^{-1}_{\theta^{\delta_1}}d^1(t)\| \\ & + 2\sqrt{\lambda_{max}(S_{n_2})}\sqrt{(\varepsilon^2(t))^T S_{n_2}\varepsilon^2(t)}\|\Delta^{-1}_{\theta^{\delta_2}}d^2(t)\|\end{aligned}$$

Recall that $\Delta^{-1}_{\theta^{\delta_i}}\delta f^i = \begin{bmatrix} \frac{1}{\theta^{\delta_i}}\delta f^i_1 \\ \vdots \\ \frac{1}{\theta^{n_i\delta_i}}\delta f^i_{n_i}\end{bmatrix}$, and using H_3)-i) and the triangular structure of the (δf^i_j)'s we obtain:

For $j = 1, \dots, n_1 - 1$; $\theta^{-j\delta_1}|\delta f^1_j| \leq c\theta^{-j\delta_1}\sqrt{(e^1_1)^2 + \dots, (e^1_j)^2}$ where c is the Lipschitz constant of the f^i's.
Hence,

$$\theta^{-j\delta_1}|\delta f^1_j| \leq c\theta^{-j\delta_1}\sqrt{(\theta^{\delta_1}\varepsilon^1_1(t))^2 + \dots, (\theta^{j\delta_1}\varepsilon^1_j(t))^2}$$

Now, taking $\theta \geq 1$, we get:

$$\theta^{-j\delta_1}|\delta f^1_j| \leq c\sqrt{(\varepsilon^1_1(t))^2 + \dots, (\varepsilon^1_j(t))^2} \leq c\|\varepsilon^1(t)\| \quad (-11)$$

For $j = n_1$, we have:

$$\begin{aligned}\theta^{-n_1\delta_1}|\delta f^1_{n_1}| &\leq c\theta^{-n_1\delta_1}\|e(t)\| \\ &= c\theta^{-n_1\delta_1}\sqrt{(\theta^{\delta_1}\varepsilon^1_1(t))^2 + \dots + (\theta^{n_1\delta_1}\varepsilon^1_{n_1}(t))^2 + (\theta^{\delta_2}\varepsilon^2_1(t))^2 + \dots + (\theta^{n_2\delta_2}\varepsilon^2_{n_2}(t))^2}\end{aligned}$$

For $\theta \geq 1$, we obtain:

$$\theta^{-n_1\delta_1}|\delta f^1_{n_1}| \leq c\|\varepsilon^1(t)\| + c\theta^{n_2\delta_2 - n_1\delta_1}\|\varepsilon^2(t)\| \quad (-11)$$

Similarly for $\theta \geq 1$, we have:

$$\theta^{-j\delta_2}|\delta f^2_j| \leq c\|\varepsilon^2(t)\| \text{ for } 1 \leq j \leq n_2 \quad (-11)$$

For $j = n_2 - 1$, n_2, and $\theta \geq 1$, we obtain:

$$\left\{\begin{array}{c}\theta^{-(n_2-1)\delta_2}|\delta f^2_{n_2-1}| \leq c\|\varepsilon^2(t)\| + c\theta^{n_1\delta_1-(n_2-1)\delta_2}\|\varepsilon^1(t)\|, (j = n_2 - 1) \\ \theta^{-n_2\delta_2}|\delta f^2_{n_2}| \leq c\|\varepsilon^2(t)\| + c\theta^{n_1\delta_1 - n_2\delta_2}\|\varepsilon^1(t)\|, (j = n_2) \\ \leq c\|\varepsilon^2(t)\| + c\theta^{n_1\delta_1-(n_2-1)\delta_2}\|\varepsilon^1(t)\| \text{ (since } \theta \geq 1)\end{array}\right. \quad (-11)$$

Finally, inequalities (24) to (27) with $\theta \geq 1$, lead to:

$$\left\{\begin{array}{c}\|\Delta^{-1}_{\theta^{\delta_1}}\delta f^1\| \leq c\|\varepsilon^1(t)\| + c\theta^{n_2\delta_2-n_1\delta_1}\|\varepsilon^2(t)\| \\ \|\Delta_{\theta^{\delta_2}}\delta f^2\| \leq c\theta^{n_1\delta_1-(n_2-1)\delta_2}\|\varepsilon^1(t)\| + c\|\varepsilon^2(t)\|\end{array}\right.$$

hence;

$$\begin{cases} \|\Delta_{\theta^{\delta_1}}^{-1}\delta f^1\| \leq \frac{c}{\sqrt{\lambda_{min}(S_{n_1})}}\sqrt{\varepsilon^1(t)^T S_{n_1}\varepsilon^1(t)} + \frac{c\theta^{n_2\delta_2-n_1\delta_1}}{\sqrt{\lambda_{min}(S_{n_2})}}\sqrt{\varepsilon^2(t)^T S_{n_2}\varepsilon^2(t)} \\ \|\Delta_{\theta^{\delta_2}}\delta f^2\| \leq \frac{c\theta^{n_1\delta_1-(n_2-1)\delta_2}}{\sqrt{\lambda_{min}(S_{n_1})}}\sqrt{\varepsilon^1(t)^T S_{n_1}\varepsilon^1(t)} + \frac{c}{\sqrt{\lambda_{min}(S_{n_2})}}\sqrt{\varepsilon^2(t)^T S_{n_2}\varepsilon^2(t)} \end{cases} \tag{-11}$$

Combining (23) and (28), we obtain:

$$\begin{aligned} \dot{V}(\varepsilon(t)) \leq &-(\tfrac{\theta^{\delta_1}\eta_1}{\lambda_{max}(S_{n_1})} - 2c\tfrac{\sqrt{\lambda_{max}(S_{n_1})}}{\sqrt{\lambda_{min}(S_{n_1})}})(\varepsilon^1(t))^T S_{n_1}\varepsilon^1(t) \\ &-(\tfrac{\theta^{\delta_2}\eta_2}{\lambda_{max}(S_{n_2})} - 2c\tfrac{\sqrt{\lambda_{max}(S_{n_2})}}{\sqrt{\lambda_{min}(S_{n_2})}})(\varepsilon^2(t))^T S_{n_2}\varepsilon^2(t) \\ &+2c(\theta^{n_2\delta_2-n_1\delta_1}\tfrac{\sqrt{\lambda_{max}(S_{n_1})}}{\sqrt{\lambda_{min}(S_{n_2})}} + \theta^{n_1\delta_1-(n_2-1)\delta_2}\tfrac{\sqrt{\lambda_{max}(S_{n_2})}}{\sqrt{\lambda_{min}(S_{n_1})}})\sqrt{(\varepsilon^1(t))^T S_{n_1}\varepsilon^1(t)}\sqrt{\varepsilon^2(t)^T S_{n_1}\varepsilon^2(t)} \\ &+2\sqrt{\lambda_{max}(S_{n_1})}\sqrt{(\varepsilon^1(t))^T S_{n_2}\varepsilon^1(t)}\|\Delta_{\theta^{\delta_1}}^{-1}d^1(t)\| \\ &+2\sqrt{\lambda_{max}(S_{n_2})}\sqrt{(\varepsilon^2(t))^T S_{n_2}\varepsilon^2(t)}\|\Delta_{\theta^{\delta_2}}^{-1}d^2(t)\| \end{aligned}$$

Setting $W_1 = \theta^{\frac{\delta_1}{2}}\sqrt{(\varepsilon^1(t))^T S_{n_1}\varepsilon^1(t)}$, $W_2 = \theta^{\frac{\delta_2}{2}}\sqrt{(\varepsilon^2(t))^T S_{n_2}\varepsilon^2(t)}$, we obtain:

$$\begin{aligned} \dot{V}(\varepsilon(t)) \leq &-(\frac{\eta_1}{\lambda_{max}(S_{n_1})} - 2c\frac{\theta^{-\delta_1}\sqrt{\lambda_{max}(S_{n_1})}}{\sqrt{\lambda_{min}(S_{n_1})}})W_1 \\ &-(\frac{\eta_2}{\lambda_{max}(S_{n_2})} - 2c\frac{\theta^{-\delta_2}\sqrt{\lambda_{max}(S_{n_2})}}{\sqrt{\lambda_{min}(S_{n_2})}})W_2 \\ &+2c(\theta^{(n_2\delta_2-n_1\delta_1-\frac{\delta_1}{2}-\frac{\delta_2}{2})}\frac{\sqrt{\lambda_{max}(S_{n_1})}}{\sqrt{\lambda_{min}(S_{n_2})}} \\ &+\theta^{(n_1\delta_1-(n_2-1)\delta_2-\frac{\delta_1}{2}-\frac{\delta_2}{2})}\frac{\sqrt{\lambda_{max}(S_{n_2})}}{\sqrt{\lambda_{min}(S_{n_1})}})\sqrt{W_1}\sqrt{W_2} \\ &+2\theta^{-\frac{\delta_1}{2}}\sqrt{\lambda_{max}(S_{n_1})}\sqrt{W_1}\|\Delta_{\theta^{\delta_1}}^{-1}d^1(t)\| + 2\theta^{-\frac{\delta_2}{2}}\sqrt{\lambda_{max}(S_{n_2})}\sqrt{W_2}\|\Delta_{\theta^{\delta_2}}^{-1}d^2(t)\| \end{aligned}$$

Set $\alpha(\theta) = c(\theta^{(n_2\delta_2-n_1\delta_1-\frac{\delta_1}{2}-\frac{\delta_2}{2})}\frac{\sqrt{\lambda_{max}(S_{n_1})}}{\sqrt{\lambda_{min}(S_{n_2})}} + \theta^{(n_1\delta_1-(n_2-1)\delta_2-\frac{\delta_1}{2}-\frac{\delta_2}{2})}\frac{\sqrt{\lambda_{max}(S_{n_2})}}{\sqrt{\lambda_{min}(S_{n_1})}})$, we get:

$$\begin{aligned} \dot{V}(\varepsilon(t)) \leq &-(\frac{\eta_1}{\lambda_{max}(S_{n_1})} - 2c\frac{\theta^{-\delta_1}\sqrt{\lambda_{max}(S_{n_1})}}{\sqrt{\lambda_{min}(S_{n_1})}} - \alpha(\theta))W_1 \\ &-(\frac{\eta_2}{\lambda_{max}(S_{n_2})} - 2c\frac{\theta^{-\delta_2}\sqrt{\lambda_{max}(S_{n_2})}}{\sqrt{\lambda_{min}(S_{n_2})}} - \alpha(\theta))W_2 \\ &+2\theta^{-\frac{\delta_1}{2}}\sqrt{\lambda_{max}(S_{n_1})}\sqrt{W_1}\|\Delta_{\theta^{\delta_1}}^{-1}d^1(t)\| \\ &+2\theta^{-\frac{\delta_2}{2}}\sqrt{\lambda_{max}(S_{n_2})}\sqrt{W_2}\|\Delta_{\theta^{\delta_2}}^{-1}d^2(t)\| \end{aligned}$$

Now using the condition of theorem 2: $\frac{2n_1-1}{2n_2-1}\delta_1 < \delta_2 < \frac{2n_1+1}{2n_2-1}\delta_1$, it follows that:

$$n_1\delta_1 - (n_2-1)\delta_2 - \frac{\delta_1}{2} - \frac{\delta_2}{2} < 0$$

Hence there exist $\theta_0 \geq 1$ such that for every $\theta \geq \theta_0$, we have:

$$\begin{cases} 2c\frac{\theta^{-\delta_1}\sqrt{\lambda_{max}(S_{n_1})}}{\sqrt{\lambda_{min}(S_{n_1})}} + \alpha(\theta) \leq \frac{\eta_1}{2\lambda_{max}(S_{n_1})} \\ 2c\frac{\theta^{-\delta_2}\sqrt{\lambda_{max}(S_{n_2})}}{\sqrt{\lambda_{min}(S_{n_2})}} + \alpha(\theta) \leq \frac{\eta_2}{2\lambda_{max}(S_{n_2})} \end{cases} \tag{-14}$$

Combining (29) and (30), and taking $\theta \geq \theta_0$, we deduce that:

$$\begin{aligned}\dot{V}(\varepsilon(t)) &\leq -\frac{\eta_1}{2\lambda_{max}(S_{n_1})}W_1 - \frac{\eta_2}{2\lambda_{max}(S_{n_2})}W_2 \\ &+ 2\theta^{-\frac{3\delta_1}{2}}\sqrt{\lambda_{max}(S_{n_1})}d^1\sqrt{W_1} + 2\theta^{-\frac{3\delta_2}{2}}\sqrt{\lambda_{max}(S_{n_2})}d^2\sqrt{W_2} \qquad (-14)\end{aligned}$$

where $d = \sup_{t\geq 0}\{\|d_i(t)\|; i = 1, 2\}$
Hence,

$$\begin{aligned}\dot{V}(\varepsilon(t)) &\leq -\frac{\eta_1\theta^{\delta_1}}{2\lambda_{max}(S_{n_1})}(\varepsilon^1(t))^T S_{n_1}\varepsilon^1(t) - \frac{\eta_2\theta^{\delta_2}}{2\lambda_{max}(S_{n_2})}(\varepsilon^2(t))^T S_{n_2}\varepsilon^2(t) \\ &+ 2d\theta^{-\delta_1}\sqrt{\lambda_{max}(S_{n_1})}\sqrt{(\varepsilon^1(t))^T S_{n_1}\varepsilon^1(t)} \\ &+ 2d\theta^{-\delta_2}\sqrt{\lambda_{max}(S_{n_2})}\sqrt{(\varepsilon^2(t))^T S_{n_2}\varepsilon^2(t)} \\ &\leq -\sigma(\theta)V(\varepsilon(t)) + \gamma(\theta)d\sqrt{2}\sqrt{V(\varepsilon(t))} \qquad (-16)\end{aligned}$$

where
$\sigma(\theta) = \min\{\frac{\eta_1\theta^{\delta_1}}{2\lambda_{max}(S_{n_1})}, \frac{\eta_2\theta^{\delta_2}}{2\lambda_{max}(S_{n_2})}\}$;$\gamma(\theta) = \max\{2\theta^{-\delta_1}\sqrt{\lambda_{max}(S_{n_1})}, 2\theta^{-\delta_2}\sqrt{\lambda_{max}(S_{n_2})}\}$
and, $d = \max\{d^1, d^2\}$.
Hence the ball $B(0, \frac{\gamma(\theta)d}{\sigma(\theta)})$ exponentially attracts $\varepsilon(t)$.
Since $e(t) = \hat{x}(t) - x(t) = [(\Delta_{\theta^{\delta_1}}\varepsilon^1(t))^T\Delta_{\theta^{\delta_2}}\varepsilon^2(t)]^T$ it follows that $e(t)$ is exponentially attracted by some $B(0, \lambda_1(\theta)d)$.

3. APPLICATION TO BINARY DISTLLATION COLUMNS

Based on the Lewis assumption (the molar overflow is constant), the model of a binary distillation column that we consider is the classical (L, V) model:

$$\begin{cases} H_1\dot{x}_1 = V(y_2 - x_1) \text{ (total condenser)} \\ H_i\dot{x}_i = L(x_{i-1} - x_i) + V(y_{i+1} - y_i)(i = 2, \ldots, f-1; \text{ rectifying section}) \\ H_f\dot{x}_f = F(Z_f - x_f) + L(x_{f-1} - x_f) + V(y_{f+1} - y_f) \text{ (feed tray)} \\ H_i\dot{x}_i = (F + L)(x_{i-1} - x_i) + V(y_{i+1} - y_i) \ (i = f+1, \ldots, n-1; \text{ stripping section}) \\ H_n\dot{x}_n = (F + L)(x_{n-1} - x_n) + V(x_n - y_n)\text{(boiler)} \end{cases} \qquad (-16)$$

Where H_i is the liquid holdup on the ith-tray supposed known, x_i, y_i are the liquid and vapor compositions on the ith-tray, f is the number of the feed tray, F, L, V are the feed, reflux, and vapor rates (measured), and ZF is the feed composition. On each tray the liquid and vapor compositions, y_i and x_i, are linked by the liquid-vapor equilibrium law:

$$y_i = \frac{\alpha x_i}{1 + (\alpha - 1)x_i} \qquad (-16)$$

where α is the relative volatility constant $(0 < \alpha < 1)$. The state of the model are the set of liquid and feed compositions $(x_i, Z_f)$of the most volatile component. The two control variables are L and V, i.e. $u = [L, V]^T$. In practice, the top and bottom product compositions x_1 and x_n are measured, i.e. $y = [x_1, x_n]^T$. In the sequel, we will use the following notations:

$$\left\{ x^1 = \begin{bmatrix} x_1^1 \\ . \\ . \\ x_{n_1}^1 \end{bmatrix} = \begin{bmatrix} x_1 \\ . \\ . \\ x_{f-1} \end{bmatrix}; x^2 = \begin{bmatrix} x_1^2 \\ . \\ . \\ x_{n_2}^2 \end{bmatrix} = \begin{bmatrix} x_n \\ . \\ . \\ x_f \\ Z_f \end{bmatrix} \right.$$

Here, $n_1 = f - 1$, $n_2 = n - f + 2$, The dynamics of Z_f is assumed to be unknown and bounded:

$$\dot{Z}_f = \varepsilon(t)$$

The extended model is then of the form (9):

$$\left\{ \begin{array}{l} \dot{x}^1(t) = f^1(x(t), u(t)) \\ \dot{x}^2(t) = f^2(x(t), u(t)) + d^2(t) \\ y(t) = (x_1^1, x_1^2)^T = (x_1, x_n)^T \end{array} \right. \quad (-16)$$

where $d_{n_2}^2(t) = \varepsilon(t)$ and $d_i^2(t) = 0$, for $1 \leq i \leq n_2 - 1$.
In order to apply the observer design of section 2, one has to check hypothesis H_3).
Let us check H_3):
H_3) -i) can be obtained by extending nonlinear dynamics by global Lipschitz one. Indeed, the state components of the system are in the interval $]\epsilon, 1]$, where $\epsilon > 0$ is the smallest concentration.
H_3) -ii) For $1 \leq i \leq n_1 - 1 = f - 2$

$$\left\{ \begin{array}{l} f_1^1 = \dfrac{V}{H_1}(y_2 - x_1^1) \\ f_i^1 = \dfrac{V}{H_i}(y_{i+1} - y_i) + \dfrac{L}{H_i}(x_{i-1}^1 - x_i^1) \\ y_i = \dfrac{\alpha x_i^1}{1 + (\alpha - 1)x_i^1} \end{array} \right.$$

hence for $1 \leq i \leq n_1 - 1 = f - 2$ we have:

$$\frac{\partial f_i^1}{\partial x_{i+1}^1} = \frac{V}{H_i}\frac{\partial y_{i+1}}{\partial x_{i+1}^1} = \frac{V}{H_i}\frac{\alpha}{(1 + (\alpha - 1)x_i^1)^2} \geq \frac{V}{H_i}\frac{\alpha}{(1 + (\alpha - 1)\varepsilon)^2} > 0 \quad (-16)$$

Similarly, for $1 \leq i \leq n_2 - 1 = n - f + 1$ we have:

$$\left\{ \begin{array}{l} f_1^2 = \dfrac{1}{H_n}[(F + L)(x_2^2 - x_1^2) + V(x_1^2 - y_n)] \\ f_i^2 = \dfrac{1}{H_i}[(F + L)(x_{i-1}^2 - x_i^2) + V(y_{i+1} - y_i)] \\ f_{n_2-1}^2 = F(Z_f - x_f) + L(x_{n_2-2}^2 - x_{n_2-1}^2) + V(y_{f+1} - y_f) \end{array} \right.$$

then,

$$\left\{ \begin{array}{l} \dfrac{\partial f_i^2}{\partial x_{i+1}^2} = \dfrac{(F + L)}{H_i} > 0; \qquad 1 \leq i \leq n_2 - 2 = n - f \\ \dfrac{\partial f_{n_2-1}^2}{\partial x_{n_2}^2} = \dfrac{F}{H_f} > 0; \text{ for } i = n_2 - 1 \end{array} \right. \quad (-16)$$

Simulation results: In order to show the performance of the proposed observer, we consider the binary methanol-ethanol distillation column with the following characteristics:

$$n = 12; f = 8; F = 1.65\ (mol/min); \alpha = 1.65$$

$$L = 4.69\ (mol/min); V = 5.94\ (mol/min); H_1 = 5.5\ mol; H_{12} = 5.5\ mol$$

$$H_i = 0.55\ mol \quad for\ 2 \leq i \leq n-1$$

The feed composition Z_f varies as a no uniform square signal from 0.75 to 0.38 ($\varepsilon(t) = 0$). The model simulation was performed under the following initial conditions:

$$\begin{aligned} x(0) = \ (& \ 0.93000585, 0.88953536, 0.84223977, 0.78925097, \\ & 0.73261728, 0.67505746, 0.61947090, 0.56837712, 0.47000139, \\ & 0.36710020, 0.27017424, 0.18748224, 0.75)^T \end{aligned} \qquad (-17)$$

The gain of the observer was synthesized following the procedure described in proofs of the above lemma 1 with:

$$r_1 = r_2 = 1,$$

$$\delta_1 = 1, \delta_2 = \frac{2n_1}{2n_2 - 1}\delta_1 = \frac{14}{11}$$

and

$$S_{n_1} = \begin{bmatrix} 1.5 & -0.5 & 0 & 0 & 0 & 0 & 0 \\ -0.5 & 2 & -1.5 & 0 & 0 & 0 & 0 \\ 0 & -1.5 & 4 & -2 & 0 & 0 & 0 \\ 0 & 0 & -2 & 8 & -3 & 0 & 0 \\ 0 & 0 & 0 & -3 & 10.5 & -4 & 0 \\ 0 & 0 & 0 & 0 & -4 & 15.5 & -5 \\ 0 & 0 & 0 & 0 & 0 & -5 & 17.5 \end{bmatrix},$$

$$S_{n_2} = \begin{bmatrix} 1.5 & -0.5 & 0 & 0 & 0 & 0 \\ -0.5 & 2 & -1.5 & 0 & 0 & 0 \\ 0 & -1.5 & 4 & -2 & 0 & 0 \\ 0 & 0 & -2 & 8 & -3 & 0 \\ 0 & 0 & 0 & -3 & 10.5 & -4 \\ 0 & 0 & 0 & 0 & -4 & 15.5 \end{bmatrix}.$$

then,

$$S_{n_1}^{-1} C_{n_1}^T = [0.7610, 0.2830, 0.1237, 0.0352, 0.0113, 0.0032, 0.0009]^T$$

and

$$S_{n_2}^{-1} C_{n_2}^T = [0.7610, 0.2830, 0.1237, 0.0352, 0.0113, 0.0032]^T$$

Since the obtained results are quite similar and in order to avoid curves redundancy, we only present here those related to three trays which respectively correspond to a tray (tray 3) in the rectifying section, to the feed tray (composition in feed tray x_f and feed composition Z_f) and finally to a tray 10 in the stripping section.

Two sets of simulation results of the observer are presented:

The first one is with free noisy outputs measurements $x_1(t)$ and $x_{12}(t)$ issued from simulation of model (33)(see Fig. 1). The second sets of simulation is with noisy outputs measurements $x_1(t)$ and $x_{12}(t)$ issued from simulation of model (33) and respectively corrupted by an adding a Gaussian noise with zero mean and an amplitude equivalent to 5% of the corresponding values (see Fig. 2).

As previously mentioned, the observer convergence can be enhanced with large values of θ. However, such values are to be avoided, since the observer generate the so-called peak phenomena (see Fig. 3.) and my become noise sensitive (see Fig. 4.). To obtain good results, the choice a value of θ is one which provided a best compromise between the fast convergence and the noise rejection (see Fig. 5. and Fig. 6.).

3. CONCLUSION

In this paper, we have extended the high gain observer proposed in [5,6] to a class of multi-output nonlinear system which are not necessarily control affine.
A new algorithm permitting to compute the gain of the observer is proposed. The construction is based on a symmetric positive definite matrix having a simple structure. Moreover, this S.P.D. matrix plays a key role in the proof of the convergence of the estimator.
The extension of this result to a multi-output systems (more then two outputs), requires an adequate structure. A interesting structure consists of defining a class of non affine systems extending the class of uniformly observable systems which are proposed in [2].

REFERENCES

1. Besancon G., and Hammouri H., 'On uniform observation of non uniformly observable systems', *systems and control leters*, **29**, 9-19(1996).
2. Bornard G., and Hammouri H., 'A high gain observer for a class of uniformly observable systems', *Proc. 30th IEEE Conference on Decision and Control*, **122** (1991), Brighton, England.
3. Deza F., 'Contribution to the synthesis of exponential observers', *PhD thesis, INSA Rouen, France*, (1991).
4. Gauthier J.P., and Bornard G., 'Observability for any u(t) of a class of nonlinear systems', *IEEE Trans. on Aut. Control.* **26** 922-926(1981).
5. Gauthier J.P., Hammouri H., and Othman S., 'A simple observer for nonlinear systems-Application to bioreactors', *IEEE Trans. on Aut. Control.* **37**, No. 6, 875-880(1992).
6. Gauthier J.P. and Kupka I.A.K., 'Observability and observers for nonlinear systems', *SIAM J. Control optim.*, **32**,975-994(1994).
7. Hermann R. and Krenner A.J., 'Nonlinear Controllability and Observability', *IEEE trans. on Aut. Control,* **AC-22**, No. 5,728-740(1977).
8. Krener A.J. and Isidori A., 'Linearization by output injection and nonlinear observers', *Syst. and Control Letters*, **3**, 47-52(1983).
9. Sussman H.J., 'Single-Input Observability of Continuous-Time systems', *Math. System Theory*, **12**, 371-393(1979).
10. Xia X.H., and Gao W.B., 'Nonlinear observer design by observer error linearization', *SIAM J. Control and Opt.*, **27**, No. 1, 199-216(1989).

FIGURES CAPTION

Fig. 1. Simulated free noisy outputs $x_1(t)$ and $x_{12}(t)$

Fig. 2. Simulated noisy outputs $x_1(t)$ and $x_{12}(t)$

Fig. 3. Comparison of estimated and simulated data from free noisy outputs (peak phenomena $\theta = 4.8$)

Fig. 4. Comparison of estimated and simulated data from noisy outputs (noisy sensitive $\theta = 4.2$)

Fig. 5. Comparison of estimated and simulated data from free noisy outputs (good results $\theta = 3.5$)

Fig. 6. Comparison of estimated and simulated data from noisy outputs (good results $\theta = 3.5$)

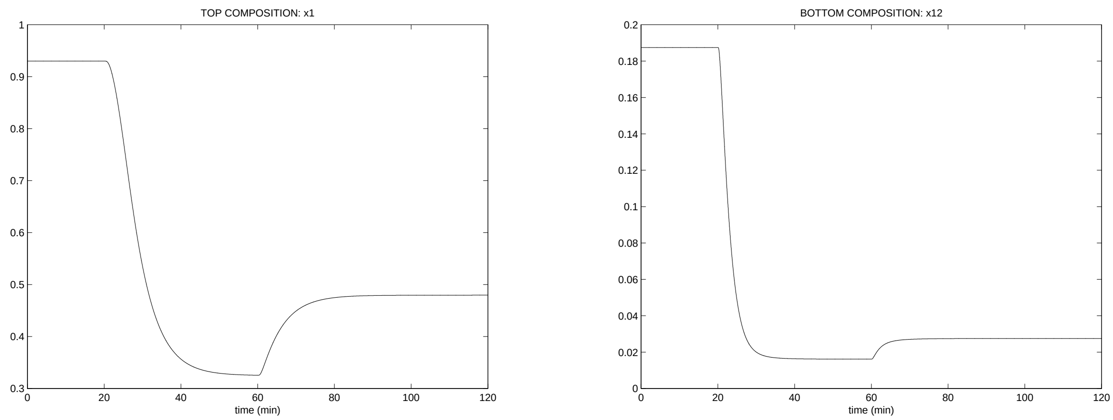

Fig. 1. Simulated free noisy outputs $x_1(t)$ and $x_{12}(t)$

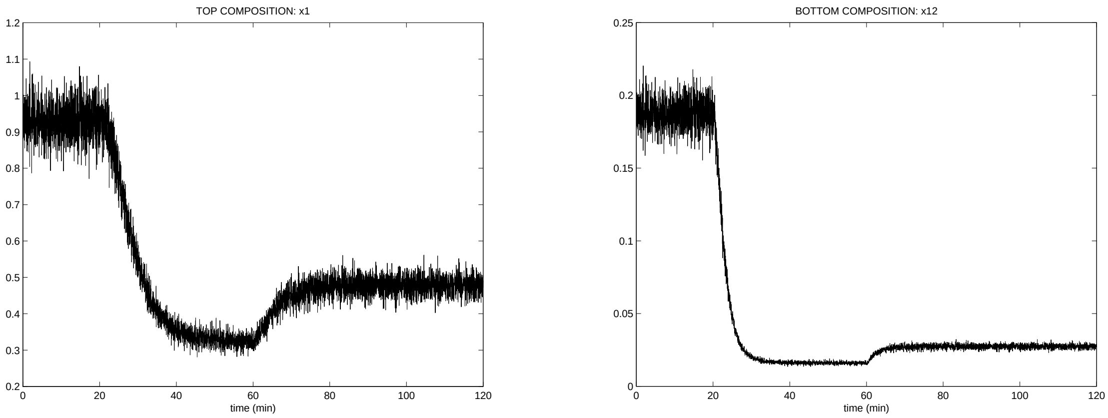

Fig. 2. Simulated noisy outputs $x_1(t)$ and $x_{12}(t)$

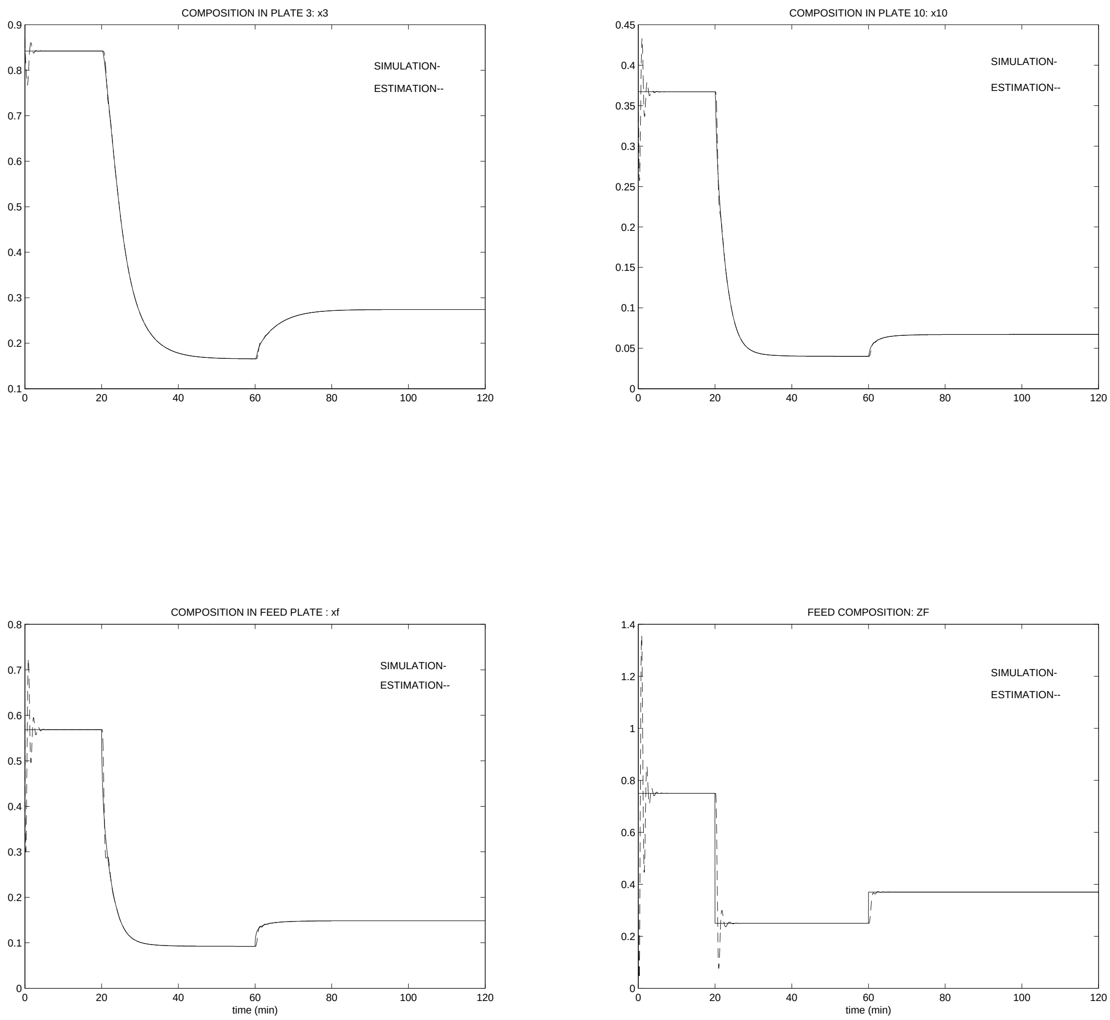

Fig. 3. Comparison of estimated and simulated data from free noisy outputs (peak phenomena $\theta = 4.8$)

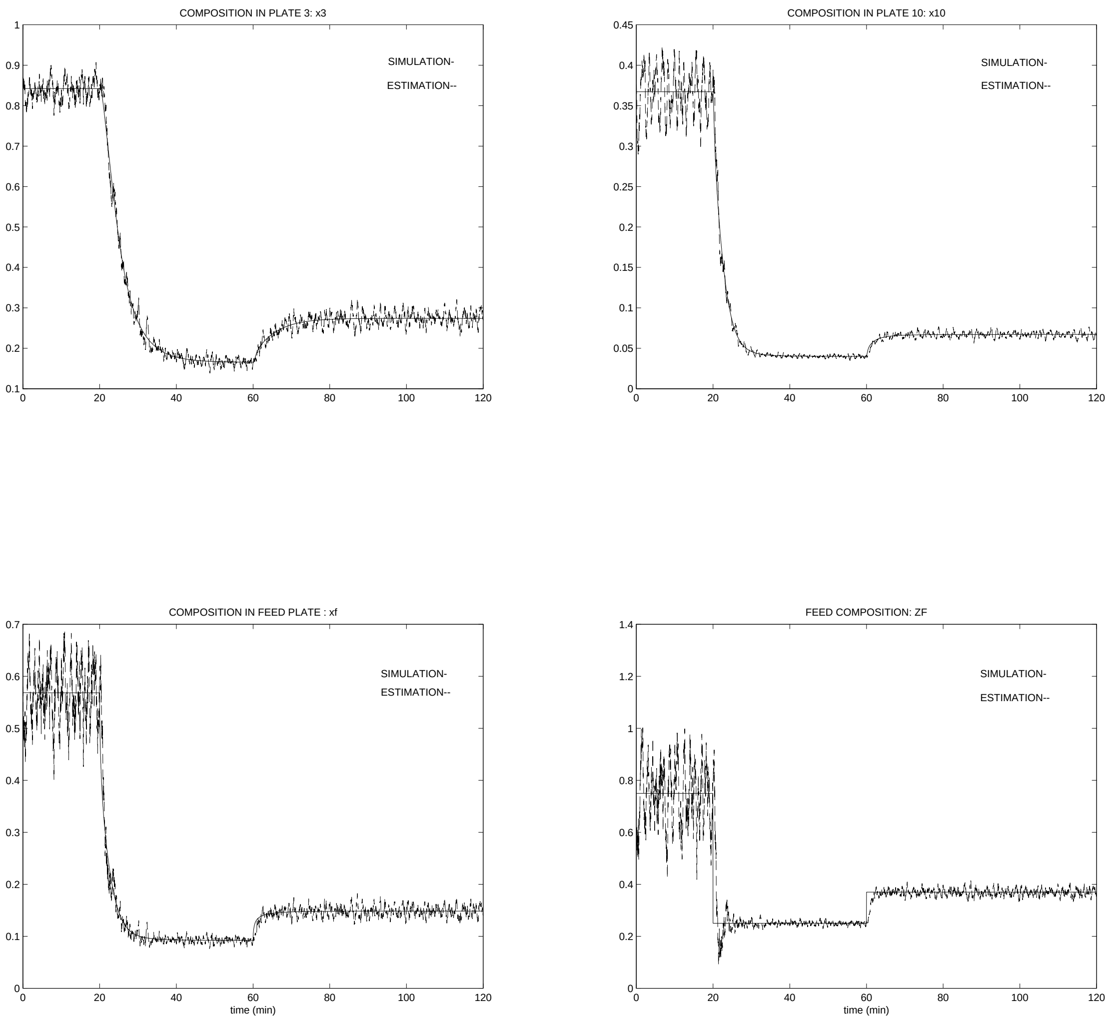

Fig. 4. Comparison of estimated and simulated data from noisy outputs (noisy sensitive $\theta = 4.2$)

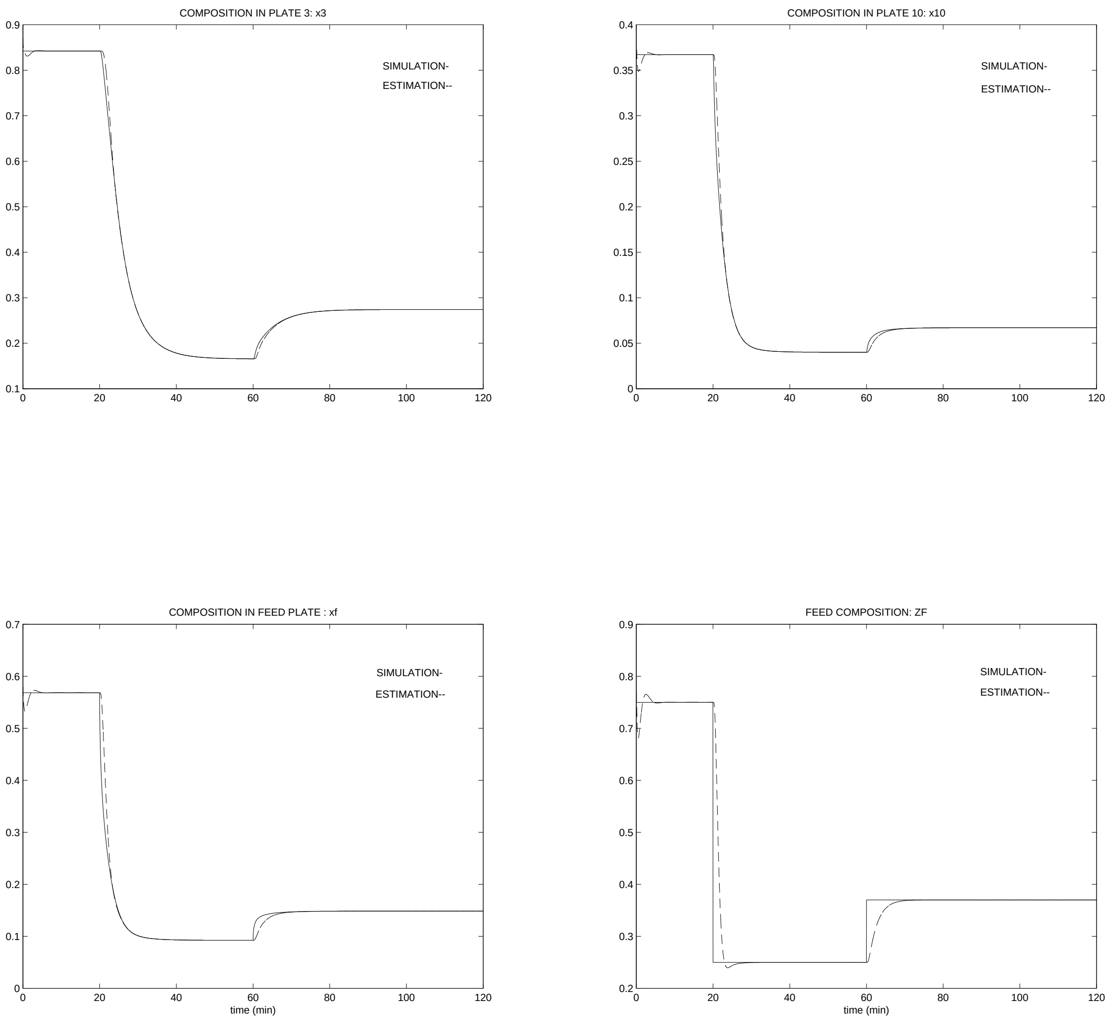

Fig. 5. Comparison of estimated and simulated data from free noisy outputs (good results $\theta = 3.5$)

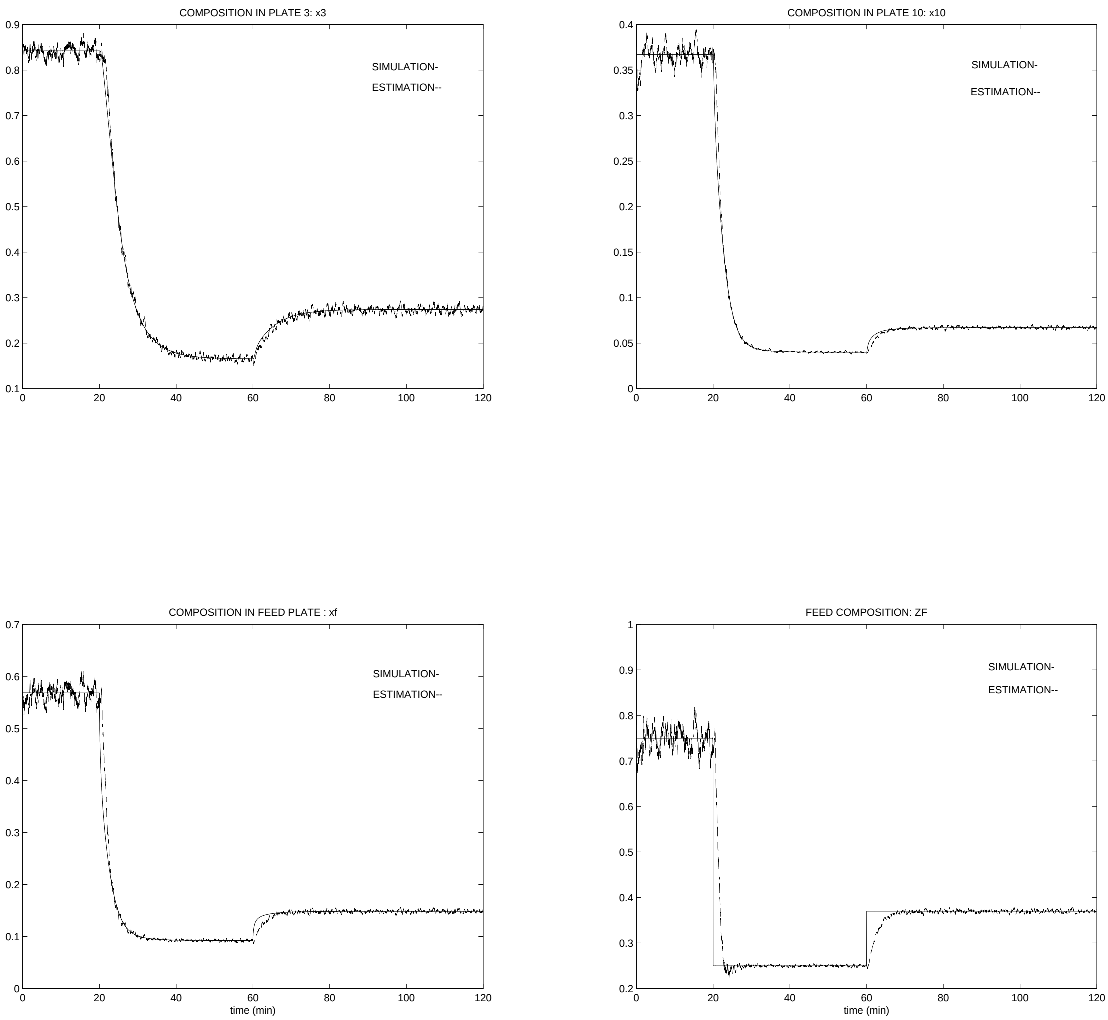

Fig. 6. Comparison of estimated and simulated data from noisy outputs (good results $\theta = 3.5$)