Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Projectional Differential Neural Network Observer with Stable Adaptation Weights

Alejandro García[1], Alexander Poznyak[1], Isaac Chairez[2], and Tatyana Poznyak[2]

***Abstract*— A class of *dynamic neural network* (DNN) observers involving a projection operator inside is considered. Such observers seem to be useful when an uncertain nonlinear system, affected by external perturbations, keeps its states in an *a priori* known compact set, defined by the given state constraints independently of the measurement noise effects. Since the projection method introduces discontinuities into the trajectory dynamics, the standard Lyapunov method is not applicable to describe the convergence property of this class of observers. This problem is suggested to be resolved using a Lyapunov-Krasovski functional including both the estimation error and the weights involved in the DNN description. The stable adaptive laws for the DNN-weights adjustment are derived. The upper bound for the estimation error is obtained based on Linear Matrix Inequality (LMI) technique implementation. An illustrative example clearly shows the effectiveness of the suggested approach. It deals with an environment control problem, related to the soil contaminants degradation by ozonation.**

I. INTRODUCTION

The majority of modern controllers assumes the availability of the current state-vector of a system to be controlled. However, in many practical situations only inputs and outputs of a system are available (measurable). Therefore, one of the frequent challenges for practical control engineers is to design a workable *state observer* (or filter), based only on the current available information [1]. Such observers are often treated as *software-sensors*. The practical usefulness of state observers is related not only with a possible system monitoring and regulation but also with possibility to detect (identify) failures occurred in the considered dynamic system. Some common examples of the observers structures are: based on the Lie-algebraic method [2], Lyapunov-like observers [3], the high gain observation [4], recent structures based on sliding mode technique [5], numerical approaches as the set-membership observers [6] and etc.

All approaches mentioned above assume a complete knowledge of the system structure (mathematical model) for their design. Therefore the presence of disturbances, uncertainties and nonlinearities pose a great challenge [1]. If the mathematical model of a considered process is incomplete or partially known, it is possible to take advantage of the function approximation capacity of the artificial *Neural Network* (NN) [7] involving it in the observer structure designing [8],[9].

(1)Department of Automatic Control, CINVESTAV-IPN, AP. 14740, Av. Instituto Politécnico Nacional No. 2508, C.P. 07360, México D.F., México. e-mail: $apoznyak(agarcia)@ctrl.cinvestav.mx$

(2)Superior School of Chemical Engineering and Extractive Industries (ESIQIE-IPN), Edif. 7, UPALM, C.P. 07738, México, D.F, México, e-mail: $tpoznyak@ipn.mx$

There are known two types of NN: *static* one, [10], and *dynamic* neural networks (DNN) [11]. The first one deals with the class of global optimization problems trying to adjust the weights of such NN to minimize an identification error. The second approach, exploiting the feedback properties of the applied DNN, permits to avoid many problems related to global extremum search converting the learning process to an adequate feedback design [12], the DNN-approach provides an effective instrument to attack a wide spectrum of problems such as identification, state estimation, trajectories tracking an etc. [12] [13]. Moreover DNN have demonstrated perfect identification properties in the presence of uncertainties and external disturbances, in other words, they provide the *robustness* property.

In this paper we deal with the state estimation problem for a special class of nonlinear uncertain systems affected by additive bounded disturbances both in the state dynamics and in the measurable outputs. Here we design a state observer based on a special class of DNN containing the projection operator. The adaptive behavior of this DNN structure is carry out solving numerically two stable matrix differential equations derived based on the stability analysis by the direct Lyapunov's method and LMI technique. The specific feature of the considered control processes is that the state-vector $x(t)$ always belongs to a given *compact set* X even in the presence of noise. For example, the so-called "nonnegative systems" evidently have this property. It seems to be natural that the generated state estimates $\hat{x}(t)$ also belong to the same compact. To provide this property a projectional operator (which mapping is never differentiable) in each integration step is introduced.

II. ESTIMATION PROBLEMS UNDER STATE CONSTRAINTS

Consider the nonlinear continuous-time model given by the following ODE:

$$\begin{array}{c} \dot{x}(t) = f\left(x(t), u\left(t\right)\right) + \xi(t),\ x(0) \text{ is fixed} \\ y(t) = Cx(t) + \eta(t) \end{array} \tag{1}$$

where $x(t) \in \Re^n$ is the state-vector at time $t \geq 0$, $y_t \in \Re^m$ is the corresponding measured output, available for a designer at any time t, the known matrix $C \in \Re^{m\times n}$ defines the state-output transformation, $u\left(t\right) \in \Re^r$ is a bounded control action $(r \leq n)$ belonging to the following admissible set $U^{adm} := \left\{u\left(t\right) : \left\|u\left(t\right)\right\| \leq \Upsilon_u < \infty\right\}$, $\xi(t)$ and $\eta(t)$ are noises in the state dynamics and in the output, respectively, $f : \Re^{n\times r} \to \Re^n$.

In many practical problems a designer knows *a priori* that the state-vector $x\left(t\right)$ always belongs to a given *compact set*

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

X (even in the presence of noise) which has a concrete physical sense. For example, the dynamic behavior of some reagents, participating in chemical reactions, always keep nonnegative the current values.

The *state estimation (observation) problem* consists in designing a vector-function $\hat{x}(t) \in \mathbb{R}^n$ depending only on the data $\{y(t), u(t)\}_{\tau\in[0,t]}$ available up to the time t in such a way that it would be "close" to its real (but non-measurable) state-vector $x(t)$. The measure of that "closeness" depends on the accepted assumptions concerning the state dynamics as well as the noise effects. The most of observers, solving this problem, also have an ODE-structure, usually given by

$$\frac{d}{dt}\hat{x}(t) = F\left(\hat{x}(t), u(t), y_{\tau\in[0,t]}, t\right), \ \hat{x}_0 \text{ is a fixed vector} \tag{2}$$

Here the mapping $F : \Re^n \times \Re^r \times \Re^m \times \Re^+ \rightarrow \Re^n$ defines the structure of the observer to be implemented. The property of an observer, which we are looking for, is to keep the generated state estimates $\hat{x}(t)$ always within the given compact set X, that is,

$$\hat{x}(t) \in X \tag{3}$$

Indeed, for example, applying a linear feed-back $u(t) = K\hat{x}(t)$ with a high-gain K, may provoke a significant instability effect of the corresponding close-loop dynamics if any changing of a sign in $\hat{x}(t)$ are admissible. As is known observation techniques consider partial or full knowledge of a system mathematical model. In this paper it is shown that for a wide class of models the projectional DNN observer permits to avoid this constrain, namely, the function $f(x, u)$ in the model description (1) is admitted to be unknown exactly (may be, belonging to some class) and fulfilling the condition (3).

III. PROJECTIONAL DNN OBSERVERS

Let us consider the following observer referred hereafter to as the *projectional observer:*

$$\hat{x}(t) = \pi_X\{\hat{x}(t\text{-}h(t)) + \int_{\tau=t-h(t)}^{t} F(\hat{x}(\tau), u(\tau), y(\tau), \tau)\, d\tau\}, \tag{4}$$

Here $t > h(0)$ and $h(t) \in C^1$ is supposed to be given and non-increasing positive function, that is, $\dot{h}(t) \leq 0$. The operator $\pi_X\{\cdot\}$ is the projector to the given convex compact set X satisfying the condition

$$\|\pi_X\{x\} - z\| \leq \|x - z\| \tag{5}$$

for any $x \in \mathbb{R}^n$ and any $z \in X$. The operator $\pi_X\{\cdot\}$ may be defined non uniquely. An example of $\pi_X\{\cdot\}$ is given below.

Example 1:

$$\pi_X\{x\} = \begin{bmatrix} \text{sat}(x_1) & \ldots & \text{sat}(x_n) \end{bmatrix}^\top \tag{6}$$

where for any $i = 1 \ldots n$

$$\text{sat}(x_i) := \begin{cases} (x_i)^- & x_i \leq (x_i)^- \\ x_i & (x_i)^- < x_i < (x_i)^+ \\ (x_i)^+ & x_i \geq (x_i)^+ \end{cases}$$

with $(x_i)^- < (x_i)^+$ as an extreme point *a priori* known.

Remark 1: Notice that with the implementation of the projectional operator, the trajectories $\{\hat{x}(t)\}$ generated by (4) are not differentiable for any $t \geq h(t) > 0$.Structure of DNN Observers

1) The complete information case: If the right-hand side $f(x(t))$ of the dynamics (1) is known then usually the structure F of the observer (2) is selected in the, so-called, Luenberger-type form:

$$F(\hat{x}(t), u(t), y(t), t) = f(\hat{x}(t), u(t)) + K(t)(y(t)\text{-}C\hat{x}(t)) \tag{7}$$

So, it repeats the dynamics of the plant and, additionally, contains the correction term, proportional to the output error (see, for example, [14], [15], [3] and [16]). The adequate selection of the matrix-gain $K(t)$ provides a good-enough state estimation .

2) The "grey-box" case: In the case when the right-hand side $f(x, u)$ of the dynamics (1) is unknown, there is suggested to apply some approximation of it, say, $\bar{f}(x(t), u(t) \mid W(t))$ where $\bar{f} \in \Re^n$ defines the approximative mapping depending on the time-varying parameters $W(t)$ which should be adjusted by a concrete "*adaptation law*" suggested by a designer. According to the DNN-approach [12], we may decompose $\bar{f}(x(t), u(t) \mid W(t))$ into two parts: first one approximates the linear dynamics part by a *Hurwitz* fixed matrix $A \in \Re^{n\times n}$ (selected by the designer) and nonlinear part is approximated by variable time parameters $W_{1,2}(t)$ with "sigmoid" multipliers, that is:

$$\begin{array}{c} \bar{f}(x(t), u(t) \mid W_{1,2}(t)) := \\ Ax(t) + W_1(t)\sigma(x(t)) + W_2(t)\varphi(x(t))u(t) \\ A \in \Re^{n\times n}, \ W_1(t) \in \Re^{n\times p}, \ \sigma(\cdot) \in \Re^{p\times 1} \\ W_2(t) \in \Re^{n\times q}, \ \varphi(\cdot) \in \Re^{q\times r} \end{array} \tag{8}$$

The activation vector-function $\sigma(\cdot)$ and matrix-function $\varphi(\cdot)$ are usually selected as functions with *sigmoid-type components*, In (8) The constant parameters A as well as the time-varying parameters $W_{1,2}(t)$ should be properly adjusted to guarantee a good state approximation. Notice that for any fixed matrices $W_{1,2}(t) = \hat{W}_{1,2}$ the dynamics (1) always could be represented as

$$\begin{array}{c} \dot{x}(t) = Ax(t) + \hat{W}_1\sigma(x(t)) + \hat{W}_2\varphi(x(t))u(t) + \tilde{f}(t) + \xi(t) \\ \tilde{f}(t) := f(x(t)) - \bar{f}\left(x(t) \mid \hat{W}_{1,2}\right) \end{array} \tag{9}$$

where $\tilde{f}(t)$ is referred to as a modelling error vector-field called the "*unmodelled dynamics*". In view of the corresponding boundedness property, the following upper bound for the unmodelled dynamics $\tilde{f}(t)$ takes place*:*

$$\begin{array}{c} \left\|\tilde{f}(t)\right\|^2_{\Lambda_f} \leq \tilde{f}_0 + \tilde{f}_1 \|x(t)\|^2_{\Lambda^1_{\tilde{f}}} \\ \tilde{f}_0, \ \tilde{f}_1 > 0; \ \Lambda_f, \Lambda^1_{\tilde{f}} > 0, \ \Lambda_f = \Lambda_f^\top, \ \Lambda^1_{\tilde{f}} = \left(\Lambda^1_{\tilde{f}}\right)^\top \end{array} \tag{10}$$

A. Structure of projectional DNN observers

Introduce the following projectional DNN observer:

$$
\begin{array}{c}
\hat{x}(t)= \\
\pi_X\left\{\hat{x}(t-h(t))+\right. \\
\int\limits_{\tau=t-h(t)}^{t}\left[A\hat{x}(\tau)+W_1(\tau)\sigma\left(\hat{x}(\tau)\right)+\right. \\
\left.\left.W_2(\tau)(\varphi\left(x(\tau)\right)u(\tau)+Ke(\tau)\right]d\tau\right\} \\
e(t):=y(t)-C\hat{x}(t)
\end{array}
\tag{11}
$$

Here the weights matrices $W_1(t)$ and $W_2(t)$ supply the adaptive behavior to this class of observers if they are adjusted by an adequate manner. We derived (see Appendix) the following nonlinear weight *updating* laws based on the Lyapunov-like stability analysis:

$$
\left.
\begin{array}{c}
\dot{W}_1(t)=-\dfrac{k_1^{-1}(t)}{2}P\Omega(t)\sigma^{\intercal}\left(\hat{x}(t)\right)-\dot{k}_1(t)\tilde{W}_1(t) \\
\Omega(t):=\Pi\tilde{W}(t)\sigma\left(\hat{x}(t)\right)+2N_{\varpi}C^{\intercal}e(t-h(t)) \\
\tilde{W}_1(t):=W_1(t)-\hat{W}_1 \\
\Pi=\left(N_{\varpi}\left(\varpi\Lambda_3+C^{\intercal}\Lambda_2C\right)N_{\varpi}P+I\right)
\end{array}
\right\}
\tag{12}
$$

$$
\left.
\begin{array}{c}
\dot{W}_2(t)=-\dfrac{k_2^{-1}(t)}{2}P\Phi(t)u^{\top}(\tau)\varphi^{\top}\left(\hat{x}(\tau)\right)-\dot{k}_2(t)\tilde{W}_2(t) \\
\Phi(t):=\Xi\tilde{W}_2(\tau)(\varphi\left(\hat{x}(\tau)\right)u(\tau)+2N_{\varpi}C^{\intercal}e(t-h(t)) \\
\tilde{W}_2(t):=W_2(t)-\hat{W}_2 \\
\Xi=\left(N_{\varpi}\left(\varpi\Lambda_7+C^{\intercal}\Lambda_6C\right)N_{\varpi}P+I\right)
\end{array}
\right\}
\tag{13}
$$

where:

$$
N_{\varpi}=\left(C^{\intercal}C+\varpi I\right)^{-1},\ \varpi>0
$$

To improve the behavior of this adaptive laws, the matrix $\hat{W}_{1,2}$ can be "provided" by one of the, so-called, *training algorithms* (see, for example, [17] and [18]),

IV. Upper Bound for State Estimation Error

A. Behavior of weights dynamics

Here we wish to show that under the adapting weights laws (12) and (13) the weights $W_1(t)$ and $W_2(t)$ are bounded.

Theorem 1: If $k_{i,t}$ $(i=1,2)$ in (12) and (13) satisfy

$$
\begin{array}{l}
\dot{k}_{1,t}\leq-\dfrac{2\left(k_1(t)\right)^2\left|tr\left\{\tilde{W}_1^{\intercal}(t)\,P\Omega(t)\sigma^{\intercal}\left(\hat{x}(t)\right)\right\}\right|}{tr\left\{\tilde{W}_1^{\intercal}(t)\,\tilde{W}_1(t)\right\}+ck_1(t)\left[k_1(t)-k_{1\min}\right]} \\
\dot{k}_{2,t}\leq-\dfrac{2\left(k_2(t)\right)^2\left|tr\left\{\tilde{W}_2(t)\,P\Phi(t)u^{\top}(t)\varphi^{\top}\left(\hat{x}(t)\right)\right\}\right|}{tr\left\{\tilde{W}_2(t)^{\intercal}\,\tilde{W}_2(t)\right\}+ck_2(t)\left(k_2(t)-k_{2,\min}\right)}
\end{array}
\tag{14}
$$

then $tr\left\{\tilde{W}_1^{\intercal}(t)\,\tilde{W}_1(t)\right\}$ is monotonically non-decreasing function.

Proof: Considering the dynamics for the weight matrix $\tilde{W}_1(t)$ and the following candidate Lyapunov function $V_w(t)$.

$$
V_w(t):=\frac{1}{2}tr\left\{\tilde{W}_1^{\intercal}(t)\,\tilde{W}_1(t)\right\}+\frac{c}{4}\left[k_1(t)-k_{1\min}\right]_+^2
$$

where

$$
[z(t)]_+:=\left\{\begin{array}{cc} z(t) & z(t)\geq 0 \\ 0 & z(t)<0\end{array}\right.
$$

one has

$$
\dot{V}_w(t):=tr\left\{\tilde{W}_1^{\intercal}(t)\left(\dot{W}_1(t)\right)\right\}+2^{-1}c\dot{k}_1(t)\left[k_1(t)-k_{1\min}\right]_+^2
$$

By (12) it follows

$$
\begin{array}{c}
\dot{V}_w(t)=tr\left\{\tilde{W}_1^{\intercal}(t)\left(-\frac{k_1^{-1}(t)}{2}\left[P\Omega(t)\sigma^{\intercal}\left(\hat{x}(t)\right)-\dot{k}_1(t)\tilde{W}_1(t)\right]\right)\right\}+ \\
2^{-1}c\dot{k}_1(t)\left[k_1(t)-k_{1\min}\right]_+\leq \\
\frac{k^{-1}(t)}{2}\left|tr\left\{\tilde{W}_1^{\intercal}(t)\,P\Omega(t)\sigma^{\intercal}\left(\hat{x}(t)\right)\right\}\right|+ \\
2^{-1}\dot{k}_1(t)\left(k_1^{-1}(t)\,tr\left\{\tilde{W}_1^{\intercal}(t)\,\tilde{W}_1(t)\right\}+2^{-1}c\left[k_1(t)-k_{1\min}\right]_+\right)
\end{array}
$$

The property $\dot{V}_w(t)\leq 0$ results from (14).It is worth notice that the learning law (12) and (13) must be realized on-line in parallel with the gain-parameter adaptation procedure (14). ■

B. Main theorem on an upper bound for the observation error

Hereafter we will assume that

A1) The class of the function $f:\Re^n\to\Re^n$ is Lipschitz continuous in $x\in X$, that is, for all $x,x'\in X$ there exist constants $L_{1,2}$ such that

$$
\begin{array}{c}
\left\|f(x,u,t)-f(y,v,t)\right\|\leq \\
L_1\left\|x-y\right\|+L_2\left\|u-v\right\|; \\
\left\|f(0,0,t)\right\|^2\leq C_1; \\
x,y\in\Re^n;\ u,v\in\Re^m;\ 0\leq L_1,L_2<\infty
\end{array}
\tag{15}
$$

A2) The pair (A,C) is observable, that is, there exists the gain matrix $K\in\Re^{n\times m}$ such that matrix

$$
\tilde{A}(K):=A-KC \tag{16}
$$

is stable.

A3) The noises ξ_t and η_t in the system (1) are uniformly (on t) bounded such that

$$
\left\|\xi(t)\right\|_{\Lambda_\xi}^2\leq\Upsilon_\xi,\ \ \left\|\eta(t)\right\|_{\Lambda_\eta}^2\leq\Upsilon_\eta \tag{17}
$$

where Λ_ξ and Λ_η are known "normalizing" non-negative definite matrices which permit to operate with vectors having components of different physical nature (for example, meters, $mole/l$, voltage and etc.).

Theorem 2: Under assumptions A1-A3 and if there exist matrices $\Lambda_i=\Lambda_i^{\intercal}>0$, $\Lambda_i\in\Re^{n\times n}$, $i=1\ldots 10$, $Q_0\in\Re^{n\times n}$, $K\in\Re^{n\times m}$ and positive parameters ϖ, μ_1,μ_2 and μ_3 such that the following LMI

$$
\left[\begin{array}{cccc}
LMI_1 & 0 & 0 & 0 \\
0 & LMI_2 & 0 & 0 \\
0 & 0 & LMI_3 & 0 \\
0 & 0 & 0 & LMI_4
\end{array}\right]>0 \tag{18}
$$

with:

$$LMI_1 := \begin{bmatrix} -\Gamma(K,\varpi,\mu_1,\mu_2) & P \\ P & R \end{bmatrix}$$

$$LMI_2 := \begin{bmatrix} \Theta_1 & \tilde{A}^\top(K)P \\ P\tilde{A}(K) & \mu_1 P \end{bmatrix}$$

$$LMI_3 := \begin{bmatrix} \Theta_2 & \hat{W}_1^\top(K)P \\ P\hat{W}_1 & \mu_2 P \end{bmatrix}$$

$$LMI_4 := \begin{bmatrix} \Theta_3 & \hat{W}_2^\top(K)P \\ P\hat{W}_2 & \mu_3 P \end{bmatrix}$$

where $\mathrm{tr}\{\Theta_i\} < 1,\ i = 1,2,3$ and

$$\Gamma(K,\delta,\mu_1,\mu_2) = \left[\tilde{A}^\top(K)P + P\tilde{A}(K) + Q(\delta,\mu_1,\mu_2,\mu_3)\right]$$

$$R^{-1} = \Lambda_1^{-1} + \Lambda_9^{-1} + \Lambda_{10}^{-1} + \hat{W}_1\Lambda_5^{-1}\left(\hat{W}_1\right)^\top + \hat{W}_2\Lambda_8^{-1}\left(\hat{W}_2\right)^\top$$

$$Q(\delta,\mu_1,\mu_2,\mu_3) = \left[\|\Lambda_5\| L_\sigma + \|\Lambda_8\| L_\varphi \Upsilon_u^2 + \mu_1 + \mu_2 L_\sigma + \mu_3 \Upsilon_u^2 L_\varphi\right] I + \varpi\left(\Lambda_3^{-1} + \Lambda_7^{-1}\right) + Q_0$$

has positive definite solution P, then the projectional DNN observer 11 with the weight's learning laws, given by (12),(13),(14), and with $h(t)$ satisfying

$$\lim_{t\to\infty} h(t) \to \varepsilon, 0 < \ \varepsilon << 1 \tag{19}$$

provides the following upper bound for the "averaged estimation" error

$$\begin{aligned} &\overline{\lim_{T\to\infty}} \frac{1}{T}\int_{\tau=0}^{T}\left(\delta^\top(\tau-h(\tau))Q_0\delta(\tau-h(\tau))\right)d\tau \le \\ &\|\Lambda_9\|\left(\left(\|K\|\left\|\Lambda_\eta^{-1}\right\|^{1/2}\Upsilon_\eta + \left\|\Lambda_\xi^{-1}\right\|^{1/2}\Upsilon_\xi\right)\right)^2 \\ &+ \|\Lambda_{10}\|\left\|\Lambda_{\tilde{f}}^{-1}\right\|\left[\tilde{f}_0 + \tilde{f}_1\|x(t)\|^2_{\Lambda^1_{\tilde{f}}}\right] \\ &+ \|K\|^2\|P\|\left\|\Lambda_\eta^{-1}\right\|^{1/2}\Upsilon_\eta \\ &+ \|P\|\left\|\Lambda_{\tilde{f}}^{-1}\right\|\left[\tilde{f}_0 + \tilde{f}_1\left\|\Lambda^1_{\tilde{f}}\right\|\mathrm{Diam}(x)^2\right] \\ &+ \|P\|\left\|\Lambda_\xi^{-1}\right\|\Upsilon_\xi + 2\Upsilon_\eta \end{aligned} \tag{20}$$

where: $\delta(t\text{-}h(t)) := \hat{x}(t\text{-}h(t))\text{-}x(t\text{-}h(t))$, $\mathrm{Diam}(x) = \sup_{x,z\in X}\|x\text{-}z\|$ and $Q_0 > 0$

The proof of this theorem is briefly exposed in the appendix A.

Remark 2: It is easy to see that in the absence of noises ($\eta_t = \xi_t = 0$) and unmodelled dynamics ($\tilde{f} = 0$), we can choose $\tilde{f}_0, \tilde{f}_1, \Upsilon_\xi$ and Υ_η such that:

$$\overline{\lim_{T\to\infty}} \frac{1}{T}\int_{\tau=0}^{T}\left(\delta^\top(\tau-h(\tau))Q_0\delta(\tau-h(\tau))\right)d\tau \to 0$$

V. Numerical Example

As it follows from the presentation above, to realized the suggested approach one needs to fulfill the following steps:1) Define the projector, 2) Select Matrices A and $\hat{W}$ (some hints are given by [17] and [18]),3) Select K such that $A - KC$ is stable, with C defined by the output of the system, 4) Find P as the solution of the LMI problem (18).5) Introduce P into the adapting weight law (12), (13) and (14) and realized it on-line.

Example 2: The next simplified model (21) describes the ozonization process when a contaminant is present in a soil just with solid and gas phases involved [19]. It is worth notice that the model is employed only as a data source, any structural information has been used in the projectional DNN observer design.

$$\left.\begin{aligned} V_{gas}\dot{x}_{1,t} = V_{gas}^{-1}\left[W_{gas}C_\tau^{in} - W_{gas}x_{1,t} - k_1 S_1 x_{4,t}x_{3,t}\right. \\ \left. -K_t^{abs}\left(Q_{\max}^{free\_abs} - x_{2,t}\right)\right] \\ \dot{x}_2 = K_t^{abs}\left(Q_{\max}^{free\_abs} - x_{2,t}\right) \\ \dot{x}_{3,t} = k_1 S_1 x_{4,t}x_{3,t},\ \dot{x}_{4,t} = -k_1 G^{-1}x_{4,t}x_{3,t} \end{aligned}\right\} \tag{21}$$

Here in (21) $y_t = x_{1,t} + \eta_t$ (see Figure 2) is the ozone concentration (mole/L) at the output of the reactor assumed to be measurable, $x_{2,t}$ (mole) is the ozone amount absorbed by the soil which is not reacting with the contaminant, $x_{3,t}$ (mole) is the ozone amount absorbed by the soil and reacting with the contaminant, and $x_{4,t}$ (mole/g) is the current contaminant concentration. The convex compact set X according to the physical system constrictions is given as:

$$X := \left\{\begin{array}{c} 0 \le x_{1,t} \le x_{1,0} \\ 0 \le x_{2,t} \le Q_{\max}^{free\_abs} \\ 0 \le x_{3,t} \le V_{gas}C^{in} \\ 0 \le x_{4,t} \le x_{4,0} \end{array}\right\}$$

and the corresponding observer parameters are defined by:

$$A = \begin{bmatrix} -2.6 & 0 & 0 & 0 \\ 0 & -1.6 & 0 & 0 \\ 0 & 0 & -2.24 & 0 \\ 0 & 0 & 0 & -0.46 \end{bmatrix}, K = \begin{bmatrix} 0.01 \\ 0.01 \\ -0.0001 \\ -0.1 \end{bmatrix}$$

Figures 1,and 2 represent the results of the x_3 and x_4 estimation from the output, comparing the projectional DNN observer against a DNN observer without projection operator in presence of "quasi-white noise" $\eta(t)$ (amplitude $= 0.6 \times 10^{-5}$) and with the same initial conditions in both cases.

VI. Conclusion

The suggested approach related to the DNN-projectional state estimate problem for a special class of partially unknown nonlinear system demonstrates good results when the plant dynamics belongs to a given compact set (assumed to be known *a priori*), even when external perturbations are essensial. The complete convergence analysis for this class of adaptive observer is presented. Also the boundedness property of the adaptive weights in DNN is proven. Since the projection method leads to discontinuous trajectories in the estimated states, a nonstandard Lyapunov - Krasovski functional is applied to derive the upper bound for estimation error (in "average sense"), which depends on a noise power (output and dynamics disturbances) and on an unmodelled dynamic. It is shown that the asymptotic stability is attained when both of these uncertainties are absent. The illustrative example confirms the advantages which the suggested of observers have being compared with traditional ones.

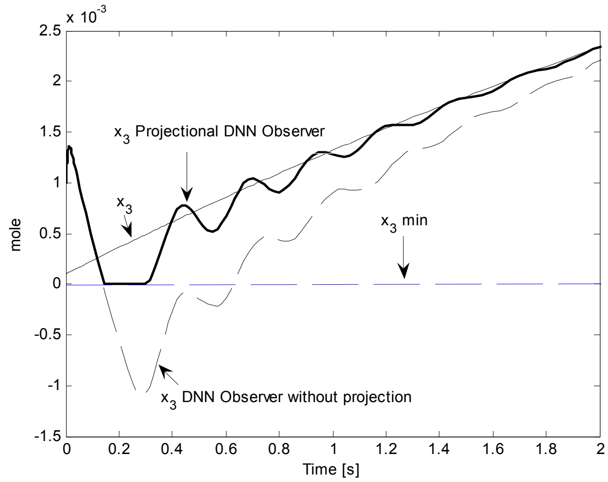

Fig. 1. Estimation of $x_3(t)$ (2 seconds)

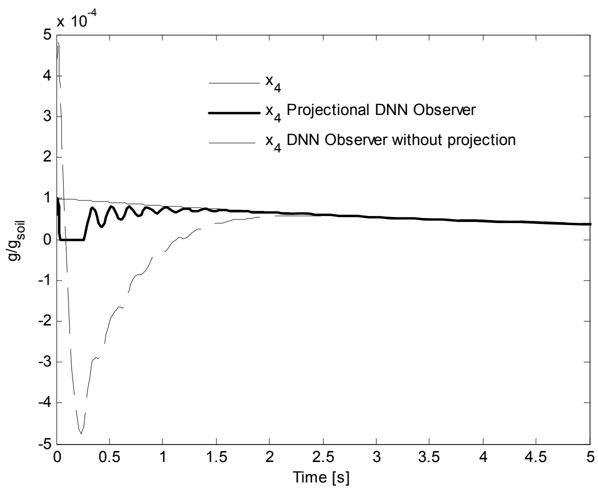

Fig. 2. Estimation of $x_4(t)$ (5 seconds)

REFERENCES

[1] A. Radke and Z. Gao, "A survey of state and disturbance observers for practitioners," *Proceedings of the American Control Conference Mineapolis, Minesota, USA*, pp. 5183–5188, 2006.

[2] H. Knobloch, A. Isidori, and D. FLocherzi, "Topics in control therory," *Birkhauser Verlag, Basel-Boston-Berlin*, 1993.

[3] H. Zak and B. L. Walcott, *State observation of nonlinear control systems via the method of Lyapunov*. in Zinober, A.S.I. (ed.), 1990, ch. Deterministic Control of Uncertain Systems, pp. 333–350.

[4] S. Singhal and L. Wu, "Training multilayer perceptrons with the extended kalman algorithm," in *Advances in Neural Information Processing Systems I*. San Mateo: Morgan Kaufmann, 1989, pp. 133–140.

[5] W. Wang and Z. Gao, "A comparison study of advanced state observer design techniques," *Proceedings of the American Control Conference*, pp. 4754–4759, 2003.

[6] T. Alamo, J. Bravo, and E. Camacho, "Guaranteed state estimation by zonotopes," *Automatica*, vol. 41, pp. 1035–1043, 2005.

[7] S. Haykin, *Neural Networks. A comprehensive Foundation*. New York: IEEE Press, 1994.

[8] A. F., T. A., and P. R., "A stable neural network-based observer with application to flexible-joint manipulators," *IEEE Transactions on Neural Networks*, vol. 17, no. 1, pp. 118–129, January 2006.

[9] W. Haddad, J. Bailey, T. Hayakawa, and N. Hovakimnyan, "Neural network adaptive output feedback control for intensie care unit sedation and intraoperative anesthesia," *IEEE Transactions on Neural Networks*, vol. 18, pp. 1049–1065, 2007.

[10] N. Cotter, "The stone-weierstrass theorem and its application to neural networks," *IEEE Transactions on Neural Networks*, vol. 1990, pp. 290–295, December 1990.

[11] G. Rovithakis and M. Christodoulou, "Adaptive control of unknown plants using dynamical neural networks," *IEEE Trans. Systs., Man and Cyben.*, vol. 24, pp. 400–412, 1994.

[12] A. Poznyak, E. Sanchez, and W. Yu, *Differential Neural Networks for Robust Nonlinear Control (Identification, State Estimation an Trajectory Tracking*. World Scietific, 2001.

[13] F. L. Lewis, A. Yesildirek, and K. Liu, "Multilayer neural-net robot controller with guaranteed tracking performance," *IEEE Trans. Neural Netw.*, vol. 7, no. 2, pp. 1–11, 1996.

[14] M. Zeitz, "The extended luenberger observer for nonlinear systems," *Syst Control Lett.*, vol. 9, pp. 149–56, 1987.

[15] E. Yaz and A. Azemi, ""robust adaptive observers for systems having uncertain functions with unknown bounds",," in *Proceedings of Amer.Contr.Conf.*, 1994, pp. 73–74.

[16] A. Poznyak, *Variable structure systems: from principles to implementation*. IEE Control Engineering Series, 2004, ch. Deterministic output noise effects in sliding mode observation, pp. 45–80.

[17] I. Chairez, A. Poznayk, and T. Poznayk, "New sliding-mode learning law for dynamic neural network observer," *IEEE Transactions on Circuits and SystemsŮII*, vol. 53, pp. 1338–1342, 2006.

[18] V. Stepanyan and N. Hovakimyan, "Robust adaptive observer design for uncertain systems with bounded disturbances," *IEEE Transactions on Neural Networks*, vol. 18, no. 5, pp. 1392–1403, 2007.

[19] T. Poznyak, A. García, I. Chairez., M. Gómez, and A. Poznyak, "Application of the differential neural network observer to the kinetic parameters identification of the anthracene degradation in contaminated model soil," *Journal of Hazardous Materials*, vol. 146, p. 661Ü667, 2007.

VII. APPENDIX

Proof: Consider the next "nonstandard" energy-like Lyapunov-Krasovkii funcational

$$V(t)=\int_{t-h(t)}^{t}\left[\|\delta(\tau)\|_p^2+k(\tau)\operatorname{tr}\left\{\tilde{W}^{\intercal}(\tau)\tilde{W}(\tau)\right\}\right]d\tau \tag{22}$$

where $\tilde{W}(\tau):=W(\tau)-\hat{W}$. Since the problem under consideration contains uncertainties and external output disturbances we won't demonstrate that the time-derivative of this energetic function is strictly negative. Instead, we will use it to obtain an upper bound for the averaged state estimation error. Taking time derivative of (22) one has

$$\begin{gathered}\dot{V}(t)=\|\delta(t)\|_p^2\text{-}\|\delta(t\text{-}h(t))\|_p^2(1\text{-}\dot{h}(t))+\\ k_1(t)\,tr\left\{\tilde{W}_1^{\intercal}(t)\tilde{W}_1(t)\right\}-\\ \left[k_1(t)\operatorname{tr}\left\{\tilde{W}_1^{\intercal}(t\text{-}h(t))\tilde{W}_1(t\text{-}h(t))\right\}\right](1\text{-}\dot{h}(t))+\\ k_2(t)\,tr\left\{\tilde{W}_2^{\intercal}(t)\tilde{W}_2(t)\right\}-\\ \left[k_2(t)\operatorname{tr}\left\{\tilde{W}_2^{\intercal}(t\text{-}h(t))\tilde{W}_2(t\text{-}h(t))\right\}\right](1\text{-}\dot{h}(t))\end{gathered}$$

let us define:

$$\begin{gathered}\tilde{A}:=A-KC,\\ \tilde{W}_i(t):=W_i(t)-\hat{W}_i\quad i=1,2\\ \tilde{\sigma}(t):=\sigma(\hat{x}(t))-\sigma(x(t))\\ \tilde{\varphi}(t):=\varphi(\hat{x}(t))-\varphi(x(t))\end{gathered}$$

Presenting the state estimation error δ_t as a function of the available output, the estimation error e_t can represented as:

$$\begin{gathered}\delta(t)=N_{\varpi}\left(-C^{\intercal}e(t)+C^{\intercal}\eta(t)+\varpi\delta(t)\right)\\ N_{\varpi}:=\left(C^{\intercal}C+\varpi I\right)^{-1}\end{gathered} \tag{23}$$

ϖ is a small positive scalar. Then, by the property (5), the assumption *A2-A3*, an upper bound for each involved term is determined:

$$\begin{gathered}\dot{V} \leq \\ h(t)\,\delta_{t-h(t)}^{\top}\left[\tilde{A}^{\top}(K)P + P\tilde{A}(K) + PR^{-1}P + \right. \\ \left. Q(\delta,\mu_1,\mu_2,\mu_3)\right]\delta_{t-h(t)} + \\ h(t)^3\left[L_{1h}(t)\right] + h(t)\left[L_{2h}(t)\right] + \\ +2\int_{\tau=t-h(t)}^{t}\left(e^{\top}(t-h(t))\,CN_{\varpi}P\tilde{W}_1(\tau)\,\sigma(\hat{x}(\tau))\right)d\tau \\ +\int_{\tau=t-h(t)}^{t}\left[\sigma^{\top}(\hat{x}(\tau))\,\tilde{W}_1^{\top}(\tau)\,PN_{\varpi}(C\Lambda_2C+\varpi\Lambda_3)\right. \\ \left. N_{\varpi}P\tilde{W}_1(\tau)\,\sigma(\hat{x}(\tau))\right]d\tau \\ +\int_{\tau=t-h(t)}^{t}\sigma^{\intercal}(\hat{x}_{\tau})\,\tilde{W}_1^{\intercal}(\tau)P\tilde{W}_{\tau}\sigma(\hat{x}_{\tau})\,d\tau + \\ k_1(t)\,\mathrm{tr}\left\{\tilde{W}_1^{\intercal}(t)\,\tilde{W}_1(t)\right\} - \\ k_1(t-h(t))\,\mathrm{tr}\left\{\tilde{W}_1^{\intercal}(t-h(t))\tilde{W}_1(t-h(t))\right\} + \\ \int_{\tau=t-h(t)}^{t}2\left(e^{\top}(t\text{-}h(t))\,CN_{\varpi}P\tilde{W}_2(\tau)(\varphi(\hat{x}(\tau))\,u(\tau)\right)d\tau + \\ \int_{\tau=t-h(t)}^{t}\left[u^{\top}(\tau)\varphi^{\top}(\hat{x}(\tau))\,\tilde{W}_2^{\top}(\tau)\,PN_{\varpi}(C^{\intercal}\Lambda_6C+\varpi\Lambda_7)\right. \\ \left. N_{\varpi}P\tilde{W}_2(\tau)(\varphi(\hat{x}(\tau))\,u(\tau)\right]d\tau + \\ \int_{\tau=t-h(t)}^{t}u^{\intercal}(\tau)(\varphi(\hat{x}(\tau))^{\intercal}\,\tilde{W}_2^{\top}(\tau)P\tilde{W}_2(\tau)(\varphi(\hat{x}(\tau))\,u(\tau)d\tau + \\ k_2(t)\,\mathrm{tr}\left\{\tilde{W}_2^{\intercal}(t)\,\tilde{W}_2(t)\right\} - \\ k_2(t-h(t))\,\mathrm{tr}\left\{\tilde{W}_2^{\intercal}(t-h(t))\tilde{W}_2(t-h(t))\right\}\end{gathered}$$

Where:

$$\begin{gathered}Q(\delta,\mu_1,\mu_2,\mu_3) := \\ \left[\|\Lambda_5\|\,L_{\sigma} + \|\Lambda_8\|\,L_{\varphi}\Upsilon_u^2 + \mu_1 + \mu_2L_{\sigma} + \mu_3\Upsilon_u^2L_{\varphi}\right]I \\ +\varpi\left(\Lambda_3^{-1}+\Lambda_7^{-1}\right) + Q_0 \\ L_{1h}(t) = \|\Lambda_1\|\left\|\tilde{A}\right\|^2\frac{L_{\delta}^2}{4} + \|\Lambda_5\|\,\frac{L_{\sigma}L_{\delta}^2}{3} + \\ \mu_2\tfrac{L_{\sigma}L_{\delta}^2}{3} + \mu_1\tfrac{L_{\delta}^2}{3} + \mu_3\tfrac{\Upsilon_u^2L_{\varphi}L_{\delta}^2}{3} + \|\Lambda_8\|\,\frac{L_{\varphi}\Upsilon_u^2L_{\delta}^2}{3}\end{gathered}$$

$$\begin{gathered}R^{-1} := \Lambda_1^{-1} + \hat{W}_1\Lambda_5^{-1}\left(\hat{W}_1\right)^{\top} + \\ \hat{W}_2\Lambda_8^{-1}\left(\hat{W}_2\right)^{\top} + \Lambda_9^{-1} + \Lambda_{10}^{-1} \\ L_{2h}(t) := \\ \|\Lambda_9\|\left(\left(\|K\|\left\|\Lambda_{\eta}^{-1}\right\|^{1/2}\Upsilon_{\eta} + \left\|\Lambda_{\xi}^{-1}\right\|^{1/2}\Upsilon_{\xi}\right)\right)^2 + \\ \|\Lambda_{10}\|\left\|\Lambda_{\tilde{f}}^{-1}\right\|\left[\tilde{f}_0 + \tilde{f}_1\,\|x(t)\|_{\Lambda_{\tilde{f}}^1}^2\right] + \|P\|\left\|\Lambda_{\xi}^{-1}\right\|\Upsilon_{\xi} + \\ 2\Upsilon_{\eta} + \|K\|^2\,\|P\|\left\|\Lambda_{\eta}^{-1}\right\|^{1/2}\Upsilon_{\eta} - \delta_{t-h(t)}^{\top}Q_0\delta_{t-h(t)} + \\ \|P\|\left\|\Lambda_{\tilde{f}}^{-1}\right\|\left[\tilde{f}_0 + \tilde{f}_1\left\|\Lambda_{\tilde{f}}^1\right\|\mathrm{Diam}(x)^2\right]\end{gathered}$$

Considering:

$$\begin{gathered}\tilde{A}^{\top}(K)\,P + P\tilde{A}(K) + \\ PR^{-1}P + Q(\delta,\mu_1,\mu_2,\mu_3) \leq 0\end{gathered}$$

and the adaptation laws (12)(13) Finally we get:

$$\dot{V} \leq h(t)\left(h(t)^2\,a + b - \delta^{\intercal}(t-h(t))\,Q_0\delta(t-h(t))\right) \tag{24}$$

where

$$\begin{gathered}a := \|\Lambda_1\|\left\|\tilde{A}\right\|^2\frac{L_{\delta}^2}{4} + \|\Lambda_5\|\,\frac{L_{\sigma}L_{\delta}^2}{3} + \\ \mu_2\tfrac{L_{\sigma}L_{\delta}^2}{3} + \mu_1\tfrac{L_{\delta}^2}{3} + \mu_3\tfrac{\Upsilon_u^2L_{\varphi}L_{\delta}^2}{3} + \|\Lambda_8\|\,\frac{L_{\varphi}\Upsilon_u^2L_{\delta}^2}{3} \\ b := \|\Lambda_9\|\left(\left(\|K\|\left\|\Lambda_{\eta}^{-1}\right\|^{1/2}\Upsilon_{\eta} + \left\|\Lambda_{\xi}^{-1}\right\|^{1/2}\Upsilon_{\xi}\right)\right)^2 + \\ +\|\Lambda_{10}\|\left\|\Lambda_{\tilde{f}}^{-1}\right\|\left[\tilde{f}_0 + \tilde{f}_1\,\|x(t)\|_{\Lambda_{\tilde{f}}^1}^2\right] \\ +\|K\|^2\,\|P\|\left\|\Lambda_{\eta}^{-1}\right\|^{1/2}\Upsilon_{\eta} + \\ \|P\|\left\|\Lambda_{\tilde{f}}^{-1}\right\|\left[\tilde{f}_0 + \tilde{f}_1\left\|\Lambda_{\tilde{f}}^1\right\|Diam(x)^2\right] + \\ \|P\|\left\|\Lambda_{\xi}^{-1}\right\|\Upsilon_{\xi} + 2\Upsilon_{\eta}\end{gathered}$$

So,

$$\delta^{\intercal}(t-h(t))\,Q_0\delta(t\text{-}h(t)) \leq \left(ah(t)^2 + b\right) - \frac{\dot{V}}{h(t)}$$

and integrating (24) we derive

$$\begin{gathered}\int_{\tau=0}^{T}\delta^{\intercal}(\tau-h(\tau))\,Q_0\delta(\tau-h(t)(\tau))\,d\tau \leq \\ \int_{\tau=0}^{T}\left[\left(ah(\tau)^2+b\right) - \tfrac{\dot{V}}{h(\tau)}\right]d\tau\end{gathered}$$

This implies

$$\begin{gathered}\int_{\tau=0}^{T}\delta^{\intercal}(\tau-h(t)(\tau))\,Q_0\delta(\tau-h(t)(\tau))\,d\tau \leq \\ a\int_{\tau=0}^{T}h(t)^2\,d\tau + bT + \frac{V_0}{h(0)}\end{gathered}$$

Dividing by T and taking the upper we finally get (20). ■