Adaptive-gain extended Kalman filter: application to a series connected DC motor

Nicolas Boizot*† Eric Busvelle† Jean-Paul Gauthier‡ Juergen Sachau*

* *LASSY, Université du Luxembourg, 6 rue Coudenhove Kalergi, L-1359 Luxembourg nicolas.boizot@uni.lu*

† *LE2I, IUT d'Auxerre, Route des plaines de l'Yonne, 89000 Auxerre, busvelle@u-bourgogne.fr*

‡ *LE2I, Université de Bourgogne, 21000 Dijon, gauthier@u-bourgogne.fr*

***Abstract*— In this paper, we present an adaptative extended Kalman filter (A-EKF). The exponential convergence of this observer is theoretically proven. An application on a series-connected DC motor is given to illustrate the performance of this A-EKF.**

***Index Terms*— Adaptive observer, Kalman, DC motor**

I. Introduction

The extended Kalman filter (EKF) is one of the most famous algorithms used to estimate unknown state variables from measurements in dynamical nonlinear systems. It is also used to estimate unknown constant or slowly varying parameters in linear systems and sometimes to perform failure detection. In this last case, it is necessary to quantify the efficiency of the EKF with time. This task is usually based on the innovation process, which is the integrated difference between actual measurements and predicted measurements. The innovation process can be monitored, and a large value of the innovation can be used to send an alarm or to switch from an old model to a new one. It can also be used to estimate the noise entering into the process or to estimate the measurement noise.

Although the extended Kalman filter is not proved to converge (except locally, see [3, 13]), it is very often used, even for critical processes. In order to increase the performance and the reliability of the EKF several engineers and researchers already tried to develop an adaptive version. Using innovation and state estimation, it seems possible to estimate parameters that characterize the state of the process. These parameters can then be used to adapt the gain matrix by online automatic tuning of some of the covariance matrices used in the computation of the gain matrix. These kind of adaptive EKF are empirical but seem to have nice behavior compared to the EKF.

Because of the difficulty to ensure robustness when adaptive quantity is continuously updated, some authors used an adaptive algorithm based on switching between several models. For instance, in [17], authors have developed an application on a highly critical process (from robustness point of view). They proposed to switch between two covariances matrix Q_1 and Q_2 depending on the state of the process.

There exist a very large number of papers dealing with adaptive observers and adaptive extended Kalman filtering ([4, 5, 15]) especially in the GPS and DGPS community ([6, 14, 16]). In [6] for instance, authors present an adaptive extended Kalman filter using innovation in order to adapt Q and R matrices, exactly in the same spirit than in the present paper, except that they do not give any theoretical proof. Nevertheless, the need for this kind of observer is clearly established.

In those papers, adaptation of the filter is done using empirical rules (genetic algorithms [18], neural networks [20], statistics [17]...), and no proofs are given. But in all cases, efficiency of the adaptive observer is highlighted. Let us remark that for neural networks based extended Kalman filters (N-EKF), the system is splitted into a linear part and a nonlinear part, and the extended kalman filter is applied to the nonlinear part, which is approximated by neurons. The weights of neurons can be calculated using EKF, making the algorithm adaptive. In this case, some proofs can be established, but only if the neural network can approximate the system.

An intuitive theoretical justification of adaptive gain is based on the high gain observer theory. It has been shown from a long time ([12]) that high gain observers have very nice theoretical properties. The first one is that they required to study the observability property of the model. This study prevents from developing an observer for a non-observable system. But high gain observers are also exponential observers: one can prove the convergence of the high gain observer. In our opinion, the convergence property is a minimum requirement for an observer which is used on some critical processes, and sometimes as a diagnostic tool. Therefore, it is a good idea to adapt the gain of observers in the following way:

- use an EKF when the estimation is close to the true state, because EKF is a good (optimal) local observer (as already stated) and
- use a high-gain observer when large perturbations occur, because these observers are nonlinear converging observers.

In several previous papers [10, 11, 13], we have introduced the high-gain extended Kalman filter (HG-EKF) which is also an exponentially converging observer, but with the property that it is more efficient in the presence of noise. Indeed, the high sensitivity of high-gain observers is a well known drawback : the high gain ensures convergence but also increases noise effects. In [7], we developed a new algorithm, based on classical and high-gain EKF. This algorithm is based on a theoretical result, which states that a time-dependant HG-EKF, which is asymptotically equivalent to a classical EKF, may be an exponentially converging observer, if the transition from HG-EKF to EKF is slow enough. But this result is based on a time-dependant observer and, in order to make its convergence property persistent, it is necessary to use several observers and to switch from one to another, depending on the innovation process. Although it is an efficient observer, as shown in the reference above, but also in [8, 9], it is rather complicated and CPU intensive. Moreover, even if the final algorithm can be considered as an adaptive high-gain extended Kalman filter (A-EKF), its implementation is far from the one of classical observers as used by engineers.

In this paper, we will present a time-independant adaptive-gain extended Kalman filter. The adaptation of θ will depend on the innovation process.

As usual for the HG-EKF, the parameter θ appears in the Riccati equation of the Kalman filter, and more precisely in the matrix Q, denoted Q_θ. But in our case, the high-gain parameter appears also in the matrix R (denoted R_θ), as in [6] (for a practical application). It is the first difference with our previous result [7]. The second difference is that θ may increase if the innovation is high and decrease if the innovation is low. This idea is the basis of practical applications: it is also the cornerstone of the proof of our theorem.

In the next section (II), we will present our main result. Because of the length of this paper, we will omit the proof. Then, in section III, we will present an application to a series-connected DC motor.

II. Theoretical result

In this paper, we consider systems of the form

$$\left\{\begin{array}{lcl} \frac{dx}{dt} & = & A(u)\,x + b(x,u) \\ y & = & C(u)\,x \end{array}\right. \tag{II.1}$$

where $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}$ and $u \in \mathcal{U}_{\mathrm{adm}} \subset \mathbb{R}^d$. Moreover, $A(u)$, $b(x,u)$, and $C(u)$ are defined as is usual for such systems in the canonical form of observability:

$$A(u) = \begin{pmatrix} 0 & a_2(u) & 0 & \cdots & 0 \\ & & a_3(u) & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & & a_n(u) \\ 0 & & \cdots & & 0 \end{pmatrix}$$

with $0 < a_m \le a_i(u) \le a_M$ for any u in $\mathcal{U}_{\mathrm{adm}}$,

$$b(x,u) = \begin{pmatrix} b(x_1,u) \\ b(x_1,x_2,u) \\ \vdots \\ b(x_1,\ldots,x_n,u) \end{pmatrix}$$

and we suppose that there exists a compact subset of $\mathbb{R}^n$ such that $b(x,u)$ is identically equal to zero for x outside this compact set, and finally

$$C(u) = (a_1(u), 0, \cdots, 0).$$

This form is now conventional, and was introduced in [12] and extensively studied in [13] for instance. Recall that although the form of the system seems to be very restrictive, every system which has the observability for any input property can be written in this form, using a change of coordinates. Some other hypothesis (bounds, compactness, etc...) have been discussed in previous papers (see [12]) and are not very restrictive for practical applications, as it will appear in the application part of our paper. We will not give technical details here, since proofs are omitted.

Because of its canonical form of observability, it is possible to apply an extended Kalman filter to this system:

$$\left\{\begin{array}{lcl} \frac{dZ}{dt} & = & A(u)Z + b(Z,u) + S^{-1}C'R^{-1}(CZ - y(t)) \\ \frac{dS}{dt} & = & -(A(u) + b(Z,u))'S - S(A(u) + b^*(Z,u)) \\ & & +C'R^{-1}C - SQS \end{array}\right. \tag{II.2}$$

where, from a deterministic point of view, Q and R are tuning matrices and are used to ensure good performance. However the Kalman filter has a stochastic meaning, where Q is the covariance matrix of the state noise and R is the covariance matrix of the measurement noise, i.e. if the system is written in a stochastic form

$$\left\{\begin{array}{lcl} dx_t & = & A(u)\,x_t dt + b(x_t,u)\,dt + dw_t \\ dy_t & = & C(u)\,x_t dt + dv_t \end{array}\right.$$

then $Q = E\left[w_t w_t^T\right]$ and $R = E\left[v_t v_t^T\right]$. Hence, both matrices depend on statistical properties of noise processes. In this context, the Kalman filter is an optimal filter. Therefore, the extended Kalman filter is close to an optimal filter when the state and its estimation are close, since it is based upon a linearization of the system around its estimated trajectory. As a matter of fact, the extended Kalman filter is a converging *local* observer (see [3, 7]).

However, the EKF is not a globally converging observer. It cannot be used to estimate the state from a poor *a priori* estimation, or when large unmodeled perturbations occurs. In this case, and in order to keep theoretical justifications and proof of convergence, one can use a high-gain observer ([12, 13]). The simplest form of a high-gain observer is the Luenberger high-gain observer: it is a constant gain observer of the form

$$\frac{dZ}{dt} = A(u)Z + b(Z,u) - K_\theta(C(u)\,Z - y(t))$$

where $K_\theta = \Delta_\theta K$, $\Delta_\theta = \mathrm{diag}\left(\theta, \theta^2, \ldots, \theta^n\right)$, and θ is the high-gain parameter. The theorem says that with some hypotheses on the system, the input, and the matrix K when θ is large enough the observer converges exponentially. The same results hold for the high-gain extended Kalman filter

$$\left\{\begin{array}{lcl} \frac{dZ}{dt} & = & A(u)Z + b(Z,u) + S^{-1}C'R^{-1}(CZ - y(t)) \\ \frac{dS}{dt} & = & -(A(u) + b(Z,u))'S - S(A(u) + b^*(Z,u)) \\ & & +C'R^{-1}C - SQ_\theta S \end{array}\right. \tag{II.3}$$

where $Q_\theta = \Delta_\theta Q \Delta_\theta$. This high-gain observer was introduced in [10, 11] and was developed in [13]. The basis of our work consists of the following two remarks:

1) if one sets θ to 1 in system II.3 then one obtains the classical extended Kalman filter, which is a local optimal observer (in the sense explained above)

2) if θ is large enough then one obtains a high-gain observer, which is a global exponential observer.

The first application of this remark was presented in [7]: in this paper, we just added the equation

$$\frac{d\theta}{dt} = \lambda\left(1-\theta\right) \tag{II.4}$$

to the system II.3. If θ is large enough (and the parameter λ small enough) then we obtain an observer which is a high-gain observer for small time and which converges asymptotically to a classical extended Kalman filter. Hence we can expect its convergence since the observer should converge exponentially to the state (high-gain observer property) and then stays in a neighborhood of the state (since extended Kalman filter is a local observer). Indeed this result has been proved in [7]. But this observer is time-dependant, since it converges exponentially only in the beginning. So, in order to construct a persistent observer, we should take into account this property. The simplest way is to use several observers of the form II.3-II.4, each one initialized at different times, and using some delays between each initialization. Thus we obtain several estimations of the state, given by each one of the observers: the final estimation is the one corresponding to the observer that minimizes the innovation process. The whole construction is clearly explained in [8, 7].

In this paper, we present a much simple observer. In place of equation II.4, we introduce the equation

$$\frac{d\theta}{dt} = F(\theta, \mathcal{I}) \tag{II.5}$$

where

$$\mathcal{I} = \int_{t-T}^{t} \left\| y(s) - \bar{y}_{t-T}(s) \right\|^2 ds = \left\| y - \bar{y}_{t-T} \right\|^2_{L^2(t-T,t)} \tag{II.6}$$

is the innovation from time $t-T$ to current time t. More precisely, in definition II.6, y represents the output, but $\bar{y}_{t-T}$ represents the prediction of the output from the state estimation at time $t-T$ (given by the observer, $Z\left(t-T\right)$): hence $\bar{y}_{t-T}(s)$ is the solution at time s of

$$\left\{ \begin{array}{rcl} \frac{d\xi}{d\tau} & = & A(u)\xi\left(\tau\right) + b(\xi\left(\tau\right), u) \\ \xi\left(t-T\right) & = & Z\left(t-T\right) \\ \bar{y}_{t-T}\left(\tau\right) & = & C\left(u\right)\xi\left(\tau\right) \end{array} \right.$$

T is a tuning parameter, representing the length of the window used to calculate the innovation. In the following theorem, the function F will be chosen in the form $F\left(\theta, \mathcal{I}\right) = \lambda\left(1-\theta\right) + \mu\left(\theta_{\max} - \theta\right)$. In fact, F can be chosen in a more general form. We will describe a more general hypothesis on F in a next paper. We will also give a version of F that is better adapted in the presence of noise in the application part of this paper. Intuitively, the role of the function F is:

- to let θ decrease if the innovation is small, because in this case the observer has already converged and a Kalman-like observer will be sufficient to correctly estimate the state
- to let θ increase if the innovation is too large, because in this case, the observer gives a bad estimation of the state and θ has to be large enough in order to ensure convergence, thanks to the exponential property of high-gain observers.

Finally, the adaptive extended Kalman filter can be written

$$\left\{ \begin{array}{rcl} \frac{dZ}{dt} & = & A(u)Z + b(Z,u) + S^{-1}C'R_\theta^{-1}(CZ - y(t)) \\ \frac{dS}{dt} & = & -(A(u) + b(Z,u))'S - S(A\left(u\right) + b^*\left(Z,u\right)) \\ & & +C'R_\theta^{-1}C - SQ_\theta S \\ \frac{d\theta}{dt} & = & \lambda\left(1-\theta\right) + \mu\left(\theta_{\max} - \theta\right) \end{array} \right. \tag{II.7}$$

We define Q_θ and R_θ from Q and R thanks to the matrix

$$\Delta = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{\theta} & 0 & & \vdots \\ 0 & 0 & \frac{1}{\theta^2} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \frac{1}{\theta^{n-1}} \end{pmatrix}$$

by $Q_\theta = \theta\Delta^{-1}Q\Delta^{-1}$ and $R_\theta = \theta^{-1}R$. Let us remark that this change of coordinates is different from the previous one as explained in [7].

Our main result is the following:

Theorem 1: Let us consider a system of the form II.1 in the canonical form of observability. We consider the adaptive-gain extended Kalman filter II.7. Let us suppose that λ, μ and $\theta_{\max}$ are three constant parameters such that λ is small enough, μ is large enough, and θ_0 is large enough. Then, II.7 is an exponentially converging observer.

We have not enough place to give a proof of this result. It is more or less an adaptation of the proof in [7]. The main difference is the fact that now, the matrix R depends on θ, which was not necessary when θ was only a decreasing parameter. The proof is then based on the following lemma:

Lemma 2: Let x_1^0, $x_2^0 \in \mathbb{R}^n$ and $u \in \mathcal{U}_{\text{adm}}$. Let us consider the outputs $y_1\left(t\right)$ and $y_2\left(t\right)$ of system II.1 with initial conditions respectively x_1^0 and x_2^0. The following condition (called persistant observability) holds:

$$\forall T > 0 \quad \forall u \in L_b^1\left(\mathcal{U}_{\text{adm}}\right) \quad \exists \lambda_T > 0$$

$$\left\| x_1^0 - x_2^0 \right\| \leq \frac{1}{\lambda_T} \int_0^T \left\| y_1\left(\tau\right) - y_2\left(\tau\right) \right\| d\tau$$

III. Application to a series-connected DC motor

The stator of a DC motor (also denoted *field*) is composed of an electromagnet or a permanent magnet that immerses the rotor in a magnetic field. The rotor (also denoted *armature*) is made of an electromagnet whose windings are connected to a commutator that switches the direction of the current flowing through it, inversing its polarity. The attraction/repeling behavior of magnets creates the rotating motion that is maintained by the commutations. A DC motor whose field and armature circuits are connected in series, and therefore fed by the same power supply, is refered to as *a series-connected DC motor.*

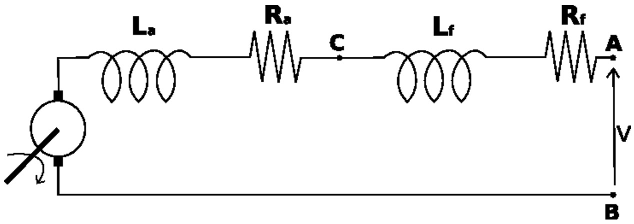

Fig. 1. Equivalent circuit

A. Motor model

The mathematical model for the series-connected DC motor is obtained from the equivalent circuit representation shown in fig.1. The state vector (x) is composed of the current flowing through the circuit (I) and the rotation speed of the shaft (ω_r). The electrical balance of the circuit is given by: $(L_a+L_f).\dot{I} = V_{AB} - (R_a+R_f).I - E$ where L_a and R_a, and L_f and R_f denote the inductance and the resistance of the armature, and of the field. E is the back electromotive force. The mechanical balance for the shaft in the hypothesis of vicscous friction and load is $J\dot{\omega_r} = T_{em} - B\omega_r - T_a$ where T_{em} denotes the electromechanical torque, T_a the (positive) load torque, and B the viscous friction coefficient. We consider the case of no magnetic saturation in the field circuit such that $T_{em} = K_m L_f I^2$ and $E = K_m L_f I \omega_r$ where K_m denotes the back e.m.f constant.

The voltage applied between terminals A and B is the input variable u, and the current I is the output variable y leading to the following SISO model:

$$\Sigma_1 \quad \begin{pmatrix} L\dot{I} \\ J\dot{\omega_r} \end{pmatrix} = \begin{pmatrix} u - RI - L_{af}\omega_r I \\ L_{af}I^2 - B\omega_r - T_a \end{pmatrix} \qquad y = I$$

where $L = L_a + L_f$, $R = R_a + R_f$ and $L_{af} = K_m L_f$. More detailed approaches can be found in [1, 2].

B. Observability

We propose identifying the load torque by means of observation when the load changes suddenly. When such a change occurs we expect the high-gain parameter to increase as a reaction to this perturbation. This identification is made considering the extended state vector (I, ω_r, T_a) and the simple equation $\dot{T_a} = 0$ for the load torque. The diffeomophism defined by

$$\begin{array}{c} \mathbb{R}^{+*}\times\mathbb{R}\times\mathbb{R} \to \mathbb{R}^{+*}\times\mathbb{R}\times\mathbb{R} \\ (I,\omega_r,T_a) \mapsto (I, I.\omega_r, I.T_a) = (x_1,x_2,x_3) \end{array}$$

transforms the model Σ_1 into the following SISO observability canonical form which implies observability for the extended system and therefore identifiability for the load torque[13].

$$\Sigma_2 \quad \begin{array}{l} \begin{pmatrix} \dot{x_1} \\ \dot{x_2} \\ \dot{x_3} \end{pmatrix} = \begin{pmatrix} 0 & -\frac{L_{af}}{L} & 0 \\ 0 & 0 & -\frac{1}{J} \\ 0 & 0 & 0 \end{pmatrix} . \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ + \begin{pmatrix} \frac{V}{L} - \frac{R}{L}x_1 \\ \frac{L_{af}}{J}x_1^3 + (\frac{V}{L}.\frac{1}{x_1} - \frac{L_{af}}{L}..\frac{x_2}{x_1} - \frac{R}{L} - \frac{B}{J}).x_2 \\ -\frac{L_{af}}{L}.\frac{x_2.x_3}{x_1} + \frac{V}{L}.\frac{x_3}{x_1} - \frac{R}{L}x_3 \end{pmatrix} \\ y = x_1 \end{array}$$

C. Observer construction

The adaption function we actually used to completely define the A-EKF is slightly different from the one presented in II:

$$F(\theta,\mathcal{I}) = \lambda(1-s(\mathcal{I})).(1-\theta) + K.s(\mathcal{I}).(\theta_{max}-\theta)$$

where $s(\mathcal{I}) = \left[1 + e^{-\beta(\mathcal{I}-m)}\right]^{-1}$, namely the sigmoid function, is defined from $\mathbb{R}$ to $]0;1[$, C^∞, increasing, and lipschitz. This function can be devided into three parts:

- values of $\mathcal{I}$ such that $s(\mathcal{I})$ is close to 0
- values of $\mathcal{I}$ such that $s(\mathcal{I})$ is close to 1
- a transition part that can be made as small as possible thanks to the parameter β

If this transition part is made small enough then $s(\mathcal{I})$ will be either close to zero or close to 1 but almost all the time. When $s(\mathcal{I})$ is close to zero the term $\lambda(1-\theta)$ will drive the evolution of θ which will decrease toward 1. Conversly when $s(\mathcal{I})$ is close to 1 then the term $K(\theta_{max}-\theta)$ will drive the evolution of θ which will increase toward θ_{max}. This adaption function introduces 5 new parameters (λ, K, m, β, θ_{max}) that need to be tuned in addition to the classical R and Q matrices. As we will explain in the next section, the choice of those parameters can easily be made.

D. Tuning of parameters

The procedure explained here is inspired by the one described in [8, part 5.2.2] regarding the parameters R, Q, and θ of a high-gain observer. We also think that only few of the additional parameters will actually need be reset at each new implementation of the observer: $\beta = 2000$, $K = \lambda = 500$, and $m = m_1 + m_2$ where $m_1 = 0.005$ may be kept to those values.

1) we determine the (symetric positive definite) matrices R and Q by using an EKF. This observer can be obtained from the A-EKF when the parameters of the adaptation function are set to 0 and the initial value of θ to 1. Large perturbations are not considered and the observer is initialised to the proper (or previously estimated) values of the state vector.
2) we then set the R and Q matrices to the values previously found and use a HG-EKF to tune θ. As above, the filter needed is obtained from our observer when the parameters of the adaption function are set to 0 and the initial value of θ is changed. Here we will try to find a value for the high-gain parameter that allows fast and reasonable convergence with respect to noise when large unmodeled perturbations are applied to the system.

3) we will now choose the parameters of the adaption function. We remark that when $m = 0$ then $s(0) = 0.5$. We then need to shift the sigmoid function to the right as we want $s(0)$ to be close to zero. Chosing y_1 as small as we want and solving the equation $s(0) = y_1$ allows us to obtain the parameter m. This solution is easily computed provided that the parameter β is known. As the sigmoid function is centered around $(0, 0.5)$ when $m = 0$, the computation of β is made by setting a length l for the transition part and solving the nonlinear equation (with $m = 0$): $s(l/2) - s(-l/2) = (1 - y_1) - y_1$. Of course, anoher approach is to graphically define β and m from trial and error.
Now that the transition part is small, we want the gain to increase and decrease quickly. If we suppose that $\theta(t) = 1$ and that we want it to reach θ_{max} within a time τ then the equation $\dot{\theta} = \frac{\theta_{max}-1}{\tau} = K.(\theta_{max} - 1)$ allows the computation of K. Since the equation used to compute K is only an approximation, a bigger value (e.g. twice the computed value) may be used. A reasonable choice for the last parameter remaining is $\lambda = K$. The delay parameter $T = 0.1$ is choosen with respect to the dynamic of the observed system and θ_{max} is taken equal to the value found in the previous stage.
One last thing has to be taken into account before continuing. Because of measurement noise the innovation will never be equal to zero and therefore the observer will stay in a high-gain mode. To avoid this problem, the parameter m is now written $m = m_1 + m_2$ where $m_1$is the previously computed quantity and $m_2$will represent the influence of the noise on the system. As a result, when $I \leq m_2$we have $s(\mathcal{I}) \leq y_1$ and θ does not increase. We denote by σ the standard deviation of the output noise, which can be estimated from output measurement. Then $m_2 = T\sigma^2$ where T is the delay used in the definition of the innovation.

E. Simulation results and discussion

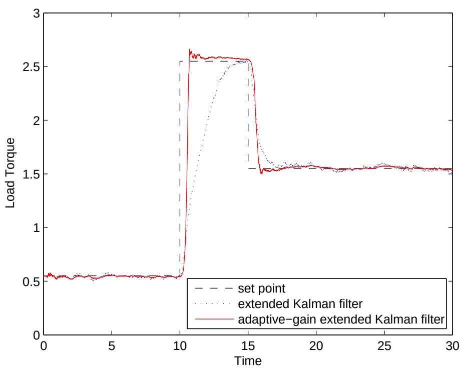

Fig. 2. Comparison with an EKF

We made our simulations with Matlab/Simulink. The system Σ_1 was used to simulate the motor and Σ_2, to implement the observer. The main equations of the observer were implemented in a continuous S-function while the computation of the innovation needed the use of a discrete S-function (to allow the computation of $\hat{y}(t)$ and of the integral). This implies that the choice of the sample time of this computation will have an impact on the behavior of the system as innovation drives the adaption function. Noise was added at both input and output (drift=0.95, sample time=10^{-4}, standard deviation=1 and 0.5 respectively) and to the load torque set points. The parameters of the model were set to $L = 1.22\ N$, $R = 5.4183\ \Omega$, $B = 0.0026\ Nm/rads$, $J = 1.22\ kgm^2$, and $L_{af} = 0.0683\ Nm/WbA$ (those values were motivated by measures made on a real system). The parameters of the observer were set to R=1, Q=[1 0 0; 0 1 0; 0 0 5], $\lambda = K = 500$, $\beta = 2000$, $m_1 = 0.005$,

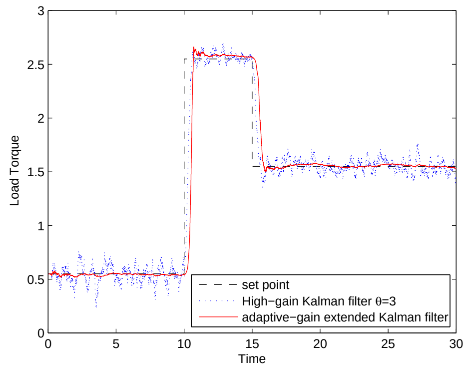

Fig. 3. Comparison with a HG-EKF

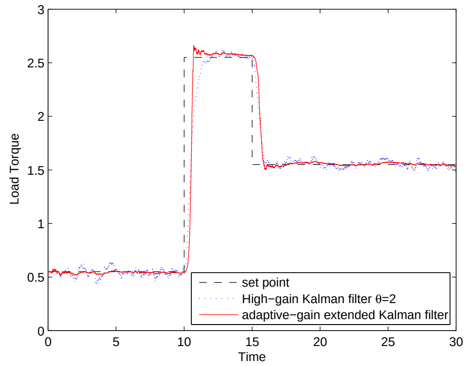

Fig. 4. Comparison with a HG-EKF when $\theta = 2$

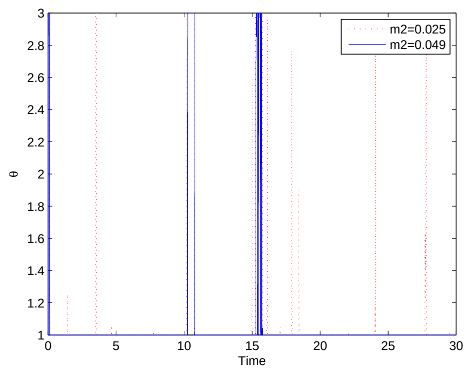

Fig. 5. Evolution of θ for different values of m_2

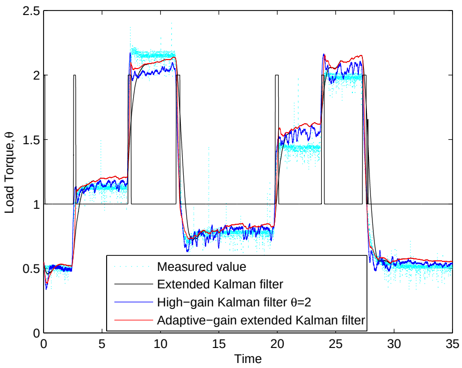

Fig. 6. Load torque estimation and evolution of θ

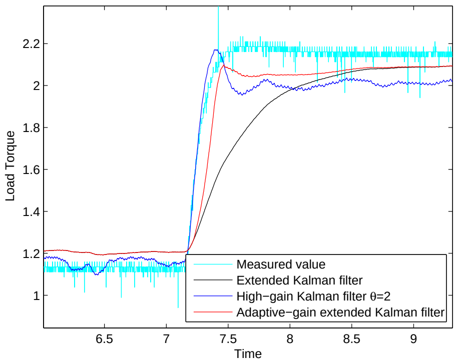

Fig. 7. Detailed view

and $T = 0.1$. From the output signal we estimated $\sigma = 0.7$ for the standard deviation , then $m_2 = \sigma^2 T = 0.049$. In order to obtain a more reactive EKF (i.e. when all the adaption coefficients were set to zero and $\theta(0) = 1$) Q was set to [25 0 0; 0 25 0; 0 0 50] .

Fig.2 shows the performance of the A-EKF compared to the one of an extended Kalman filter. We observe that when no unmodeled pertubations occur, the A-EKF ensures a good noise rejection just as the EKF does. On the other hand (and with a delay), the convergence of our observer is much faster than the EKF when a significant perturbation occurs. This delay is explained by the fact that θ only increases when the pertubation has an impact on the innovation. Therefore it is connected to the amplitude of the perturbation and to the sampling time choosen for the (discrete) computation of the innovation. Fig.3 compares the same performance with the one of a HG-EKF whose high-gain value was set to θ_{max}. As we expected the convergence speed of the A-EKF is comparable to the convergence of the high-gain one but with a delay caused by the computation of the innovation. Noise rejection is of course much more effective in the case of the adaptive-gain observer. Fig.4 shows what happens when we set θ in the HG-EKF to a smaller value in order to lessen the influence of noise on the estimation: even if there's a delay due to the adaption procedure, the A-EKF is more efficient in this situation. In Fig.5 we want to stress the importance of the parameter m_2 which represents the influence of noise on the innovation. The dotted curve shows that underestimation of this parameter leads to an increase of θ only due to the noise and not due to a large pertubation, and then that the observer switches to a high-gain mode that only amplifies noise. Still we think that the estimation of m_2 from the ouput signal can be done quite accuratly and easily.

F. *Estimation from real Data*

The next step in the developement of our observer is now to apply it in a real environment. Our system is composed of a DC motor from the German company LN and that can be connected in series, parallel, or in a compound maner. The brake (magnetic powder system) and its module were also provided by the same company. An I/O card connected to a desktop PC allowed us to get and set the values of the different quantities in play. The RTAI/linux system was used as our realtime environment. Even if our goal is to use our observer as a SISO system with the voltage as input and the current as output, we measured the speed and the load torque in order to compare those values with the estimated ones. Those measures were also used to estimate model parameters. In reality we found out that our model did not fit that well to the values obtained from the real motor and therefore we modified the first equation described in III-A with the addition of a second degree polynomial in ω_r. As we see below, we still observe a small bias in our estimations. We tuned the parameters of the observer according to the procedure described above.

Results of the estimation of the torque load from real data (in a batch mode) are shown in Fig.6 and Fig.7. All the data were obtained with our observer only by changing the values of the adaption function. As told before, the estimation is not that precise but still lies in an acceptable range. The surimposed curve in Fig.6 shows the evolution of θ with respect to time. We see that the high-gain parameter reacts as expected. We insist on the fact that m_2, the important parameter representing the influence of noise, has been set only from the measured output, namely the current, and has been kept to the very first value we choose: we did not tune it by trial and error. Fig.7 presents a detailed view of what happens around time 7 after a sudden change in the load torque. The perfomance of the A-EKK is the same as in the simulations.

A last point has to be stressed here: $\mathcal{I}$ is not likely to be equal to zero because of the errors which will always occur in the system modeling. This problem is solved by the addition of an integrator acting as a high-pass filter on the innovation.

IV. Conclusion

The main advantage of our adaptive extended Kalman filter is the proof of its convergence. Because of the θ parameter, matrices Q and R may be chosen to satisfy nice local performance, as usual for extended Kalman filtering. The parameter θ is sufficient to ensure the convergence when a large disturbance appears. Moreover, monitoring θ will give information on the system, and especially on the presence of disturbances, as seen in the application.

In order to validate our approach practically, we applied our observer on a real system. The application shows exactly what we expected from the proof: the A-EKF performs as an EKF when no perturbation occurs and performs as a high-gain observer when a perturbation occurs. Therefore, it is as robust regarding to noise as the extended Kalman filter, and converges as quickly as high-gain observers. Among other objectives, we now want to assess the perfomance of the A-EKF when it is implemented in a realtime environment.

REFERENCES

[1] *S. Mehta, J. Chiasson,* Nonlinear Control of a DC Motor: Theory and Experiment, *IEEE transactions on Industrial Electronics, Vol. 45, No. 1, Feb. 1998.*

[2] *P. C. Krause, O. Wasynczuk, S. D. Sudhoff,* Analysis of Electric Machinery and Drive Systems, 2nd edition, *IEEE Series on Power Engineering, Wiley-Interscience, 2002.*

[3] *J. S. Baras, A. Bensoussan, M. R. James,* Dynamic Observers as Asymptotic Limits of Recursive Filters: Special Cases, *SIAM J. Appl. Math.,* **48,** 1147–1158, *1988*

[4] *G. Besançon, J. de Leon-Morales, O. Huerta-Guevara* On adaptive observers for state affine systems, International Journal of Control, Vol. 79, No. 6, p 581-591, *2006*

[5] *M. Gevers, G. Bastin,* A Stable Adaptive Observer for a class of Nonlinear Second-order Systems, Analysis and optimization of systems, ed. by Bensoussan, A., and Lions, J. L. Springer-Verlag, June *1986*, pp 143-155

[6] *F.D. Busse, J.P. How, J. Simpson,* Demonstration of Adaptive Extended Kalman Filter for Low-earth-orbit Formation Estimation using CDGPS Navigation. Journal of the Institute of Navigation, v 50, n 2, p 79-93, *2003*

[7] *E. Busvelle, J.-P. Gauthier,* High-Gain and Non-High-Gain Observers for Nonlinear Systems, Contemporary Trends in Nonlinear Geometric Control Theory and its Applications, *World Scientific*, p. 233–256 *2002*

[8] *E. Busvelle, J.-P. Gauthier,* Observation and Identification Tools for Nonlinear systems. Application to a Fluid Catalytic Cracker. *International Journal of Control*, Vol. 78, 3, *2005*

[9] *E. Busvelle, J.-P. Gauthier,* On determining unknown functions in differential systems, with an application to biological reactors, *ESAIM: COCV* **9**, p. 509–552 *2003*

[10] *F. Deza, E. Busvelle, J.-P. Gauthier, D. Rakotopara,* High Gain Estimation for Nonlinear Systems, *Systems & Control Letters*, **18** *1992*

[11] *F. Deza, E. Busvelle, J.-P. Gauthier, D. Rakotopara,* A stability result on the continuous-continuous and continuous-discret extended Kalman filters, *Compte-Rendus de l'Académie des Sciences de Paris*, **314**, *Février 1992*

[12] *J.-P. Gauthier, H. Hammouri, S. Othman,* A simple observer for nonlinear systems, *IEEE Trans. Aut. Control,* **37**, pp. 875–880, *1992*

[13] *J.-P. Gauthier, I. Kupka,* Deterministic observation theory and applications, Cambridge University Press, *2001*

[14] *C. Hu, W. Chen, Y. Chen, D. Liu,* Adaptive Kalman filtering for vehicle navigation, Journal of Global Positioning Systems, Vol. 2, No 1, p42–47 *2003*

[15] *A. Johansson, A. Medvedev,* An observer for systems with nonlinear output map, Automatica 39, pp 909-918, *2003*

[16] *D. Jwo, H. Huang,* Neural network aided adaptive extended Kalman filtering approach for DGPS positioning, The journal of navigation, 57, p 449-463, *2004*

[17] *Kent K. C. Yu, N. R. Watson, J. Arrillaga,* An adaptive Kalman filter for dynamic harmonic state estimation and harmonic injection tracking, IEEE Trans. on power delivery, Vol. 20, No 2, *2005*

[18] *L. La Moyne , L. L. Porter II, K. M. Passino,* Genetic adaptive observers, Engng Applic. Artif. Intell. Vol. 8, No 3, p261-269, *1995*

[19] *L. Shen, H. Wang,* Adaptive observer design for general nonlinear systems with linear output structure, 4th International Conference on Control and Automation, 2003. ICCA '03. Proceedings, June *2003* On page(s): 48- 52

[20] *S. Stubberud, R. Lobbia, M. Owen,* An adaptive extended Kalman filter using artificial neural networks, The international journal on smart system design, Vol. 1, pp 207-221 *1998*